university of groningen

faculty of science and engineering

Feasibility study of the branching fraction measurements of $B_c^+ \to \tau^+ \nu_\tau$ and $B^+ \to \tau^+ \nu_\tau$ at LHCb

Jelte Rinus de Jong

university of groningen / faculty of science and engineering

University of Groningen

Feasibility study of the branching fraction measurements of $B_c^+ \to \tau^+ \nu_\tau$ and $B^+ \to \tau^+ \nu_\tau$ at LHCb

Master's thesis

Jelte Rinus de Jong (s3086496)

First examiners: dr. K.A.M. de Bruyn & dr. M.C. van Veghel
Second examiner: prof. dr. D. Boer

June 15, 2022

Contents

Chapter 1

Introduction

Curiosity is in human nature. We fantasized about the white dots in the night sky and found out that the world we see in relation to the Universe is like a grain of sand on the beach. At the beginning of the 20^{th} century, one started to realize that those grains of sand themselves contain an at least as interesting world on its own. The human imagination was put to the test, as intuitive patterns no longer suffice to explain Nature at these smaller scales. The Universe, a complex system of the large and small, has fascinated humanity - and still does – to explain the phenomena we see around us.

The ancient Greeks showed already interest in the building blocks of Nature [1]; Democritus believed there must exist some fundamental matter that is indivisible. It had to wait until the 20^{th} century for rapid progress to be made in the discovery of various fundamental particles. Our knowledge about these particles and their interactions are described by the Standard Model of particle physics (SM). This framework has proved successful in predicting the world of particle physics, but it is generally agreed that it contains certain deficiencies.

The SM only describes the ordinary matter (and anti-matter) particles and the interactions between them, except for gravity. These particles are divided into different types, or flavours. The SM predicts that the so-called charged leptons interact similarly with their corresponding interaction particles. This property is called lepton flavour universality (LFU). Various measurements hint towards deviations from the SM in this LFU, implying that this universality may be broken [2–5]. Other imperfections include the possible existence of additional types of matter [6]. A hypothetical candidate is that of *dark matter*, where dark refers to the fact that the candidates does not (or only very weakly) interact with ordinary matter. Physics beyond the Standard Model (BSM) is searched for by exploring higher energy scales using particle colliders.

The Large Hadron Collider (LHC) has been built to help answering some of the open questions. One of the detectors at the LHC, LHCb, is specifically designed to probe new physics (NP) using b and c quarks, heavy particles in the SM. In particular, a measurement on the branching fraction of $B_c^+ \to \tau^+ \nu_\tau$

and $B^+ \to \tau^+\nu_\tau$ provide an independent measurement on LFU, complementing previously mentioned studies. This thesis considers **the feasibility to measure the $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ branching fractions using the LHCb-detector**. These signal decays are chosen to be reconstructed using $\tau^+ \to \pi^+\pi^+\pi^-\nu_\tau$. The B_c^+ decay has not yet been experimentally observed; the branching fraction of the B^+ decay has been measured by Belle and Babar with a relative precision of $\sim 23\%$ [7, 8].

Measurements of these decay types at LHCb are challenging as they contain two neutrinos that can not be detected. One of the consequences is that the B_c^+ decay vertex can not be reconstructed; information of the ν_τ is missing. For that reason, the B_c^+-direction can not be inferred in the usual way. A novel B-tracking tool, measuring hits of the B_c^+ itself instead of its decay products, has shown to be worthwile to investigate further [9]. This thesis builds further on this work.

During the proton-proton (pp) collisions at LHCb, many different particles are produced besides the B_c^+, the decaying particle in our signal. These different particles can decay into similar final states as the signal decays, therefore contributing as background decay. A measurement is only feasible if the signal decay can be identified and distinguished from all these backgrounds. The search for potentially dangerous backgrounds is a significant part of this thesis, for which simulation frameworks are used. Then, to simplify the analysis, only the most dangerous background is considered. The worst-case scenario is assumed, where all pp-collisions lead to the particular background. This study aims to answer whether the experiment is feasible in this worst-case scenario. If not, conservative assumptions can be looked at more closely.

The outline of this thesis is as follows. The SM framework is introduced in chapter 2 where there is elaborated on how the LHCb detector and its subdetectors help to probe BSM physics. The steps that are taken in this study are discussed in more detail in chapter 3. The choice of the B_c^+ signal mode is explained as well and the simulation method is discussed. Chapter 4 discusses the selection procedure that is used to distinguish signal from background decays. In chapter 5, various background types are compared and the most dangerous background is determined. Chapter 6 and 7 build on determining the feasibility of the measurements by considering the most dangerous worst-case background. A discussion on the results is given in chapter 8, where also suggestions on further improvements are given.

Chapter 2

Probing the Standard Model using the Large Hadron Collider

2.1 The Standard Model of particle physics

The Standard Model of particle physics (SM) is a framework that describes the matter and antimatter that we observe in nature. It also explains how this matter interacts with each other. Matter particles are described by spin-$\frac{1}{2}$ particles, called fermions. Fermions consist of quarks and leptons, where leptons are further divided into charged leptons and neutrinos, the latter being neutral particles. Neutrinos only interact via the weak interaction; charged leptons interact via the weak and electromagnetic force; quarks interact via the strong, weak and electromagnetic interaction.

Associated to each type of interaction corresponds a force carrier, also called gauge bosons. The weak interaction is mediated by the massive $W^{\pm}$ and Z^0 bosons, which makes that it is the weakest of the above interactions. The electromagnetic interaction occurs by the mediation of a massless photon γ. The force carrier associated to the strong interaction is the massless gluon g.

Most of the SM particles are massive; only the gluon, photon and neutrinos[1] are massless. This mass originates from the interaction of the particles with the Higgs field. In the SM framework, particles are excitations of their associated fields. This also holds for the Higgs boson, which is the excitation of the Higgs field. The SM framework consisting of the matter particles, force carriers and Higgs boson is shown in figure 2.1.

Fermions are divided into three generations. Each generation contains an up- and down-type quark, a charged lepton and a neutrino, where each particle has an associated antiparticle. The strong interaction bounds quarks together, such that no isolated quarks are observed. Particles that are composed of multiple quarks are called *hadrons*. Protons (uud) and neutrons (udd) are examples of *baryons*, which are made up of three quarks. The B_c^+ ($\bar{b}c$) and B^+ ($\bar{b}u$) are

[1]The SM assumes massless neutrinos, however measurements on neutrino oscillations show that at least two have non-zero mass [10, 11].

examples of *mesons*, which are particles containing a quark and anti-quark.

Quark and charged leptons are more massive and less stable for higher generations. Neutrinos are assumed to be massless in the SM. The SM predicts that the charged fermions in each generation are treated identically, accounting for the mass difference and the associated coupling strength. In the case of leptons, this is called lepton flavour universality (LFU). As will be discussed in the following section, various measurements show hints to LFU violation, implying BSM physics.

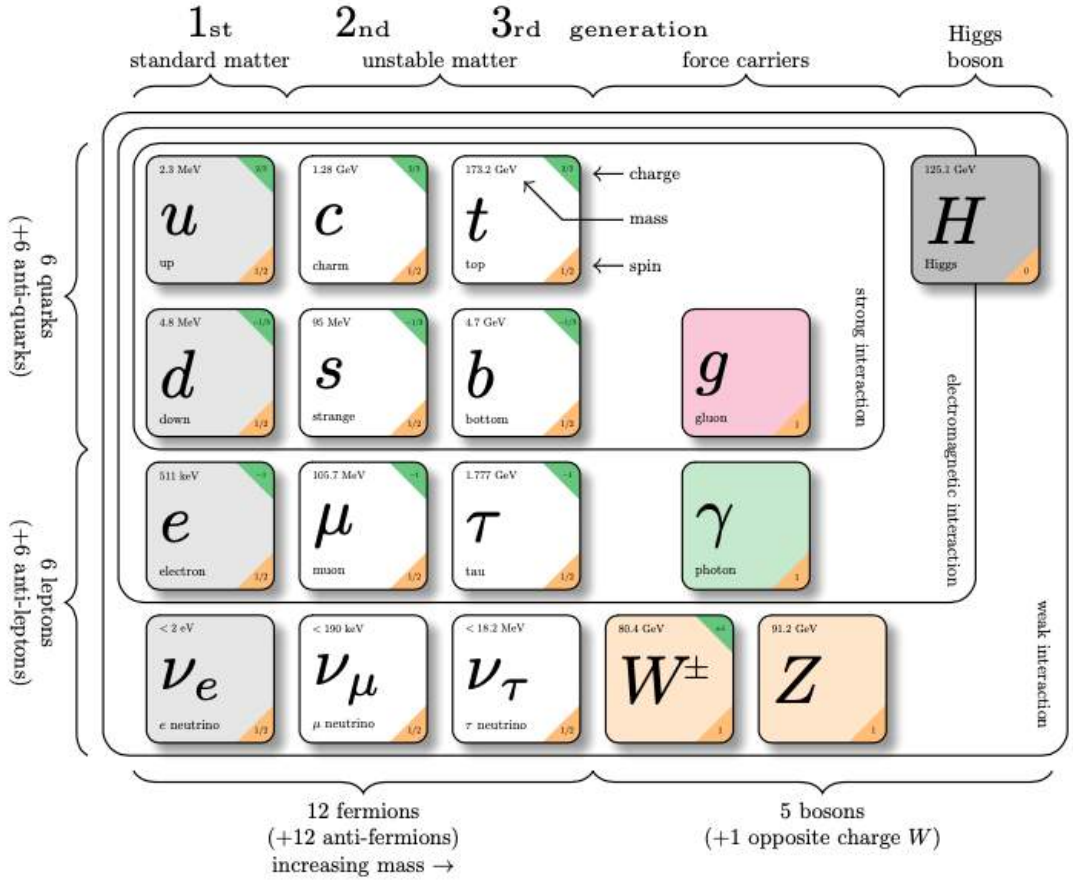

Figure 2.1: The particle content of the Standard Model, adapted from Ref. [12]. The matter particles consist of quarks and leptons and are divided into three generations of increasing mass. They interact via the exchange of gauge bosons. The possible interactions between matter particles and gauge bosons is indicated with regions. The Higgs boson gives mass to all particles it interacts with, leading to the gluon, photon and neutrinos remaining massless.

2.1.1 Lepton Flavour Universality Violation

One way to search for BSM physics is by probing LFU. Typical tests often comprise the comparison of the branching fractions of decays that only differ in their lepton flavour. A transition that can probe LFU is $b \to c\, l^+ \nu_l$, which

is involved in decays of beauty mesons into charm mesons. Signs of NP are observed using

$$\mathcal{R}(D^{(*)}) = \frac{\mathcal{BR}(\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau)}{\mathcal{BR}(\overline{B} \to D^{(*)}l^-\overline{\nu}_l)}$$

where the observed ratio is enhanced compared to the SM with 1.4σ and 2.8σ for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, respectively [4]. In particle physics, an observation is only considered to be a discovery if the result of a single measurement deviates at least 5σ from the SM prediction. Different BSM effects are proposed to explain the hints for the discrepancies in the observations. For example, branching fractions may be enhanced by NP effects of hypothetical particles like leptoquarks or additional types of Higgs bosons [13].

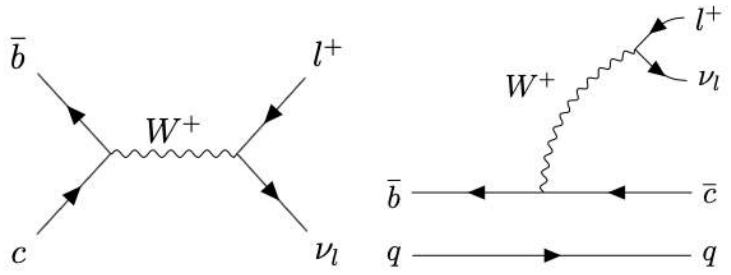

Figure 2.2: Feynman diagrams of the $bc \to l^+\nu_l$ (left) and $b \to c\,l^+\nu_l$ (right) transitions, where the spectator quark $q \in \{u, d, c\}$.

Probing LFU from a different perspective can help to further understand the patterns of BSM contributions. On the quark-level, the $b \to c\,l^+\nu_l$ transition is identical to $bc \to l^+\nu_l$, as is shown in figure 2.2. The latter transition effectively describes the decay of a B_c^+ meson. An experiment has already been done with the $B_c^+ \to J/\psi\, l^+\nu_l$ decay, where a 2σ deviation from the SM prediction has been observed [5]. A measurement on the branching fraction of $B_c^+ \to \tau^+\nu_\tau$ described by the same transition, may shed new light on the observations on LFU. Besides using $B_c^+ \to \tau^+\nu_\tau$, this study also considers $B^+ \to \tau^+\nu_\tau$ as well, which can be used to probe LFU using the $bu \to l^+\nu_l$ transition.

2.2 The Large Hadron Collider

In order to probe NP using $B_{(c)}^+ \to \tau^+\nu_\tau$, b-hadrons have to be produced. First of all, this can be done by using electron-positron colliders, where a center-of-mass (CoM) energy corresponding to the $\Upsilon(4S)$ mass is used [14]. This resonance dominantly decays into B^+ and B^0 mesons. Instead of the $\Upsilon(4S)$, a Z^0-resonance can be used. This is proposed in the feasibility studies focusing on the use of the future Circular Collider (FCC-ee) [15] and future Circular Electron Positron Collider (CEPC) [16]. Since the majority of the resonance decays into b-hadrons, electron-positron colliders suffer less from large backgrounds compared to hadron

colliders. Despite this fact, the production cross-sections of b-hadrons are much larger. Thus, if background contributions can be sufficiently rejected, this added benefit of generating more signal decays is useful. Currently, the Large Hadron Collider (LHC) is the most powerful hadron collider and was designed to achieve a CoM energy of 14 TeV [17]. It is located at CERN in Geneva, Switzerland.

The LHC consists of four interaction points, at which the opposing proton beams collide. The particles that are produced from the proton-collisions are measured by the detector that has been built around each interaction point. ATLAS and CMS were built to claim an observation or exclusion of the Higgs particle; they are general purpose detectors. ALICE was built to understand the Quark Gluon Plasma, a state of matter that is believed to have occurred at the early beginning of the Universe. Heavy ions are collided to reproduce this highly dense plasma. LHCb is specifically designed to use beauty hadrons to observe the matter-antimatter asymmetry (i.e. CP-violation) and search for rare-decays [18]. This section discusses the LHCb-detector during run 3, the third data-taking period of LHC.

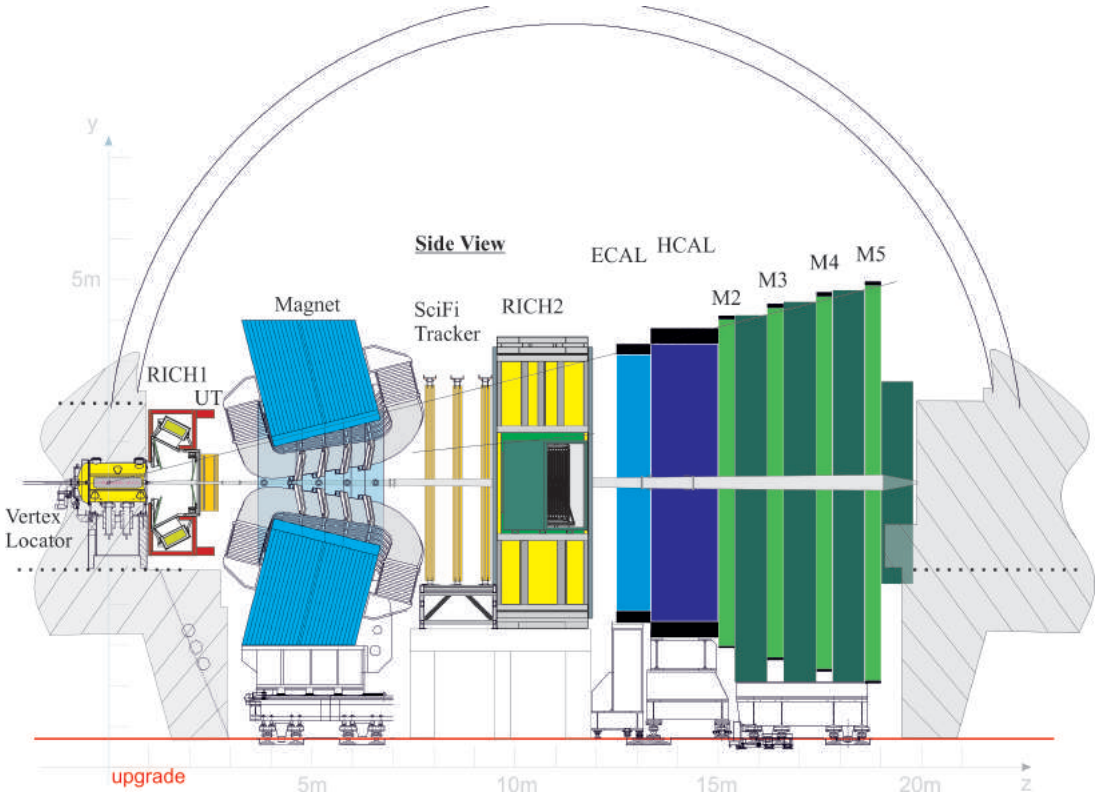

Figure 2.3: A schematic overview of the upgraded LHCb upgrade detector during run 3 [19]. Opposing proton beams travelling along the z-direction collide at the interaction point at (y,z) = (0,0). Particle tracking is done using the Vertex Locator (VELO) and tracking stations (UT, SciFi Tracker). Particle identification is done using the RICH1, RICH2, calorimeter (ECAL, HCAL) and muon stations (M2-M5).

2.2.1 The LHCb detector

The LHCb-detector is specifically built to search for signs of BSM physics using b- and c-hadrons. After being produced by the pp-collisions, these hadrons traverse the detector with small angles θ with respect to the proton beam axis. This explains why the detector is built as a forward spectrometer, as is shown in figure 2.3. It covers the pseudo-rapidity range of $2 < \eta < 5$, where $\eta = -\ln\tan(\theta/2)$. In order to be able to (partly) reconstruct a pp-collision, the detector consists of layers of sub-detectors. The sub-detectors are divided into tracking systems and detectors responsible for particle identification. Both are further discussed in the following sections, after which the trigger system - being responsible for real-time data selection - is introduced.

Tracking

The tracking system, which consists of the Vertex Locator (VELO) and tracking stations (UT, SciFi Tracker), is responsible for the reconstruction of the particle tracks. A particle track that is reconstructed using hits in the VELO, UT and SciFi stations, is called a *long track*. Most analyses use long tracks for signal reconstruction as they contain the most information. Charged particles are deflected by the magnet that is placed between the UT and SciFi tracking stations. The amount of deflection relates to the particle momentum, such that hits from the tracking stations can be used to infer the momenta of charged particles. The purpose of the VELO detector is especially relevant in this study and hence discussed in more detail.

The VELO detector is the detector that surrounds the interaction point, where the two proton bundles collide at each other. This is illustrated in figure 2.4, where the position of the VELO modules during run 3 is shown in the xz-plane. Opposing proton bundles travel along the z-direction and collide in the region around (x,z) = (0,0), illustrated by the grey dots. The bounds of the LHCb acceptance are denoted as well. The VELO consists of 52 modules in total, which are distributed according to figure 2.4. The distance between the module edges to the beam axis is 5.1mm [20].

The VELO is located close to the collision point of the two proton bundles to be able to accurately determine their interaction point, or primary vertex (PV). The main purpose of the VELO is to differentiate between this PV and the secondary decay vertices (SV). An example of the latter type is the 3π-vertex that occurs in the signal decays. Both types of vertices are identified by combining detector hits of charged particles into straight tracks and following them to their common origin. For b-hadrons produced by 13TeV proton collisions at LHCb, typical distances are given by their decay length; these are about one centimeter.

Usually, the decay vertex (DV) of the B-meson is used to infer the direction of the meson. For our particular signal decays, $B^+_{(c)} \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$, this is more involved. Namely, the LHCb detector is not able to measure both

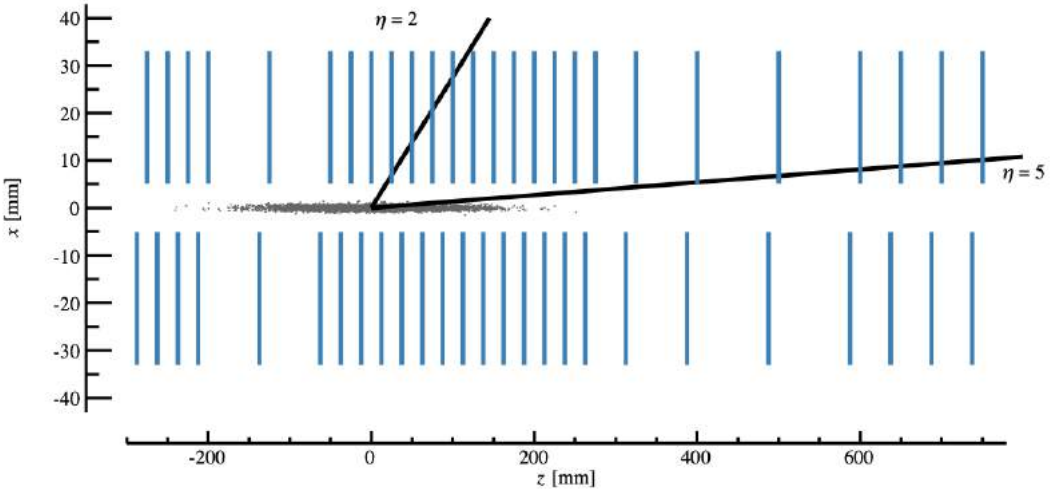

Figure 2.4: An illustration of the upgraded VELO-detector [21] in the xz-plane during run 3. Opposing proton beams are travelling in the z-direction and colliding in the region around the interaction point at (x,z)=(0,0), illustrated by the grey dots. The LHCb acceptance is shown as well.

neutrinos. Thus, the τ^+ is the only detectable particle that originates from the DV. The τ^+-direction can not be exactly determined due to the missing $\overline{\nu}_\tau$. In the end, this means that no information can be obtained on the exact location of the DV.

The B-tracking tool is introduced to determine the $B^+_{(c)}$-direction in a different way. Instead of inferring the direction using the DV, the first hit in the VELO-detector is assumed to come from the $B^+_{(c)}$-meson. This method is elaborated on in chapter 4 as it part of the selection procedure that happens during the data-taking of an experiment.

Particle identification

The remaining sub-detectors are used for particle identification. The two ring-imaging Cherenkov detectors (RICH1 and RICH2) are used to distinguish between different charged hadrons, in particular between kaons and pions. Using the Cherenkov angle measured by the RICH and the momentum measured by the tracking stations, the particle mass can be obtained which is used to identify the particle. The two RICH-detectors complement each other as they cover different momentum ranges.

The calorimeter measures the energy deposition of the particle to each layer. It is placed after the RICH2 detector. From front to back, it consists of the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). Besides the detection of charged particles, the calorimeter is the only sub-detector that is able to measure neutral particles. The ECAL measures the energy of elec-

tromagnetically interacting particles, so of photons, electrons and neutral pions. Then, the energy of hadrons is determined by the HCAL.

The majority of the particles that pass the calorimeter are muons, as they minimally interact with the detector material. For this purpose, the muon stations (M2-M5) are placed at the end of the detector. The few hadrons that have also passed the calorimeter are absorbed by iron layers that are placed between the muon stations.

Trigger

During run 3, the LHCb detector operates at an LHC collision rate of 30MHz and 6 proton-proton collisions per bunch crossing [22]. An enormous amount of data is collected in this way. However, many events contain information that is not relevant to the purpose of the measurement, i.e. many events only contain background decays. The purpose of the trigger is to select and save the useful events in real-time, or *online*. Signal-like decays are identified using a set of selection requirements. If an event meets the criteria, it is saved for the offline analysis. Events that do no meet the selection requirements will be lost. The selection procedure of the online selection (trigger) and offline selection for the our signal decays in particular is explained in more detail in chapter 4.

Chapter 3

Analysis strategy

The structure and frameworks of this feasibility study are introduced in this chapter. The argumentation of using the particular signal modes is discussed in section 3.1. Section 3.2 introduces the strategy that is followed after which the simulation framework is discussed in section 3.3.

3.1 Choice of signal modes

The feasibility of measuring the branching fractions will be studied for both the $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ decay. These decays are referred to as signals. The former is not yet observed so that the SM prediction of $\mathcal{BR}(B_c^+ \to \tau^+\nu_\tau) = 1.95\%$ is assumed [15]. The latter is observed to be $\mathcal{BR}(B^+ \to \tau^+\nu_\tau) = 1.09 \cdot 10^{-4}$ [23]. Charge conjugation of decay modes is always implied. Since the τ-lepton is unstable, both signal decays must be observed indirectly, via subsequent decay products. The τ-lepton can decay in many different modes, so a decision must be made as to which decay mode(s) the signals will be reconstructed through.

It is not desirable to use electrons and neutral pions for the reconstruction of τ^+. This is because the probability of reconstruction and selection for offline analysis are small compared to that of charged pions and muons. Despite its relatively small branching fraction, $\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau$ is useful because of two reasons [24,25]. Firstly, as three pions are produced, the tau decay vertex can be reconstructed more easily compared to the case where only a single track is involved. Without the knowledge of the position of this vertex, a lot of observables can not be derived as will be explained in section 4.3.

Secondly, the decay of $\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau$ mainly occurs via intermediate $a_1(1420)$ and $\rho(770)$ resonances [26], such that $\tau^+ \to a_1(\to \rho\,(\to \pi\pi)\,\pi)\,\overline{\nu}_\tau$. These resonances lead to additional kinematic structures that give an additional possibility to distinguish the signal pions from background, for example by using Dalitz plots [24]. As will be substantiated in section 3.3, resonances structures are not considered in this study.

This study assumes that the reconstruction of signals only occurs via $\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau$, where $\mathcal{BR}(\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau) = 9.31\%$ [23]. Other tau decay modes are not considered as their branching fractions are suppressed with respect to

the above mode.

3.2 Workflow

The purpose of this study is to demonstrate whether measurements of $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ are possible at LHCb. The various steps that are taken to achieve this goal are discussed here.

Data selection:
In an experiment, the data is filtered such that only potentially interesting events remain. As opposed to using experimental data, this study uses data that is generated by the RapidSim simulation package, discussed in section 3.3. The filtering of the data is done by imposing certain selection requirements. A basic introduction on the selection procedure is given in chapter 4. The amount of data that passes the selection is assumed to depend on the B-tracking, reconstruction and misidentification efficiency. Estimates of the latter two efficiencies are given in section 4.2.2 and 4.2.3, respectively.

Yields and shapes for background identification:
The potentially most dangerous backgrounds are determined by estimating their yields, using the discussion in section 5.1. For the backgrounds with the largest yields, we look further at their corrected mass distributions. An extensive number of backgrounds is discussed in chapter 5, where both the yields and shapes are compared. Section 5.6 concludes on the background that is considered to be the most dangerous.

Defining the worst-case scenario:
In order to determine the feasibility of measuring the signal modes, a worst-case but simplified scenario is considered. This scenario only considers the occurrence of the most dangerous background. On top of that, it assumes that all produced b-hadrons lead to the specific background. To quantify this, its branching fraction is assumed to be 100%.

Multivariate analysis:
An important part in the selection procedure discussed in chapter 4 is the multivariate analysis (MVA). Therefore, it is discussed separately in chapter 6. The MVA is used to optimally separate signal from background events. The distinguishing features are obtained from differences in their observables. These observables are discussed in section 4.3. The MVA output together with the corrected mass are used as input for a two-dimensional likelihood fit.

Likelihood fit in the worst-case scenario:
Various likelihood fits are done in chapter 7, where the worst-case scenario is considered. In the end, the fit performances determine the feasibility to mea-

sure the signal modes. To understand the relation between the reconstruction efficiency and the feasibility of the measurements, likelihood fits are done for different reconstruction efficiencies.

3.3 Simulation framework

The RapidSim simulation package [27] is used to understand the kinematic distributions of the signals and backgrounds. The software is very time-efficient due to the fact that the complete detector response is not simulated. Also, multiple assumptions are imposed on the spectra of the decay modes. In this way, potentially dangerous backgrounds can be explored as a first approximation. This makes that it perfectly serves the purpose of a feasibility study. The backgrounds that turn out to be most problematic can be looked at in more detail in further studies, using the full LHCb simulation software.

Variables of interest are defined in a *configuration* file and are generated for a particular decay defined in the *decay* file. Distributions are generated such that the parent particle is generated within a pseudo-rapidity range of $1 < \eta < 6$ and a momentum range of $0 < p < 100$ GeV/c. Settings that are decay-specific are defined in the *configuration* file. For example, the neutrinos that remain undetected and other particles that are not reconstructed are defined to be *invisible*. Also, an option *smear* can be applied on a (pseudo-stable) particle, which mimics the finite detector resolution. Namely, particle tracks cannot be reconstructed with infinite precision.

Simulated data is generated based on pre-defined distributions on for example momentum of the B-meson. The RapidSim package does not provide information on resonances that may be involved in certain decays. Instead, it is only able to generate decays based on the phase-space. Therefore, including resonances structures is beyond the scope of this project. To obtain more complete datasets in further research, RapidSim may be extended with additional simulation models like TAUOLA [28] and EvtGen [29].

Chapter 4

Selection

This chapter discusses the selection of signal-like decays and the assumptions that arise from this. The trigger is responsible for the data selection that takes place during the data taking (online) of the experiment. This is further discussed in section 4.1. Events that satisfy a list of requirements are saved for the offline analysis. The offline selection, where backgrounds are further separated from signals, is discussed in section 4.2. The observables that are used to distinguish signal from background decays are discussed in section 4.3.

4.1 Online selection

The purpose of the trigger is to select signal-like events for further analysis. The neutrinos that are produced in the signal decays $B^+_{(c)} \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ can not be detected by LHCb. Therefore, the only particles that can be used for the identification of the signals are the three charged pions in the final state. On top of that, a selection requirement is set on the number of hits that are left in the VELO-detector by the $B^+_{(c)}$-meson. These are discussed below.

During the online data selection, the trigger system checks if an event contains at least three charged pions that have similar properties as the signal pions (i.e. signal-like). For example, there is being searched for a 3π combination that is compatible with originating from a τ^+-lepton. The trigger line only checks if a 3π signal-candidate is present in the event. Thus, there is not searched for additional particles that correspond to the same decay.

For all events that contain at least three signal-like pions, the trigger imposes an additional requirement. Namely, the B-candidate that corresponds to the same decay-chain should leave at least a single hit in the VELO-detector (referred to as VELO-hits). This requirement is used as the direction of the B-meson is derived from the first VELO-hit. Therefore, without VELO-hits, the direction of the B-meson is unknown. During the data-taking, VELO-hits are searched for using the B-tracking tool. This study will mimic this tool using some approximations that are discussed in section 4.3.

As an example, let's look at a potential background decay of the form $B^+ \to$

$\pi^+\pi^+\pi^- X$, where X denotes all additional (pseudo-stable) particles that are produced in the decay. Decays of this type pass the trigger if the three charged pions obey the selection requirements. The additional particles are detected by LHCb, however they are not searched for. Hence, the trigger does not consider them in the selection. For all decays of this form, that contain three signal-like pions, the trigger system checks if the corresponding B-candidate has left a VELO-hit. If this is the case, the particular decay event is saved for offline analysis.

B-tracking efficiency

As was explained in section 2.2.1, the distance to the VELO-modules is quite significant compared to the average decay length of a 13 TeV B-meson. Therefore, only a fraction of the B-candidates leave VELO-hits. This fraction is quantified as the B-tracking efficiency $\epsilon_{tracking}$. It is defined as the fraction of decays having at least a single VELO-hit (from the B-candidate) compared to the total number of decays. This efficiency depends on the specific decay kinematics, as the particle momenta and relative angles between the particles lead to different particle trajectories. In turn, the momentum and angular distribution of a decay is related to the mass of the B-candidate. Therefore, the B-tracking efficiency is determined for each individual decay mode using the output results of RapidSim simulations.

As mentioned before, the VELO detector is only able to detect charged particles. Therefore, any neutral particle involved in a decay does not result in VELO-hits. Some background modes may have neutral mother particles (e.g. B^0) but are still identified as B-candidate. The B^0 itself does not leave any VELO-hits, but decay products (e.g. τ^+) may be able to leave VELO-hits. It may seem that the VELO-hit from the τ^+ comes from the B-candidate. In this way, a decay that has a neutral mother particle but has passed the trigger may lead to a background contribution. In general, these decays lead to smaller B-tracking efficiencies as the number of reconstructable particles in the decay is smaller. For backgrounds, a small B-tracking efficiency is advantageous as it results in less tracks being reconstructable. Signals benefit from a larger B-tracking efficiency as this leads to larger signal yields.

4.2 Offline selection

To repeat, an event is saved by the trigger if it contains at least three signal-like pions. Additionally, the corresponding B-candidate should have left at least a single VELO-hit. As such, the trigger has rejected a significant fraction of the background already. However, there are still backgrounds that have passed the selection requirements. These are further rejected during the offline analysis.

The backgrounds that have passed the trigger can be divided into various types. First of all, there are backgrounds where the B-candidate decays to exactly three charged pions. This background can be reduced by using mass vetoes,

as will be explained in section 4.2.1. Then, there are backgrounds where at least three pions are produced from the B-candidate. As only the three signal-like pions are reconstructed, these are referred to as *partially reconstructed* backgrounds. They can be reduced by using isolation variables, which are discussed in section 4.2.2. Backgrounds where one (or more) of the signal pions are misidentified are called *misidentified* backgrounds. They are discussed in section 4.2.3. Combinations of the above types may occur as well. An additional background type is the *combinatorial background*, where the signal-like pions are unrelated but are reconstructed as if all three come from the same vertex. The focus of this study lies on the partially reconstructed and misidentified backgrounds; the combinatorial background is not considered.

As will become clear later on, partially reconstructed and misidentified backgrounds are the most difficult to reject. For these backgrounds, a multivariate analysis (MVA) can help in the rejection. As the use of the MVA is an important part of this thesis, chapter 6 is particularly dedicated to this.

Each background type comes with an associated efficiency. For the mass veto and isolation method, this is a measure of the background fraction that remains after applying the tool. For the misidentified background category, the efficiency is a measure of the probability of misidentification. The following sections not only discuss the tools used for background rejection, but the associated efficiencies are estimated as well.

4.2.1 Mass vetoes

Decays that include an intermediate particle decaying to exactly three pions can be rejected by using mass vetoes. As information of all particles in the final state is available, the invariant mass spectrum of the 3π can be precisely measured. It sharply peaks at the mass of their mother particle, such that a cut-based selection is able to reject this background. For that reason, these backgrounds are assumed to be entirely eliminated and are not considered in this study.

4.2.2 Background reduction by isolation

Partially reconstructed decays, where only three signal-like pions are reconstructed, is one of the background types that may be passed by the trigger line. All additional particles are not reconstructed. These B-meson decays typically have large boosts. As large boosts lead to smaller opening angles, it is assumed that all particles stay within LHCb's geometrical acceptance of $2 < \eta < 5$. This comes down to assuming that all particles are reconstructable, meaning that all additional particles leave hits in the detector which can be used for track reconstruction.

To discard a partially reconstructed decay as a signal decay, it is sufficient to observe a single additional particle X in the vicinity of the three pions. These additional particles are searched for by so-called isolation variables. They are used to identify whether additional tracks are present around the three pions

and are corresponding to the same decay.

In general, LHCb does not reconstruct 100% of all tracks produced in the pp collision. The variable that describes the probability for a particle to be reconstructed by LHCb is the reconstruction efficiency ϵ_{rec}. Only in the cases where X can be reconstructed, the isolation variables can be used to identify the background decay. In the remaining cases, the background seems to be signal-like as in that case only the 3π is reconstructed. Effectively, if the decay involves a single additional particle, this means that isolation variables reduce the background yield with a factor ϵ_{rec}. The values that are assumed if multiple additional particles are produced together are discussed in the following section. Also, estimates on the reconstruction efficiencies are given.

Reconstruction efficiency

The reconstruction efficiency is defined for a single particle, where the probability of reconstruction depends on its charge. Namely, the only sub-detector that is able to measure neutral particles is the calorimeter, whilst charged particles can be detected by multiple sub-detectors. Therefore, it is expected that charged particles are more probable to be detected (with an efficiency of ϵ^{ch}_{rec}) than neutral particles (with an efficiency of ϵ^{n}_{rec}). As the reconstruction efficiency is linked to the use of isolation variables, the reconstruction efficiency of charged (neutral) particles is interchangeably referred to as charged (neutral) isolation.

The reconstruction efficiency for charged particles ϵ^{ch}_{rec} is experimentally obtained by using the long tracks of the $J/\psi \to \mu^+\mu^-$ decay. One of the muons together with the J/ψ particle is fully reconstructed; the other muon only partially. The reconstruction efficiency of the latter muon is found by searching for the corresponding fully reconstructed track. Although muons are used, the obtained efficiency holds for every type of charged particle. This is because there is only focused on track reconstruction rather than on particle identification. The average charged particle reconstruction efficiency is found to have a lower bound of 96% for the LHCb phase space [30], i.e. for momentum $5 < p < 200$ GeV/c and pseudo-rapidity $2 < \eta < 5$. This is shown in figure 4.1. In this study, a conservative value of 95% is used.

The reconstruction of neutral particles is more difficult as only the calorimeter is able to measure properties of neutral particles. As an example, neutrals pions are reconstructed by searching for two photon-like particles. The photons may subsequently decay into an electron-positron pair. Measurements on the reconstruction efficiency for neutral particles have not been done in general; only values are obtained for specific analyses [31, 32]. In the context of neutral isolation, it is only relevant to measure energy deposition coming from other particles than the signal pions. However, it is not necessary to identify the particles that were responsible for the extra energy deposition. As the studies above have obtained values for ϵ^{n}_{rec} where such effects are taken into account, they are very conservative for the purpose of this study.

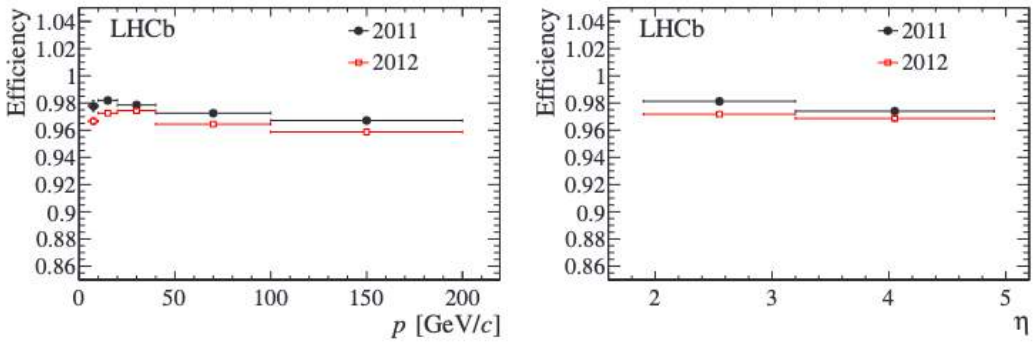

Figure 4.1: The reconstruction efficiency for charged particles as function of momentum p and pseudo-rapidity η [30]. The average efficiency over the LHCb phase space is above 96%. In this study, a conservative value of 95% is used.

Thus, the reconstruction efficiency for neutral particles is assumed to be $\epsilon^n_{rec} \sim 90\%$. As the purpose of this calculation is only to estimate the order of magnitude of the background yield, the exact values of the efficiencies does not affect the argumentation significantly. Also, the efficiencies mentioned above are averages. Note that in reality, the reconstruction efficiency depends on the pseudo-rapidity and momentum of the mother particle, as was also shown in figure 4.1. Thus, the reconstruction efficiency depends on the specific particle as well as on the kinematics of the decay. These dependencies are not taken into account, as here only order of magnitudes are of interest.

As mentioned above, the reconstruction efficiency is defined for a single particle. However, in decays like $B^+ \to \pi^+\pi^+\pi^- X$, X generally is a combination of (not reconstructed) charged and neutral particles, rather than a single particle. If X consists of multiple particles, the effective reconstruction efficiency becomes larger as the probability of reconstructing no additional particle decreases. Therefore, to stay conservative, $\epsilon_{rec} = \epsilon^{ch}_{rec}$ is assumed when X only consists of charged particles; $\epsilon_{rec} = \epsilon^{n}_{rec}$ is assumed when X only consists of neutral particles. As it is expected that $\epsilon^n_{rec} < \epsilon^{ch}_{rec}$, the conservative estimate of $\epsilon_{rec} = \epsilon^n_{rec}$ is assumed if X is a combination of charged and neutral particles.

4.2.3 Particle misidentification

Backgrounds where at least a single particle is misidentified as a signal-like pion, may have passed the trigger as a misidentified background. For example, a charged kaon may be misidentified as a charged pion, such that a decay like $B^+_c \to (\tau^+ \to K^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ is reconstructed as $B^+_c \to (\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$. As three signal-like pions are reconstructed, the trigger selects the decay for the offline analysis. However, as one of the pions is misidentified, this so-called *misidentified* background contributes to the background yield.

The probability that a particle X is wrongly identified as a different particle Y is described by the misidentification efficiency $\epsilon^{X\to Y}_{mis}$. It is a measure of how

often a certain particle and thus a certain decay is expected to be misidentified as a signal decay. In this way, $\epsilon_{mis}^{X\to Y}$ can be used to estimate the background yield of a certain decay due to misidentification. Estimates on $\epsilon_{mis}^{X\to Y}$ are given in the following section.

Misidentification efficiency

Experiments have been done to determine the probability of identifying a kaon as a kaon ($\epsilon^{K\to K}$) and the misidentification efficiency of the pion to kaon ($\epsilon_{mis}^{\pi\to K}$) [30]. The results are shown in figure 4.2. Two selection requirements are imposed to obtain good performance for either $\epsilon^{K\to K}$ and $\epsilon_{mis}^{\pi\to K}$. The former selection requirement results in an average misidentification efficiency of approximately 95% over the momentum range between 2 and 100 GeV/c. The fraction of kaon to pion misidentification does not imply an equal fraction for pion to kaon misidentification. In other words, $\epsilon_{mis}^{K\to\pi}$ can not be easily derived from $\epsilon_{mis}^{\pi\to K}$. This is because besides the pion, the kaon can be misidentified as different particles as well (e.g. as electron, muon or proton). The assumption can be made that the kaon is mostly misidentified as pion, such that $\epsilon_{mis}^{K\to\pi} \sim 1 - \epsilon^{K\to K} \sim 5\%$. As a conservative estimate, the same efficiency is assumed for electron and muon to pion misidentification.

The reverse is also possible, where pions are misidentified as other particles ($\epsilon_{mis}^{\pi\to X}$). Taking into account these effects results in lower yields for both signal and background modes. Namely, three signal-pions will be reconstructed less often. A reduction in the background yield is advantageous, whereas a reduction in signal yield has the opposite effect. Including $\epsilon_{mis}^{\pi\to X}$ thus leads to a less conservative estimate and is therefore not considered.

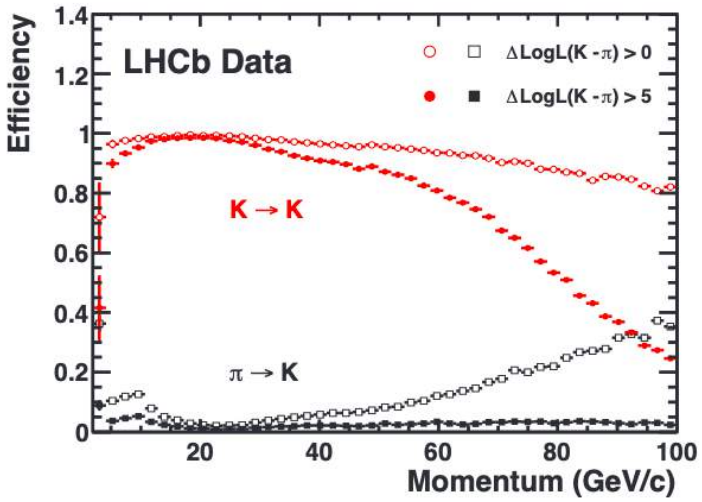

Figure 4.2: Experimental data showing the kaon identification efficiency and the pion misidentification efficiency as function of momentum [30]. Two selection requirements are chosen to optimise either efficiency.

4.3 Observables for signal identification

All background that has passed the trigger can be further reduced using an MVA, where signal and background decays are distinguished based on differences in their decays. These differences become apparent in their observables. The presented observables should not be regarded as a complete list. Instead, this section only discusses the most important observables that are used in the MVA. Chapter 6 elaborates on the MVA itself.

The signal decays $B^+_{(c)} \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ involve the production of a neutrino at both the $B^+_{(c)}$ decay vertex and the 3π-vertex. These decays are shown schematically in figure 4.3. Because neutrinos are not detected by LHCb's detector systems, they can not be used for acquiring observables. Therefore, the identification of the decay fully relies on the three charged pions that are reconstructed.

First of all, the invariant mass of the pions and their corresponding momentum can be determined. The secondary vertex (SV) is found by combining the charged pion tracks. The true directions of the τ^+- and 3π slightly differ because of the production of the $\overline{\nu}_\tau$-neutrino. As no information on this neutrino is obtained, it is assumed that the τ^+-direction is equal to the 3π-direction.

Furthermore, the $B^+_{(c)}$-direction is assumed to be the direction between the primary vertex (PV) and the first VELO-hit. Using the PV, SV, the first VELO-hit and additional information from the pions, the opening angle θ_{corr}, the impact parameter (IP), the flight distance (FD) and the number of VELO-hits are determined. They are discussed in the following sections.

4.3.1 Mass variables

The masses of the three charged pions can be used to identify their mother particle, so the particle that has decayed into the three pion combination. The invariant mass of this combination is given by

$$m_{3\pi} = \sqrt{E^2_{3\pi} - \vec{p}^2_{3\pi}} \tag{4.1}$$

where $E_{3\pi}$ and $\vec{p}_{3\pi}$ are the energy and momentum of the three pions and $m_{3\pi} \in [3m_\pi, m_\tau]$. For our signal, it corresponds to the invariant mass of the a_1-resonance. In the current study only decays according to phase-space are considered, rather than via intermediate resonances. Therefore, at this point, the invariant mass is not particularly meaningful. Instead, other variables have been invented that aim to retrieve some of the information that is lost through the neutrino. An alternative is the corrected mass defined as

$$m_{corr} = \sqrt{m^2_{3\pi} + |\vec{p}_\perp(3\pi)|^2} + |\vec{p}_\perp(3\pi)| \tag{4.2}$$

where $\vec{p}_\perp(3\pi)$ is the momentum of the 3π combination perpendicular to the B-meson flight direction. If not a neutrino but a massive (not reconstructed) particle is produced besides the three pions, $p_\perp(3\pi)$ and therefore m_{corr} differs. In

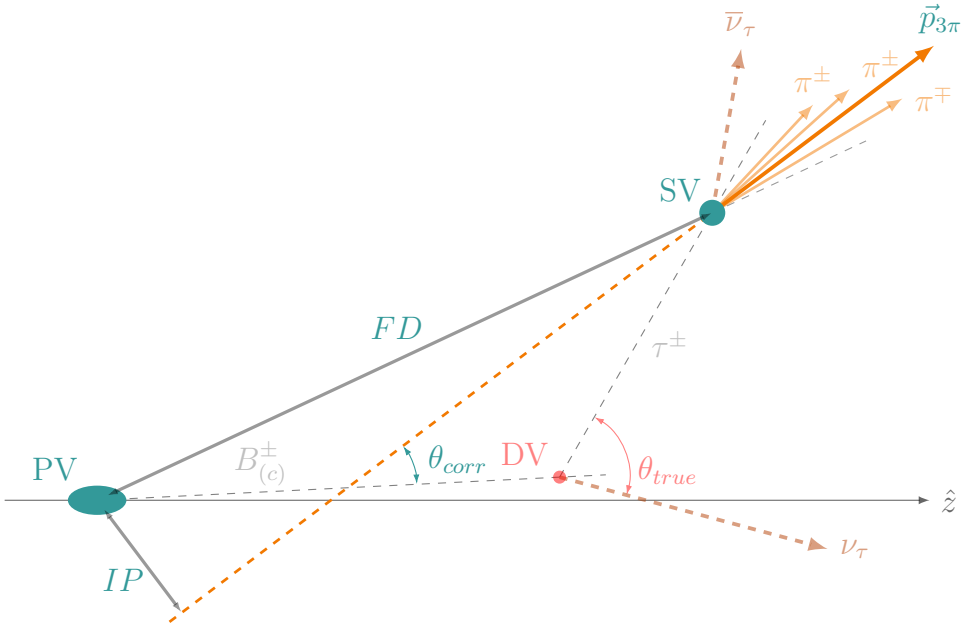

Figure 4.3: A schematic picture showing the observables that are used in this study. Observables that can be inferred experimentally are the PV, SV, FD, IP, $\vec{p}_{3\pi}$ and θ_{corr} (green). The B decay vertex (DV) and θ_{true} (red) can not be determined from experiment.

this way, the corrected mass is able to distinguish between signal and background decays.

4.3.2 Opening angle

The true opening angle θ_{true} as indicated in figure 4.3, can not be determined experimentally because the B decay vertex (DV) is unknown. Instead, an angle that can be inferred experimentally is the corrected angle θ_{corr}. It is the angle between the $B^+_{(c)}$ and 3π flight directions. As the boosted particles in typical B-meson decays have small opening angles, θ_{corr} is a good approximation to θ_{true}.

4.3.3 Impact parameter and flight distance

The IP is defined as the distance of closest approach of a certain track. This variable is especially effective in identifying particles that did not originate from the PV, but from a relatively long-lived intermediate particle. The IP distribution is strongly affected by the lifetime of this particle. As the B_c^+-lifetime is approximately three times shorter than the B^+ lifetime, the IP distribution of the $B_c^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ and $B^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ decays look very different.

Another observable that depends on the $B^+_{(c)}$ and τ^+ lifetimes is the flight distance (FD). It is defined as the distance between the PV and SV. Just as for

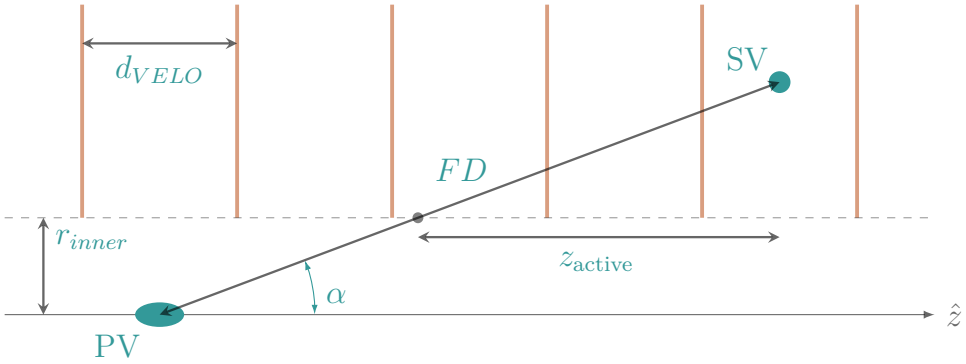

Figure 4.4: A representation of a few VELO modules (orange) together with the variables that are used to estimate the number of VELO-hits, mimicking the B-tracking tool.

the IP, this variable is helpful in distinguishing decays containing (or lacking) long-lived intermediate particles.

4.3.4 Number of VELO-hits

As mentioned before, the B-tracking tool is used in an experiment to search for VELO-hits by the B-candidate. As this study only considers simulated data that does not include the actual B-tracking tool, an approximation to the tool is used. This estimation is explained in this section.

To determine the number of VELO-hits, the VELO-detector as discussed in section 2.2.1 is approximated as shown in figure 4.4. The PV, SV and FD are equally depicted as in figure 4.3. Note that the line between the PV and SV does not represent the particle trajectory, so VELO-hits do not come from the intersection of this line with the VELO-modules. The angle α is defined as the angle between the z-direction and PV-SV-direction.

The inner radius of the VELO detector $r_{inner} = 5.1\text{mm}$; the distance between all subsequent VELO modules is assumed to be $d_{VELO} = 3\text{cm}$[1]. The active distance z_{active} is defined as the distance along the z-direction where the particle can leave VELO-hits. It is given by $z_{active} = \text{FD}\cos(\alpha) - r_{inner}\cot(\alpha)$. The number of VELO hits (N_{VELO}) is then determined using

$$N_{VELO} = \left\lfloor \frac{z_{active}}{d_{VELO}} + \delta \right\rfloor \tag{4.3}$$

where the random number $\delta \in [0, 1]$ is included to account for the PV that is uniformly distributed between two VELO-modules.

[1]This study assumes the same d_{VELO} as in Ref. [9]. The VELO-detector used in run 3 has $d_{VELO} = 25\text{mm}$ for most modules [21].

Chapter 5

Backgrounds

Chapter 4 explained that a selection procedure is done to keep as much as possible events containing signal-like decays. Still, there are different backgrounds that pass the selection requirements. The aim of this chapter is to determine the potentially dangerous backgrounds based on their estimates yield and corrected mass distribution. Section 5.1 starts by discussing how to estimate each of these background yields. Section 5.2 explains what are the potential background candidates, regarding our signals. General assumptions on these backgrounds are summarized in section 5.3 after which charm and beauty backgrounds are respectively discussed in section 5.4 and 5.5. To simplify the analysis, only one background is selected in section 5.6, which is assumed to be the most dangerous background in the remaining part of the thesis.

5.1 Estimation of the yields

The expected yield $N_{M\to Q}$ for the signal and background decay of the general form $M \to Q$ is determined by

$$N_{M\to Q} = N_M \cdot \epsilon_{M\to Q} \cdot \mathcal{BR}(M \to Q) \tag{5.1}$$

where $\mathcal{BR}(M \to Q)$ is its branching fraction and $\epsilon_{M\to Q}$ is the corresponding efficiency. In this context, it is interpreted as the probability for the decay to pass the selection requirements up to the MVA. N_M is the number of produced mesons of type M in a pp-collision and depends on the luminosity L and cross-section $\sigma(M)$ as $N_M = L \cdot \sigma(M)$. To easily compare the signal and background channels, the yields are determined assuming $L = 1\,\text{fb}^{-1}$; they are listed in in table A.4 of appendix A. $\sigma(B^+)$ is assumed to be 87 µb [33]. $\sigma(B_c^+)$ is determined to be 0.65 µb by accounting for the ratio of b-hadron production fractions $f_c/f_u = 7.56 \cdot 10^{-3}$ [34]. For backgrounds involving charm-meson decays, $\sigma(D^+)$ and $\sigma(D_s^+)$ are assumed to be 834 µb and 353 µb, respectively [35].

For signal decays, the efficiency $\epsilon_{M\to Q} = \epsilon_{tracking}$ as there is selected on the B-candidates having at least a single VELO-hit. For backgrounds, the overall efficiency includes the additional factors mentioned in section 4.2, such that

$$\epsilon_{M\to Q} = \epsilon_{tracking} \cdot (1 - \epsilon_{rec}) \tag{5.2}$$

where it is implied that all efficiencies correspond to $M \to Q$. It was assumed that in $1 - \epsilon_{rec}$ of the cases, additional particles are not reconstructed, so that these events contribute to the background yield. If the specific background involves the misidentification of at least one of the signal-like pions, an extra factor of ϵ_{mis} is applied as well.

5.2 Identification of potential backgrounds

As mentioned previously, the trigger identifies signal-like decays based on the three pions and requires that the B-meson candidate has at least a single VELO-hit. It was explained that backgrounds that have passed the trigger can be divided into three categories, namely

- **partially reconstructed**, where additional particles are not reconstructed and therefore cannot be eliminated through the use of isolation variables, in $1 - \epsilon_{rec}$ of the cases.
- **misidentified**, where a particle is identified as a different particle in ϵ_{mis} of the cases.
- **combinatorial**, where a random combination of signal-like pions is reconstructed as coming from the same vertex.

Also decays that are a combination of the above may be selected by the trigger, however they become suppressed quickly as different efficiency factors become involved. The focus of this study lies on the partially reconstructed and misidentified backgrounds; the combinatorial background is not considered. In the following, additional particles are denoted by X and not reconstructed, unless indicated otherwise.

Decays are potential backgrounds if the $\pi^+\pi^+\pi^-$ (or 3π) fall in a similar mass window as the signal decays, which is considered to be $0 < m_{corr} < 10$ GeV/c^2. Especially pions coming from $D^+_{(s)}$-mesons are difficult to distinguish from tau-leptons due to the large similarity in mass and lifetime [23]. Therefore, a B-candidate producing a $D^+_{(s)}$-meson instead of τ^+ is expected to have a significant background contribution.

5.3 General assumptions

Various assumptions on the backgrounds are made in order to simplify the analysis where possible. Approximations that are background-specific are discussed in the corresponding sections in 5.4 and 5.5. Assumptions that are applicable to each background can be summarized as:

- Production and decay vertices are well distinguished.

- For decays of the form $B \to DX$ where $D \in \{D_s^+, D^+, \overline{D}^0\}$ decays into at least three signal-like pions, only the largest contribution of the $D \to \pi^+\pi^+\pi^- X$ decay is considered.

- The (inclusive) charm excited state $\overline{D}^{**0}$ is assumed to be $\overline{D}_2^*(2460)^0$ and decays via $\overline{D}_2^*(2460)^0 \to D^-\pi^+$.

Every assumption is elaborated upon in the subsections below in the same order.

5.3.1 Distinguishability of vertices

Let's consider the example in which $B^+ \to D^-(\to \pi_1^-\pi_2^-\pi_3^+ X)\,\pi_4^+\pi_5^+$, where the subscripts are added for explanatory purposes. This decay passes the trigger if at least three signal-like pions are reconstructed.

A simple case is where $\pi_1^-\pi_2^-\pi_3^+$ are identified as signal-like. The decay passes the trigger if X and $\pi_4^+\pi_5^+$ are not reconstructed. Different configurations are possible as well. For example, $\pi_1^-\pi_2^-\pi_4^+$ may be identified as signal-like pions. The decay passes the trigger if X and $\pi_3^+\pi_5^+$ are not reconstructed. This configuration is different than the first one because in the second case the signal-like pions originate from different vertices. In an experiment, there is a finite resolution in which the production and decay vertex of the D^- can be reconstructed. Hence, decays of the second configuration contribute to background as well. However, here it is assumed that only decays of the first form contribute to background.

5.3.2 *D*-meson sub-decays

As will be shown, there are many possible background decays of the form $B \to DX$, where $D \in \{D_s^+, D^+, D^0\}$ decays into at least three signal-like pions. The largest contributions of these D-decays are shown in table A.1, A.2 and A.3 of appendix A, respectively. As a first approximation, only the largest contribution of every D-meson is assumed to occur. For D_s^+, D^+ and D^0, these are respectively

- $D_s^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\,\nu_\tau$
- $D^+ \to \pi^+\pi^+\pi^-\,K_S^0$
- $D^0 \to \pi^+\pi^+\pi^-\,K^-$

where only $\pi^+\pi^+\pi^-$ are reconstructed.

5.3.3 Charm excited states

Some B-meson decays will be considered where the charm excited state $\overline{D}^{**0}$ is involved. Firstly, there are decays where $\overline{D}^{**0}$ is not reconstructed. Secondly, the signal-pions are produced from the $\overline{D}^{**0}$-meson.

Decays containing this meson are difficult to simulate as $\overline{D}^{**0}$ is the notation for the sum of the $\overline{D}_1(2420)^0$, $\overline{D}_2^*(2460)^0$ and $\overline{D}_1(j = 1/2)^0$ L = 1 charm

meson states [36]. Neutral isolation is assumed for the decays where $\overline{D}^{**0}$ is not reconstructed as some of mesons above are allowed to decay into neutral particles.

For the decays where 3π is reconstructed as originating from $\overline{D}^{**0}$, $\overline{D}^{**0}$ is assumed to be the $\overline{D}_2^*(2460)^0$-meson state. This is because $B^+ \to \overline{D}^{**0}X$ occurs at a relative fraction of $\overline{D}_1(2420)^0 : \overline{D}_2^*(2460)^0 : \overline{D}_1(j = 1/2)^0 = 1 : 6.7 : 2$ [36]. $\overline{D}_2^*(2460)^0 \to D^-\pi^+$ and $\overline{D}_2^*(2460)^0 \to D^{*-}\pi^+$ are observed as possible decay modes [23]. As conservation of isospin and angular momentum predict that $\mathcal{B}(\overline{D}_2^*(2460)^0 \to D^{*-}\pi^+) = 1/5$, it is assumed that $\overline{D}_2^*(2460)^0 \to D^-\pi^+$ occurs mostly. Combining the above reasoning, $\overline{D}^{**0}$ is assumed to decay as

$$\overline{D}_2^*(2460)^0 \to D^-\pi^+$$

Note that this assumption is only a rough estimate and does not include meson states where $\overline{D}^{**0} = \overline{D}^{*0}n\pi$. If backgrounds involving $\overline{D}^{**0}$ turn out to be potentially dangerous, the assumptions need to be revised.

5.4 Charm backgrounds

Charm background candidates are decays where at least three signal-like pions are produced from a D-meson. The following charm contributions are considered:

- $D^+ \to \tau^+\nu_\tau$, $\tau^+ \to 3\pi\,\overline{\nu}_\tau$
- $D_s^+ \to \tau^+\nu_\tau$, $\tau^+ \to 3\pi\,\overline{\nu}_\tau$

Both charm backgrounds are shown in table A.4 of appendix A. Isolation variables do not reduce the background yield here, as the only reconstructable particles in the final are the signal-pions. For the D^+ and D_s^+ background, the expected yields per fb are determined to be $1.73 \cdot 10^5$ and $3.42 \cdot 10^5$, respectively.

The study on which this thesis continues [9] has shown that both charm contributions show narrow peaks in the range of $1 < m_{corr} < 2$ GeV/c^2. Moreover, these backgrounds were shown to be well distinguishable based on the opening angle and the corrected mass. The normalised corrected mass distributions of the backgrounds are compared to both signal distributions and shown in figure 5.1. The height of the bin denotes the number of candidates in that bin relative to the other bins. They are obtained by an independent RapidSim simulation. The background shapes show small overlap with the signal shapes and can thus be well distinguished using m_{corr}. Therefore, charm backgrounds are considered to be not dangerous with respect to their mass shapes. However, their expected yields are large compared to the signal yields.

5.5 Beauty backgrounds

Beauty background candidates are decays where a B-meson produces at least three signal-pions. In particular, the decay of a B^+-meson is considered. Different types of B^+-decays are divided into the following categories:

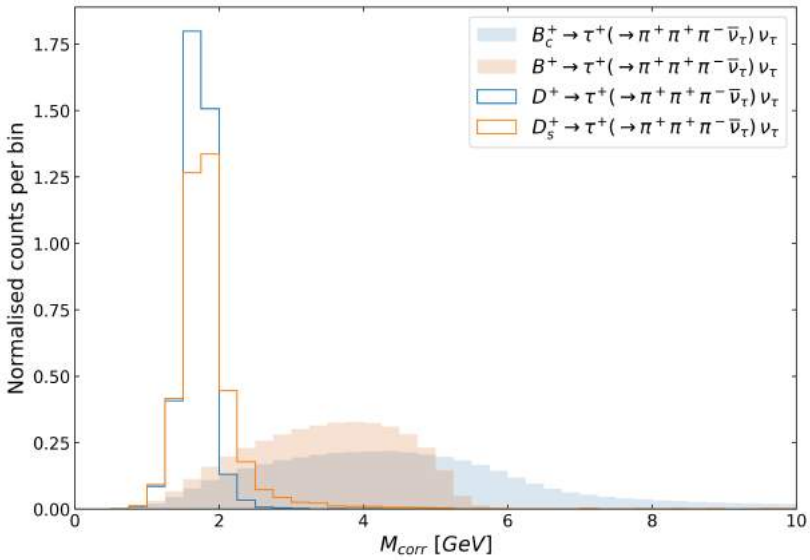

Figure 5.1: The normalised yield in bins of corrected mass for $B^+_{(c)} \to \tau^+\nu_\tau$ signals and *charm* background contributions.

- **Semileptonic, $\boldsymbol{\tau \to 3\pi}$**: $B^+ \to \tau^+\nu_\tau X$, $\tau^+ \to 3\pi\,\overline{\nu}_\tau$
- **Semileptonic, $\boldsymbol{(q\overline{q}) \to 3\pi}$**: $B^+ \to l^+\nu_l X\,(q\overline{q}), (q\overline{q}) \to 3\pi X$
- **Direct $\boldsymbol{B}$-decay**: $B^+ \to X\,3\pi$
- **Intermediate charm**: $B^+ \to D^{(*)+}_{(s)} X$, $D^{(*)+}_{(s)} \to 3\pi X$

where only the 3π are reconstructed. The neutrino is not detected by LHCb, $(q\overline{q})$ is a meson state (e.g. $D^{(*,**)0}$) and $l = e, \mu$. The latter three categories are similar in the sense that they are all of the form $B^+ \to 3\pi\,X$. However, for explanatory purposes it is decided to divide them into these separate categories.

The complete list of considered beauty backgrounds is also shown in table A.4. Possible extensions on this list are given in section 8. For all backgrounds, the corrected mass distribution of the 3π-combination of the signal and background decays are compared in the following.

5.5.1 Semileptonic, $\boldsymbol{\tau \to 3\pi}$

Decays that look similar to the signal decay but are partially reconstructed may be interesting (semi)leptonic backgrounds. They are of the form $B^+ \to \tau^+(\to 3\pi\,\overline{\nu}_\tau)\nu_\tau X$. Decays where 3π is produced from a meson state are considered in section 5.5.2. The Particle Data Group [23] shows that $B^+ \to \overline{D}^{(*)0}\tau^+\nu_\tau$ are the only (semi)leptonic decays for which the branching fraction is explicitly measured for the τ-lepton. The remaining (semi)leptonic decays are only measured for $l = e, \mu$. For that reason, all measurements on l^+ are assumed to hold for τ^+ as well, such that $\mathcal{BR}(B^+ \to X\,l^+\nu_l) = \mathcal{BR}(B^+ \to X\,\tau^+\nu_\tau)$. This estimate

is conservative considering the measurements of the $\mathcal{R}(D^{(*)})$-ratio [37], which suggest that $\mathcal{BR}(B^+ \to \overline{D}^{(*)}\,\tau^+\nu_\tau) < \mathcal{BR}(B^+ \to \overline{D}^{(*)}\,l^+\nu_l)$.

Regarding the yields, the potentially most dangerous backgrounds in this category are determined to be

$$B^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau\,\overline{D}$$

where $\overline{D} \in \{\overline{D}^0, \overline{D}^{*0}, \overline{D}_2^*(2460)^0\}$ is not reconstructed. Neutral isolation is assumed as $\overline{D}$ can decay into both charged and neutral particles. The backgrounds have respective yields of $3.10 \cdot 10^4$, $7.46 \cdot 10^4$ and $7.31 \cdot 10^4$. Figure 5.2 shows the normalised corrected mass distributions of the backgrounds and signals. As can be seen, the three background distributions are similar in shape. Compared to the charm backgrounds, the distributions are broader and peak more towards signal. However, their yields are significantly smaller than the charm yields.

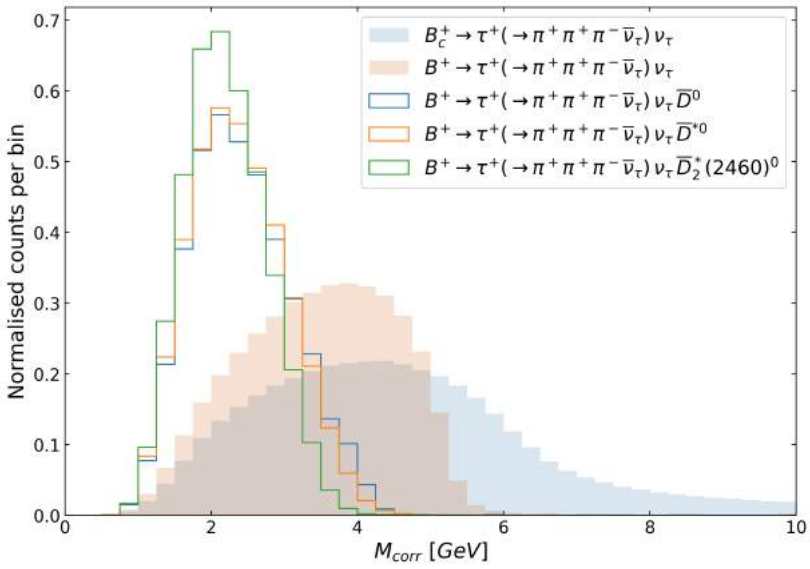

Figure 5.2: The normalised yield in bins of corrected mass for $B^+_{(c)} \to \tau^+\nu_\tau$ signals and *semileptonic,* $\tau \to 3\pi$ background contributions.

5.5.2 Semileptonic, $(q\overline{q}) \to 3\pi$

This (semi)leptonic category includes background decays with the 3π originating from a meson state. The potentially most dangerous backgrounds are determined to be of the form $B^+ \to \overline{D}(\to \pi^+\pi^+\pi^+X)\,l^+\nu_l$, where $\overline{D} \in \{\overline{D}^0, \overline{D}^{*0}, \overline{D}_2^*(2460)^0\}$ and $l = e, \mu$ are not reconstructed. $l = \tau$ is not considered here as the $\mathcal{R}(D^{(*)})$-ratio suggests that these modes are suppressed with respect to $l = e, \mu$. Moreover, on average, heavier particles lead to smaller corrected masses (i.e. having less overlap with the signal decays). Note that this category involves similar decays as in section 5.5.1, the difference being that the 3π-combination now originates from a meson instead of a lepton.

Based on the assumptions in section 5.3, the three largest backgrounds contributing to this category are considered to be

- $B^+ \to \overline{D}^0\, l^+\nu_l$, where $\overline{D}^0 \to K^+\,\pi^-\pi^-\pi^+$
- $B^+ \to \overline{D}^{*0}\, l^+\nu_l$, where $\overline{D}^{*0} \to \overline{D}^0(\to K^+\,\pi^-\pi^-\pi^+)\,\gamma$
- $B^+ \to \overline{D}^*_2(2460)^0\, l^+\nu_l$, where $\overline{D}^*_2(2460)^0 \to D^-(\to K^0_S\,\pi^-\pi^-\pi^+)\,\pi^+$

where it is assumed that only $\overline{D}^{*0} \to \overline{D}^0\gamma$ occurs (i.e. ignoring the $\overline{D}^{*0} \to \overline{D}^0\pi^0$ mode). Only the 3π-combination is reconstructed. Charged isolation is assumed for all decays as the charged lepton $l = e, \mu$ is not reconstructed. For $l = e$ (similar values are obtained for $l = \mu$), the respective yields are estimated to be $2.76 \cdot 10^3$, $6.68 \cdot 10^3$ and $7.90 \cdot 10^3$. Their corrected mass distributions are shown in figure 5.3 and compared with the signal distributions. Compared to the (semi)leptonic decays of the previous section ($\tau \to 3\pi$), their background shapes are similar but their yields are approximately 10 times smaller.

Note in table A.4 that $\epsilon_{tracking}$ of the latter background is approximately three times as large compared to the other backgrounds. This is due to the fact that it has an intermediate charged D-meson that may lead to additional VELO-hits, whilst the $\overline{D}^0$-meson is neutral and thus not leaving any hits.

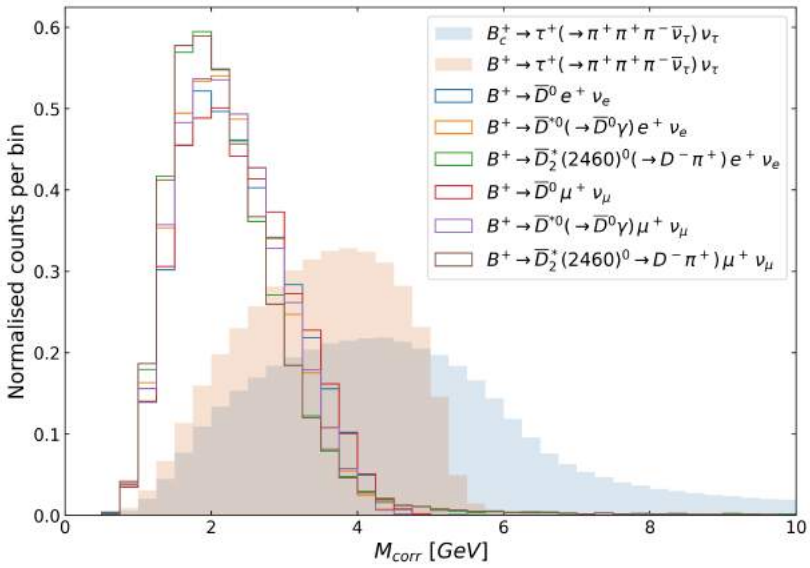

Figure 5.3: The normalised yield in bins of corrected mass for $B^+_{(c)} \to \tau^+\nu_\tau$ signals and *Semileptonic,* $(q\overline{q}) \to 3\pi$ background contributions. $\overline{D}^0 \to K^+\,\pi^-\pi^-\pi^+$ and $D^- \to K^0_S\,\pi^-\pi^-\pi^+$ are assumed.

5.5.3 Intermediate charm

There are several decays where three charged pions are produced from different charged D-mesons. The potentially most dangerous backgrounds in this category are of the form $B^+ \to D_s^{(*)+}\overline{D}$, where $\overline{D} \in \{\overline{D}^0, \overline{D}^{*0}, \overline{D}^*_2(2460)^0\}$ is not

reconstructed. Again, neutral isolation is assumed as $\overline{D}$ can decay in charged and neutral particles. The backgrounds that are considered are

- $B^+ \to D_s^+ \overline{D}^0$
- $B^+ \to D_s^{*+} \overline{D}^0$, with $D_s^{*+} \to D_s^+ \gamma$
- $B^+ \to D_s^+ \overline{D}^{*0}$
- $B^+ \to D_s^{*+} \overline{D}^{*0}$, with $D_s^{*+} \to D_s^+ \gamma$
- $B^+ \to D_s^{(*)+} \overline{D}_2^*(2460)^0$, with $D_s^{*+} \to D_s^+ \gamma$

where only $D_s^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ is considered. The $D_s^{*+} \to D_s^+\pi^0$ mode is ignored. Note that the last listed background is inclusive; no measurements have been done for the separate D_s^+ and D_s^{*+} contributions. It is assumed that the branching fractions of the B-meson decaying into $D_s^+\overline{D}_2^*(2460)^0 : D_s^{*+}\overline{D}_2^*(2460)^0$ is $1:2$ [36]. In this way, the expected yields are determined to be $3.48 \cdot 10^3$, $2.81 \cdot 10^3$, $3.13 \cdot 10^3$, $6.16 \cdot 10^3$, $3.19 \cdot 10^3$ and $6.13 \cdot 10^3$, respectively. All background contributions are shown in figure 5.4 and compared with signal contributions. Compared to the (semi)leptonic backgrounds, the distributions are peaking more towards signal. This holds especially for the $D_s^+\overline{D}^0$, $D_s^{*+}\overline{D}^0$ and $D_s^+\overline{D}^{*0}$ backgrounds. Their yields are comparable to the semileptonic $(q\overline{q}) \to 3\pi$ background yields.

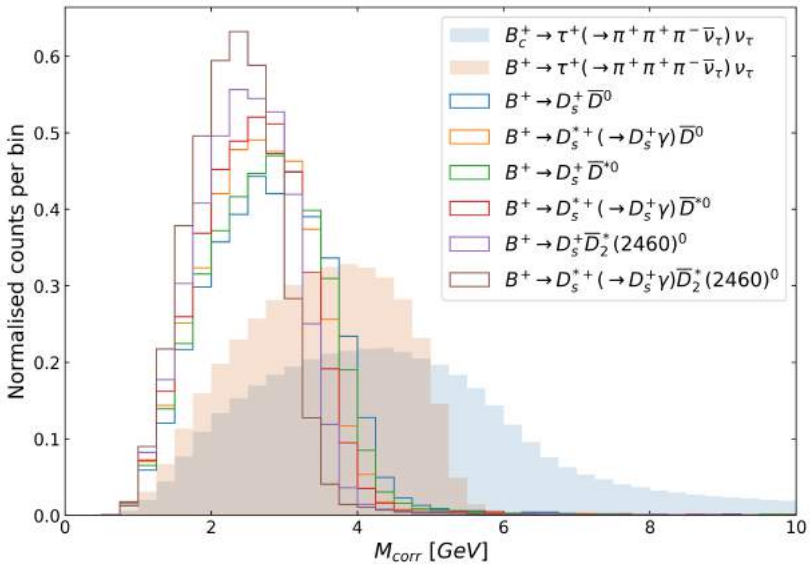

Figure 5.4: The normalised yield in bins of corrected mass for $B_{(c)}^+ \to \tau^+\nu_\tau$ signals and *intermediate charm* background contributions. $D_s^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ is assumed.

The same decays may also contribute to background if instead $\overline{D} \to 3\pi\, X$ but $D_s^{(*)+}$ is not reconstructed. However, this contribution is not considered

because in that case charged isolation can be used. Also, corresponding branching fractions are smaller.

There also exist similar decays that are of the form $B^+ \to D^{(*)+}\overline{D}$, where $\overline{D} \in \{\overline{D}^0, \overline{D}^{*0}, \overline{D}_2^*(2460)^0\}$. Their expected yields are determined to be approximately 10 times smaller than the $B^+ \to D_s^{(*)+}\overline{D}$ modes. Therefore, there is not further looked into these backgrounds.

5.5.4 Direct *B*-decay

Potential backgrounds may be due to B-mesons that directly decay into three charged pions and additional particles X that are not reconstructed. They are of the form $B^+ \to \pi^+\pi^+\pi^- X$. Three backgrounds that are considered in this category are

- $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$
- $B^+ \to \overline{D}^{*0}\,\pi^+\pi^+\pi^-$
- $B^+ \to \overline{D}^{*0}\,\pi^+\pi^+\pi^-\,\pi^0$

where the D-meson and neutral pion are not reconstructed. Neutral isolation is assumed as the D-meson can decay into both charged and neutral particles. The respective yields are determined to be $1.84 \cdot 10^5$, $3.38 \cdot 10^5$ and $5.92 \cdot 10^5$. The latter background has the largest yield, however as figure 5.5 shows, especially $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$ and $B^+ \to \overline{D}^{*0}\,\pi^+\pi^+\pi^-$ coincide significantly more with the mass shapes of the signals.

The difference between $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$ and $B^+ \to \overline{D}^{*0}\,\pi^+\pi^+\pi^-$ is that an additional (not reconstructed) particle is produced in the latter case. Namely, $\overline{D}^{*0}$ decays into $\overline{D}^0$ with the production of an additional π^0 or γ. As less (not reconstructed) additional particles are produced for the former background, the probability for that decay to pass the trigger is larger. For this reason, $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$ is considered to be the most dangerous background in this category.

5.6 The most dangerous background

In the previous section, corrected mass spectra of various background categories were shown and compared with signal distributions. Both the shapes and yields were considered important measures for identifying potentially dangerous backgrounds. For comparison purposes, the considered background shapes are shown together in figure 5.6. The background with the largest yield within each category is shown; an exception is the direct B-decay $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$.

The most overlapping shapes are observed in the *direct B-decay* category, where the background distributions peak at similar corrected masses as the signals. All other background categories peak to lower corrected masses and show less overlap.

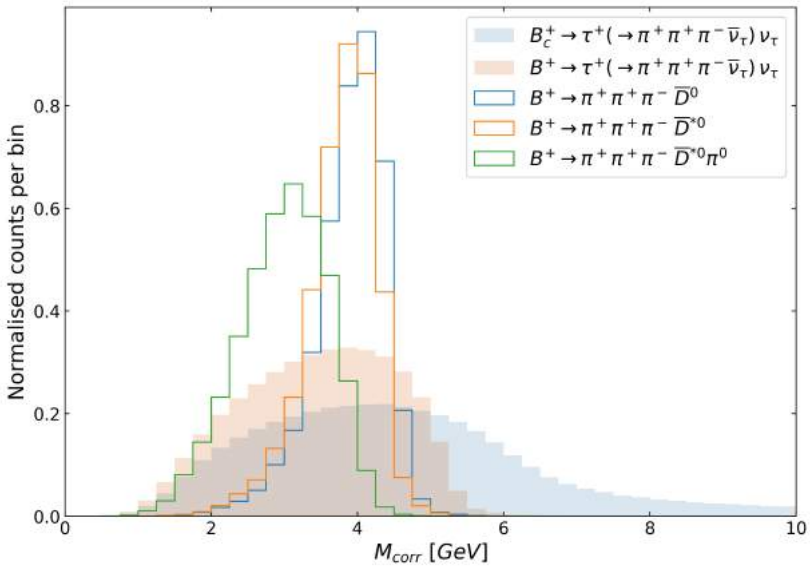

Figure 5.5: The normalised yield in bins of corrected mass for $B^+_{(c)} \to \tau^+\nu_\tau$ signals and the *direct B-decay* background contributions.

Table A.4 shows that the yields corresponding to the *direct B-decays* are the largest. The *charm* yield is comparable to the *direct B-decay* yield, but its distribution peaks very sharply at lower corrected mass (i.e. it is more distinguishable from signal). Remaining background categories have significantly smaller yields with respect to the *direct B-decays*.

In summary, both the corrected mass distributions and yields of the *direct B-decays* are considered to be the most dangerous. As section 5.5.4 motivated, the $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$ is considered the most dangerous background within the category.

For the purpose of the feasibility study, chapter 6 and 7 will look into a simplified scenario. This scenario is obtained by only considering the most dangerous background, $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$. By assuming that all produced b-hadrons lead to this background, so that $\mathcal{BR}(B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^- = 1)$, the scenario is turned into a worst-case scenario. If the worst-case measurements of the $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ branching fractions are feasible, it will also be in a more realistic scenario.

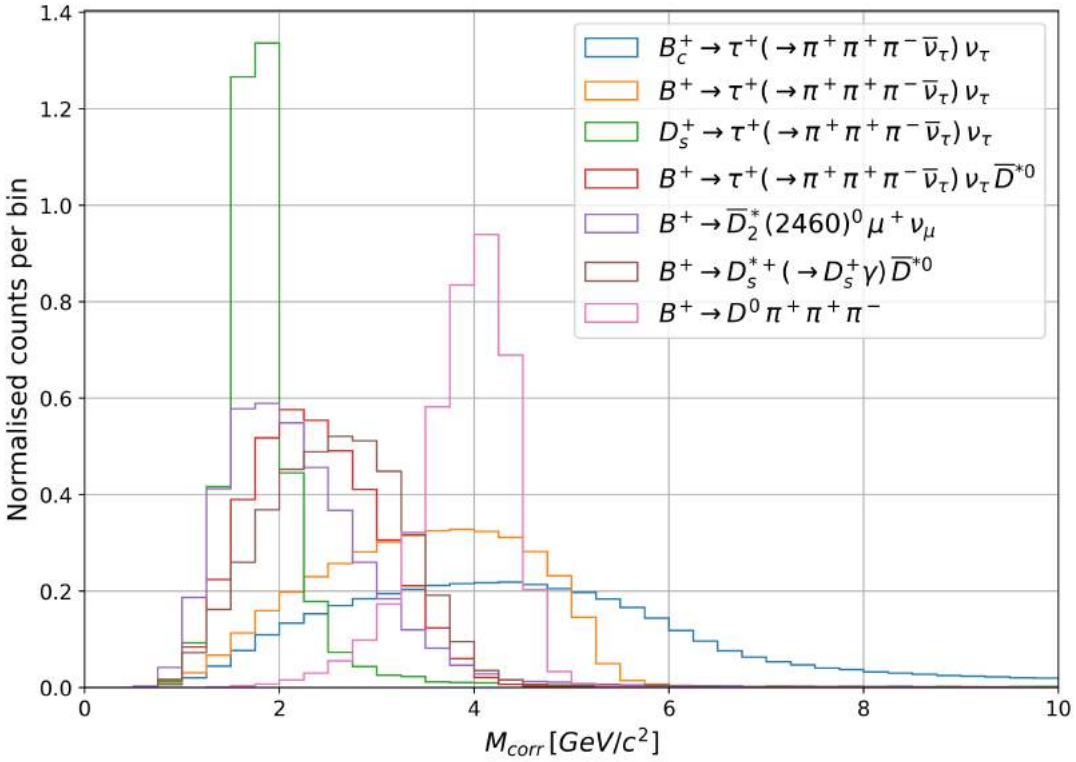

Figure 5.6: The normalised yield in bins of corrected mass for $B^+_{(c)} \to \tau^+\nu_\tau$ signals is shown. For each category, the background with the largest yield is shown as well; an exception is the direct B-decay $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$.

Chapter 6

Multivariate analysis for signal classification

Chapter 5 concluded with the observation that the *direct B-decay* category is considered to be the most dangerous. Especially, $B^+ \to D^0\,\pi^+\pi^+\pi^-$ is given a closer a look in the following chapters. Section 6.1 introduces the concept of multivariate analysis (MVA) which is applied on the background above in section 6.2.

6.1 Multivariate analysis

An experimental dataset contains mixed events, i.e. events corresponding to signal and background decay modes. To obtain a large as possible signal to background ratio, the two types of events have to be distinguished. This is done by the observables mentioned in section 4.3. In general, a larger discrepancy between the observables leads to more effective separation power. For the data classification, a gradient boosting classifier algorithm is used from the scikit-learn package [38].

6.1.1 Defining the test-statistic

The separation of signal and background is done by introducing a test-statistic. By the Neyman-Pearson lemma, the best test-statistic is defined as the likelihood ratio between signal and background [39]. Conceptually, this ratio tells in which region of the parameter space there is found more signal than background. For these likelihoods, explicit equations describing the multidimensional signal and background shapes are needed. This makes it difficult to use the likelihood ratio as test-statistic. Therefore, the test-statistic is built using MVA techniques, which makes use of both the range and shape of the observables. In this way, complex correlations between variables can be taken into account.

The MVA combines multiple observables into a single classifier that ranges between 0 and 1. A score close to 1 (0) means that the event is most probable

to be signal (background). For an optimal classification, the score corresponding to signal and background events is desired to be as distinct as possible.

6.1.2 Training phase

In order for the MVA to correctly classify signal and background events, it is trained on the corresponding dataset. The training dataset consists of the observables of both signal and background events and is obtained from RapidSim simulations. It is used as input for the MVA algorithm, such that the MVA iteratively learns to distinguish its signal from background events.

Specifically, the `GradientBoostingClassifier` algorithm aims to minimize a loss function by building up a boosted decision tree (BDT). The tree branches out by slicing the dataset using a single observable at each node. After a user-specified condition is reached (e.g. by defining a maximum number of iterations), every tree results in a set of final nodes (called the leaves). Every leaf is associated to an event being signal or background, depending on the contribution of each after applying the successive data cuts. The test-statistic that is obtained is henceforth referred to as BDT.

At every iteration during the MVA training, the improved decision tree results in a reduction of the loss function. However, it may occur that the model is trained too well on the training data, leading to a phenomenon called overtraining. A model that is overtrained tends to classify events that are due to statistical fluctuations. This is undesirable because for a similar dataset these fluctuations are different. Boosting is able to solve this problem. For each iteration, it assigns larger weights to misclassified events so that the new decision tree focuses on eliminating these. Verifying the MVA on a similar dataset is recommended, to check whether the MVA is performing as expected and is not overtraining the dataset.

6.1.3 Verification phase

During the verification phase, it is important to use a dataset that is independent of the training dataset, in order to prevent any biasing. The trained MVA predicts a certain event to be either signal or background using the BDT-score.

Note that this classification is not always correct. For example, the MVA may classify an event to be background-like whilst it is flagged as signal. Such an event is called a false negative. Instead, an event that is flagged as signal and classified correctly is called a true positive. Analogously, an event being flagged as background and classified as signal (background) is called false positive (true negative). An MVA that is able to classify many events correctly, has many true positives and negatives. The amount of true positives (i.e. true positive rate, or TPR) and false positives (i.e. false positive rate, or FPR) gives a characterization of the MVA performance.

The TPR versus the FPR can be plotted to form a receiver operating characteristic (ROC) curve. For every point on the ROC curve, the dataset is sliced

at a certain BDT-score. Depending on the particular cut, the relative amount of signal (the so-called purity) in the sliced dataset changes. This results in a different amount of misclassified events, i.e. FPR and TPR.

The MVA performance can be quickly interpreted by the curve. Perfect classification corresponds to the point where (FPR, TPR) = (0, 1). On the diagonal line that connects the bottom left to the top right corners, there are as many true positives as false positives. Thus, slicing the dataset at the corresponding FPR does not improve the purity of the dataset. Points above (below) this diagonal correspond to better (worse) classification.

Besides indicating the MVA performance, ROC curves can be used to identify signs of overtraining as well. In general, ROC curves obtained from the training and verification dataset may differ significantly for an MVA that is overtrained.

A different measure of the MVA performance is the area under the curve (AUC). Perfect classification corresponds to an AUC of 1, whilst a diagonal curve corresponds to an AUC of 0.5. Thus, a high AUC score corresponds to a well-performing MVA for that particular dataset.

6.2 Multivariate analysis on $B^+ \to \overline{D}^0 \pi^+\pi^+\pi^-$

In this section, different MVAs are applied at different datasets to compare their performances. Simulated datasets corresponding to $B_c^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$, $B^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ and $B^+ \to D^0\pi^+\pi^+\pi^-$ are denoted with B_c^+, B^+ and BG for short, respectively. The respective datasets that are used contain approximately 943.000, 5.4 million and 67 million events. MVA1 is trained using $\{B_c^+, BG\}$ data; MVA2 is trained using $\{B^+, BG\}$ data; MVA3 is trained using $\{B_c^+, B^+\}$ data. Their corresponding BDT-scores are respectively referred to as BDT1, BDT2 and BDT3 and used as input for the likelihood fits in chapter 7.

All MVAs are trained using 100.000 signal and background events. Verification of the MVA1 and MVA2 is done on B_c^+, B^+ and BG; verification of MVA3 is only done on the former two datasets. The MVAs are checked for overtraining by comparing ROC-curves of the training and verification data, which should be very similar. Also, the counts in bins of BDT-score should be similar for training and verification data, when there is accounted for the different size of the datasets.

6.2.1 Input variables

Multiple observables are used to train the MVA. As defined in section 4.3, the (transverse) momentum of the 3π, the flight distance between the PV and 3π-vertex, the corrected opening angle θ_{corr}, the number of VELO-hits and the impact parameter of the 3π are used as input variables. The corrected mass m_{corr} will be used in a two-dimensional likelihood fit in chapter 7 together with the BDT-score obtained from the MVA.

6.2.2 Obtaining BDT1: training on B_c^+ and BG

The first MVA is trained using B_c^+ and BG events as input data. The six observables described above are visualised in figure 6.1 for both categories. Large differences are seen in the transverse momentum and impact parameter of the 3π. These differences can be explained by the fact that the B_c^+-meson is more massive and short-lived than the B^+-meson.

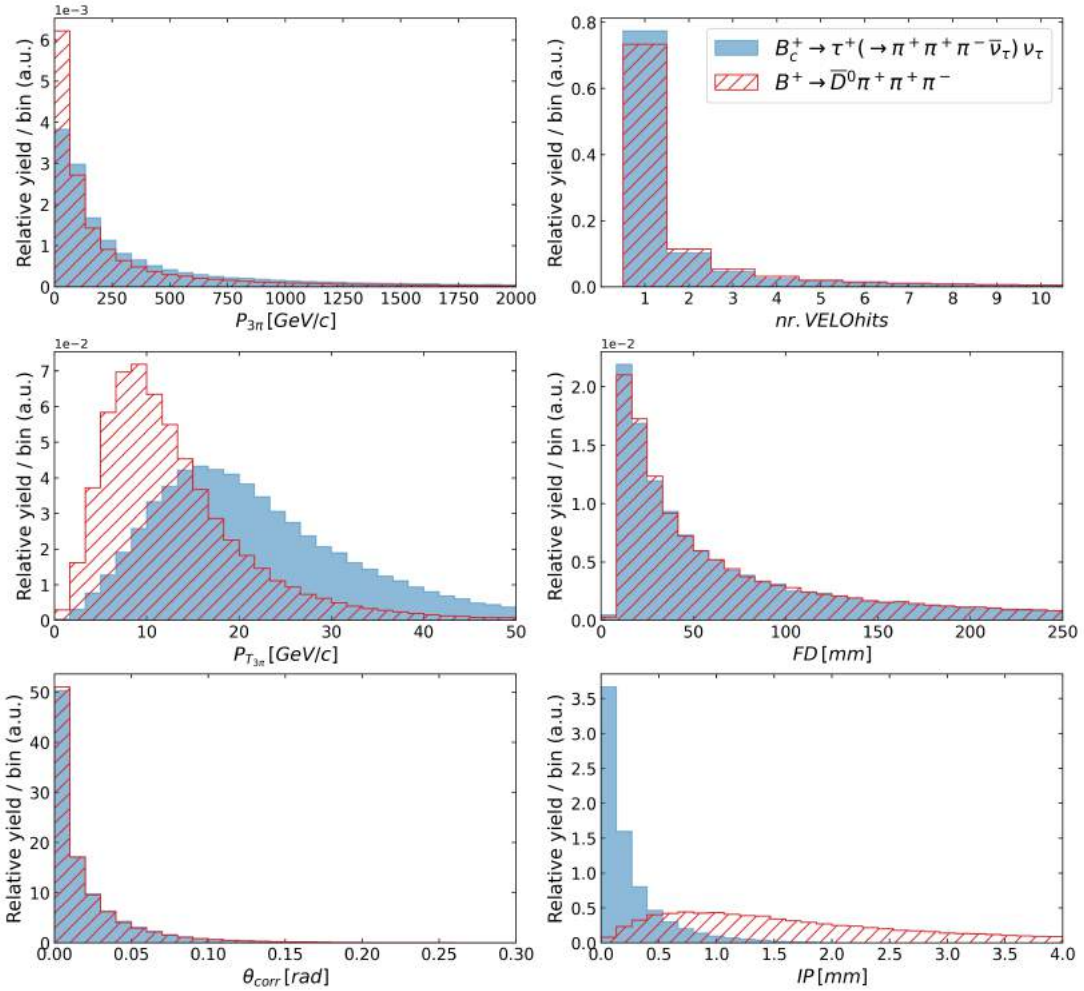

Figure 6.1: A comparison between $B_c^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ and $B^+ \to D^0\pi^+\pi^+\pi^-$ on the distribution of the normalised yield for different observables.

After training, the BDT-score is determined for every event in the verification data. A BDT-score close to 1 means that an event is most probable to be signal-like (i.e. a B_c^+-event). The normalised number of counts is plotted versus the BDT-score and shown in figure 6.2 for all three components. The B_c^+ contribution peaks towards 1, whilst the BG component peaks towards 0.

Although the MVA is not trained to distinguish the B^+ contribution from the others, the MVA is verified on it as well. This is done in order to check the MVA performance on a contribution that it was not trained on. The number of

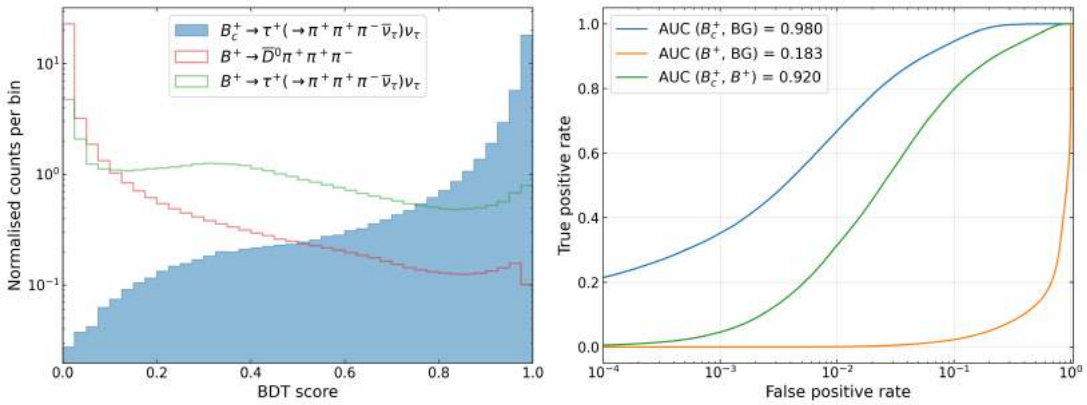

Figure 6.2: Both plots are obtained by training on $\{B_c^+, BG\}$. Left: The normalised counts per bin versus the BDT-score for both signals and $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$. Right: ROC curves obtained using different combinations of the same three components.

counts in bins of BDT-score is significantly more flattened with respect to the other contributions. As there is significantly more overlap between the $\{B_c^+, BG\}$ distribution compared to the $\{B_c^+, B^+\}$ distribution, the MVA is expected to distinguish less well for the latter case.

This is verified by the corresponding ROC-curve that is shown in the same figure. An AUC-score of 0.980 is obtained after verification on $\{B_c^+, BG\}$. As expected from the BDT-distribution in the same figure, the MVA performance on $\{B_c^+, B^+\}$ data is worse compared to the $\{B_c^+, BG\}$ case. However, still many B^+ events can be distinguished from B_c^+, judging from the AUC-score of 0.920.

As the BG and B^+ both tend to be classified as background-like, their BDT-distributions are similar. This can be explained by the fact that the MVA is trained to separate the datasets based on the IP of the pions, making use of the shorter lifetime of B_c^+ compared to B^+. The similar BDT-distribution explains the low AUC-score of 0.183. The score being lower than 0.5 is due to convention of labelling the two classes. Interchanging the labels gives an AUC-score of 0.817. The particular ROC-curve is shown in this way to emphasize the overlap between the B^+ and BG distribution in the BDT-score.

6.2.3 Obtaining BDT2: training on B^+ and BG

The second MVA is trained using B^+ and BG events as input data. The same input observables as above are used and are shown in figure 6.3. Much less discrepancy in the IP and $p_{T_{3\pi}}$ is seen, as both decays have B^+-mother. As the distributions of the B^+ signal are different with respect to B_c^+ signal, the MVA performance is expected to be different as well.

The MVA performance is shown in figure 6.4 where the normalised counts per bin is plotted versus the BDT-score. The BG distribution peaks at low BDT-scores, whilst the peak is not as sharp for B^+ and B_c^+. Nevertheless, as expected, B^+ peaks at a different BDT-score than BG does. B_c^+ tends to be classified as signal-like. Interestingly, although the MVA is trained on $\{B^+, BG\}$,

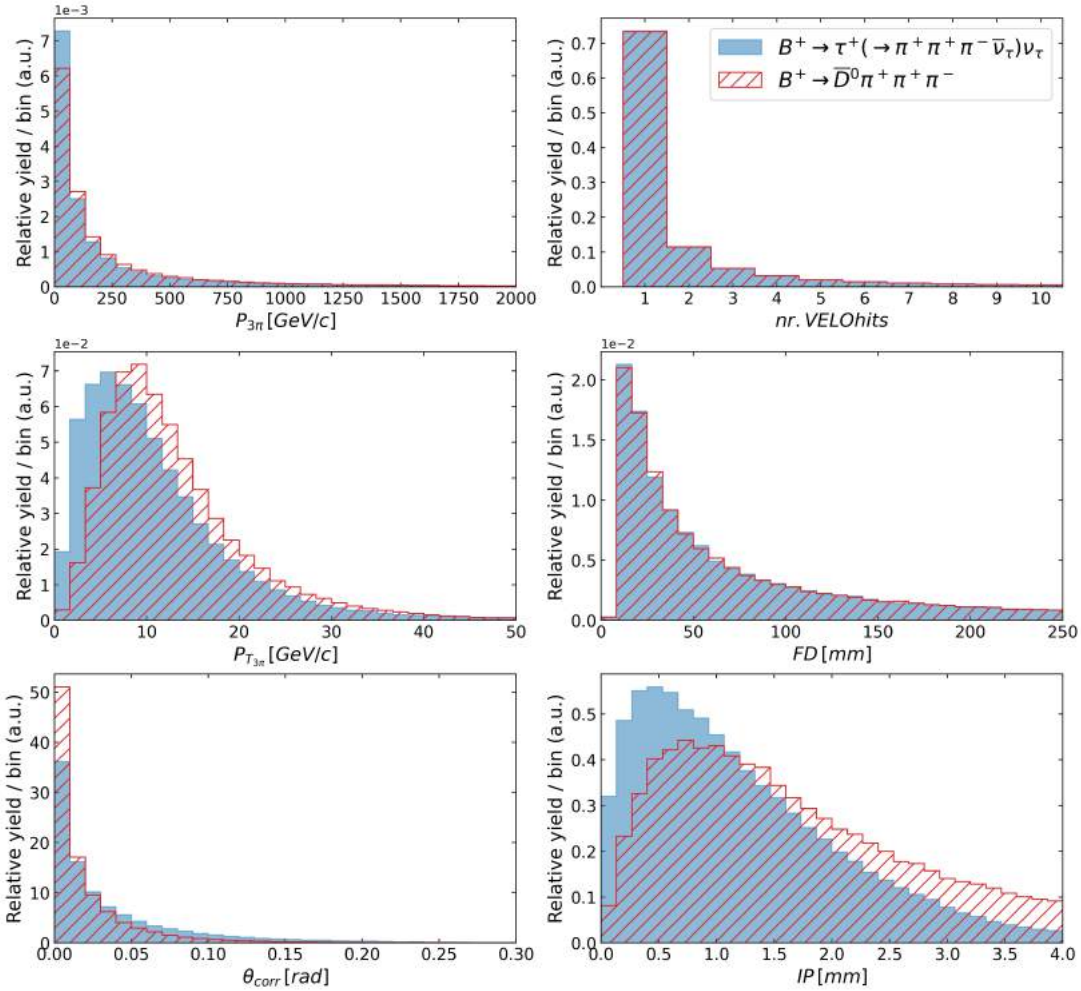

Figure 6.3: A comparison between $B^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ and $B^+ \to D^0\pi^+\pi^+\pi^-$ on the distribution of the normalised yield for different observables.

the separation between B^+ and B_c^+ is larger than between B^+ and BG. This shows that this MVA mainly classifies events based on the final states rather than on the B-mesons. The short-lived B_c^+-mother may be more discriminating compared to BG, resulting in a large separation. This means that this MVA can be used for classification of B_c^+ as well.

The performance of classification is verified in the corresponding ROC-curve, also shown in figure 6.4. The verification of the MVA on $\{B^+, BG\}$ gives an AUC-score of 0.910. A better score of 0.948 is obtained for $\{B_c^+, BG\}$. A score of 0.331 is obtained for $\{B^+, B_c^+\}$, reflecting the significant overlap between the two datasets. Again, the AUC-score below 0.5 is due to the labelling of the classes.

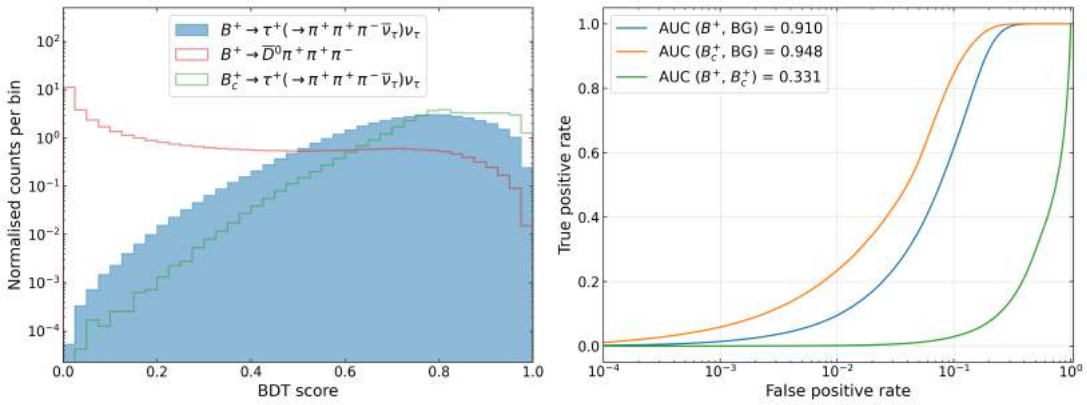

Figure 6.4: Both plots are obtained by training on $\{B^+, BG\}$. Left: The normalised counts per bin versus the BDT-score for both signals and $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$. Right: ROC curves obtained using different combinations of the same three components.

6.2.4 Obtaining BDT3: training on B_c^+ and B^+

MVA3 is trained on a dataset containing $\{B_c^+, B^+\}$. The distributions of the observables for B_c^+ and B^+ datasets are not shown, as they are already included in figure 6.1 and 6.3, respectively. Differences in the distributions are especially seen for the $p_{T_{3\pi}}$ and IP, just as for the observables in MVA1. Again, this can be explained by the difference between the decaying B-meson.

Figure 6.5 shows the normalised counts per bin versus the BDT-score. The corresponding ROC-curve is shown as well. As expected, the B_c^+ signal peaks at high BDT-scores, whilst B^+ peaks at low BDT-scores. An AUC-score of 0.923 is obtained. Interestingly, the performance of MVA3 is very similar to that of MVA1 when there is verified on the $\{B_c^+, B^+\}$ dataset. If we compare the AUC-scores from figures 6.2 and 6.5, the separation between B_c^+ and B^+ is similar in both cases.

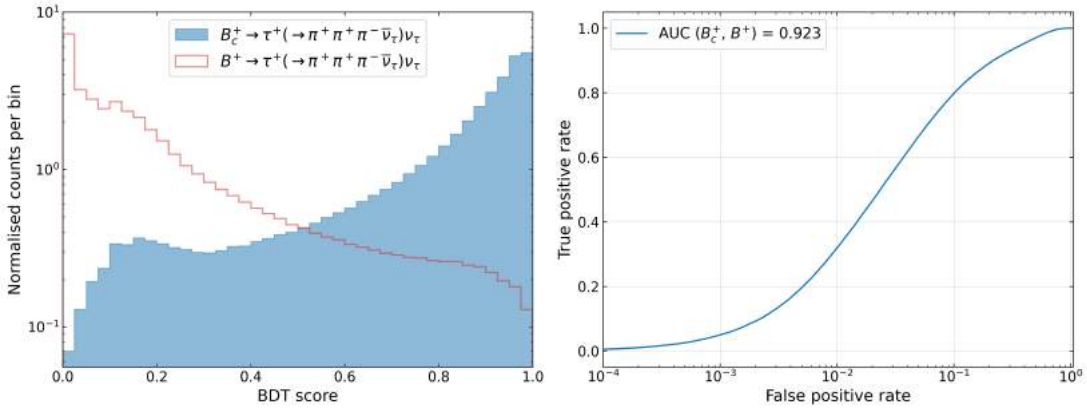

Figure 6.5: Both plots are obtained by training on $\{B_c^+, B^+\}$. Left: The normalised counts per bin versus the BDT-score for B_c^+ and B^+. Right: The ROC curve corresponding to the $\{B_c^+, B^+\}$ dataset.

Chapter 7

Likelihood fit

This chapter report on the significance to observe the $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ decays. For that, a likelihood fit is performed. Section 7.1 explains the theory between the maximum likelihood model and discusses the frameworks that are used. The model is applied to the worst-case background in section 7.2.

7.1 Maximum likelihood model

The signal yields and their corresponding uncertainties are determined by performing a likelihood fit. For this a probability density function (PDF) that describes the dataset has to be made. In particular, the total PDF consists of the $B_{(c)}^+ \to \tau^+\nu_\tau$ signals and the $B^+ \to \overline{D}^0\pi^+\pi^+\pi^-$ background. In the following, these are abbreviated with B_c^+, B^+ and BG, respectively. Each PDF is described by $\vec{x} \in (BDT, m_{corr})$, which are the two dimensions for the fit. Specifically, the total PDF is given by

$$\mathcal{P}(\vec{x}|\theta) = f_{B_c^+}\mathcal{P}_{B_c^+}(\vec{x}|\theta) + f_{B^+}\mathcal{P}_{B^+}(\vec{x}|\theta) + (1 - f_{B_c^+} - f_{B^+})\mathcal{P}_{BG}(\vec{x}|\theta) \qquad (7.1)$$

such that it consists of the individual components $\mathcal{P}_X(\vec{x}|\theta)$ with $X \in \{B_c^+, B^+, BG\}$. Each component has a particular shape that is generally described by parameters θ. The relative contribution of each PDF to $\mathcal{P}(\vec{x}|\theta)$ is determined by $f_{B_c^+}$ and f_{B^+}, which denote the B_c^+- and B^+-fraction with respect to the total expected yield $N = n_{B_c^+} + n_{B^+} + n_{bg}$. As the sum of the fractions should be 1, the background fraction can be written as $1 - f_{B_c^+} - f_{B^+}$. Hence, the only free parameters in the fit are $f_{B_c^+}$ and f_{B^+}. The signal yield n_X can then be determined by $n_X = f_X \cdot N$. Equation 7.1 can also be rewritten such that the signal yields are the free parameters in the fit.

The best estimate of the fit is found by maximizing the likelihood function. Besides the signal fractions being free parameters, the total expected yield is not fixed as well. This is accounted for by multiplying the likelihood by an additional Poisson term, such that an extended likelihood is obtained. It is given by [40]

$$\mathcal{L}(\vec{x}|\theta) = \frac{(n_{B_c^+} + n_{B^+} + n_{bg})^{N_{obs}} e^{-(n_{B_c^+} + n_{B^+} + n_{bg})}}{N_{obs}!} \cdot \prod_{i=1}^{N_{obs}} \mathcal{P}(\vec{x}_i|\theta) \qquad (7.2)$$

where N_{obs} is the number of observed events in the dataset. The likelihood function is obtained by determining $\mathcal{P}_X(\vec{x}|\theta)$ at all observed datapoints and multiplying the outcomes with each other. By varying the signal fractions, the outcome of equation 7.2 changes accordingly. The joint probability, given by the product term, is a measure of how well the PDF corresponds to the observed dataset. A larger joint probability means that the PDF resembles the dataset more closely. Therefore, maximizing the likelihood gives the best estimate for the free parameters and thus the most compatible PDF. Because of computational reasons, in practice it is more convenient to minimize $-\ln\mathcal{L}(\vec{x}|\theta)$ instead.

Obtaining the uncertainty in the yield

To illustrate how the uncertainty in the signal fraction is obtained from the likelihood fit, let's assume that $\mathcal{L}$ is approximately Gaussian and only depends on θ. This means that $-\ln\mathcal{L}$ is approximately parabolic. Taylor expanding $-\ln\mathcal{L}$ around its minimum $\hat{\theta}$ gives [39]

$$-\ln\mathcal{L}(\theta) = -\ln\mathcal{L}(\hat{\theta}) - \frac{\partial\ln\mathcal{L}}{\partial\theta}\bigg|_{\theta=\hat{\theta}} - \frac{1}{2!}\frac{\partial^2\ln\mathcal{L}}{\partial\theta^2}\bigg|_{\theta=\hat{\theta}}(\theta-\hat{\theta})^2 - \cdots \tag{7.3}$$

where its dependence on $\vec{x}$ is implicitly assumed. The first derivative vanishes at the minimum; the second derivative can be determined by filling in $\ln\mathcal{L}$ explicitly. If higher order terms are neglected, equation 7.3 can be approximated as

$$-\ln\mathcal{L}(\theta) \simeq -\ln\mathcal{L}_{min} + \frac{(\theta-\hat{\theta})^2}{2\sigma_\theta^2} \tag{7.4}$$

where $\mathcal{L}_{min} = \mathcal{L}(\hat{\theta})$ and σ_θ is the uncertainty in $\hat{\theta}$. The upper and lower bounds corresponding to n standard deviations around $\hat{\theta}$ (i.e. $\hat{\theta}\pm n\sigma_\theta$) can then be determined using

$$-\ln\mathcal{L}(\hat{\theta}\pm n\sigma_\theta) = -\ln\mathcal{L}_{min} + \frac{n^2}{2} \tag{7.5}$$

This shows that the width of $-\ln\mathcal{L}(\theta)$ around the best estimate value is a measure of the uncertainty of that value. In this study, $\hat{\theta}$ is the best estimate of the signal fractions. All uncertainties that will be determined in section 7.2 are 1σ uncertainties.

Frameworks for fitting

Equation 7.2 describes the likelihood for unbinned data. Instead, the use of binned data is simpler and sufficient for a feasibility study. The principle of obtaining the signal fractions and their uncertainties is the same. The datasets of the signal and background components are obtained using RapidSim simulations, meaning that there are no free parameters that determine the PDF shape. The ROOT framework [41] is used to convert the datasets to binned data, by subsequently applying `RooDataSet` and `RooDataHist`. The PDFs of the signals and

background are combined into a total PDF by `RooAddPdf`, following equation 7.1.

Pseudo-data is generated based on the total PDF, so that it closely follows the PDF. So, instead of using observed data from an experiment, the maximum likelihood is determined using this pseudo-data. The maximization of the likelihood function is done numerically using the MINUIT package [42]. In particular, the minimization of $-\ln \mathcal{L}(\vec{x}|\theta)$ is performed using a combination of MIGRAD and HESSE. If the components in the fit are too correlated, the minimization procedure becomes more difficult such that convergence is not guaranteed.

Reliability check

A single likelihood fit gives an estimation on the signal fraction and uncertainty. The reliability of the fit can be checked by performing multiple independent likelihood fits, or toy studies, using the ROOFIT framework [43]. `RooMCStudy` is used to generate multiple pseudo-datasets in a user-friendly way, resulting in multiple likelihood fits. For every fit, a best estimate on its signal fraction and uncertainty is obtained together with its pull. The pull p is described by

$$p = \frac{n_f - n_t}{\sigma_{n_f}} \tag{7.6}$$

where $n_f \pm \sigma_{n_f}$ is the estimated fit result of a single pseudo-experiment; n_t is the true value of the signal fraction, i.e. the expected signal fraction. In the limit of a large number of fits, the signal fraction, uncertainty and pull show a Gaussian behaviour. Specifically, the likelihood fit is unbiased if the mean of the pull distribution is compatible with 0; the estimates for the uncertainties are reliable if its standard deviation is compatible with 1. The uncertainty in the signal fraction is extracted from the obtained error distribution by determining its mean value.

The likelihood fit is used to understand if the $B^+_{(c)} \to \tau^+\nu_\tau$ decay can be observed. An observation can be quantified using the (statistical) significance of the fit result. If the toy study shows a Gaussian behaviour, σ_{n_f} is the 1σ uncertainty of the result. The significance of the fit result is therefore given to good approximation by n_f/σ_{n_f}. A null hypothesis can be defined where the $B^+_{(c)} \to \tau^+\nu_\tau$ decay does not exist, whereas the alternative hypothesis represents the existence of the decay. A fit result having a large significance means that it has a low probability of occurring assuming the null hypothesis is correct. A significance exceeding 5σ would suggest that the decay mode can be observed. Fit results are obtained assuming SM branching fractions for $B^+_{(c)} \to \tau^+\nu_\tau$. Therefore, the significance is a measure of the deviation from the expected SM yield to a zero yield.

7.2 Likelihood fit on $B^+ \to \overline{D}^0 \pi^+\pi^+\pi^-$

In chapter 6, different MVA performances have been compared. In particular, MVA1 was trained on $\{B_c^+, BG\}$, resulting in the BDT1 test-statistic; MVA2 was trained on $\{B^+, BG\}$, resulting in the BDT2 test-statistic; MVA3 was trained on $\{B_c^+, B^+\}$, resulting in the BDT3 test-statistic. These BDT-scores are used as input for the likelihood fits that are done in this section. The B_c^+ and B^+ signals are fitted together with the most dangerous background that chapter 5 concluded upon. The worst-case scenario is assumed where $\mathcal{BR}(B^+ \to D^0\,\pi^+\pi^+\pi^-) = 1$. This section aims to understand if $B_{(c)}^+ \to \tau^+\nu_\tau$ can be observed in this scenario, which is quantified in terms of significance.

At first, an attempt was made to do a likelihood fit using absolute yields as free parameters. For these, the minimization procedure did not succeed in finding a convergent result. Using signal fractions as free parameters instead, results in convergence. The total PDF follows the form of equation 7.1. A two-dimensional binned likelihood fit is performed, where $m_{corr} \in [0, 10]\,\mathrm{GeV/c^2}$ is divided in 40 bins and BDT $\in [0, 1]$ is divided in 20 bins. The signal fractions, errors and pull are obtained by doing 1000 toy studies.

SM values are assumed for the signal branching fractions. The branching fraction of the background is assumed to be 100% and therefore not SM. The signal fractions according to the SM are determined using equation 5.1 and listed in table 7.1 for multiple reconstruction efficiencies. The particular efficiencies are chosen such that $1 - \epsilon_{rec}$ is halved for subsequently increasing ϵ_{rec}. An exception to this is the additional $\epsilon_{rec} = 88\%$. These SM signal fractions are used as input values for the free parameters in the fit. During the minimization process, each parameter varies in a user-defined range. Based on table 7.1, this range is chosen to be $f_{B_{(c)}^+} \in [-10^{-4}, 1]$. Small negative values are included in order to increase the symmetry in the likelihood function, especially because both input signal fractions lie close to zero.

Table 7.1: The estimated signal fractions determined from SM values for different reconstruction efficiencies, in the case where $\mathcal{BR}(B^+ \to D^0\,\pi^+\pi^+\pi^-) = 1$. These fractions are used as input values for the free parameters in the fit.

ϵ_{rec}	$B_c^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$	$B^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$
84 %	$6.517 \cdot 10^{-6}$	$9.206 \cdot 10^{-5}$
88 %	$8.688 \cdot 10^{-6}$	$1.227 \cdot 10^{-4}$
92 %	$1.303 \cdot 10^{-5}$	$1.841 \cdot 10^{-4}$
96 %	$2.606 \cdot 10^{-5}$	$3.681 \cdot 10^{-4}$
98 %	$5.210 \cdot 10^{-5}$	$7.360 \cdot 10^{-4}$
99 %	$1.041 \cdot 10^{-4}$	$1.471 \cdot 10^{-3}$
99.5 %	$2.079 \cdot 10^{-4}$	$2.937 \cdot 10^{-3}$

As a first step, a likelihood fit is done in the $\{\text{BDT1}, m_{corr}\}$ space[1]. This fit is discussed in section 7.2.1. Although MVA1 is not explicitly trained to separate B_c^+ from B^+, it may contain information to do so. MVA3 is specifically trained to separate B_c^+ from B^+. If MVA1 is able to separate B_c^+ from B^+ equally well as MVA3, it means that MVA3 does not add useful information to MVA1. In that case, only using MVA1 in the fit is sufficient. For this purpose, a *crosscheck* fit is done in section 7.2.2. Thus, this crosscheck aims to understand if the fit result using BDT1 can be improved by adding information about BDT3.

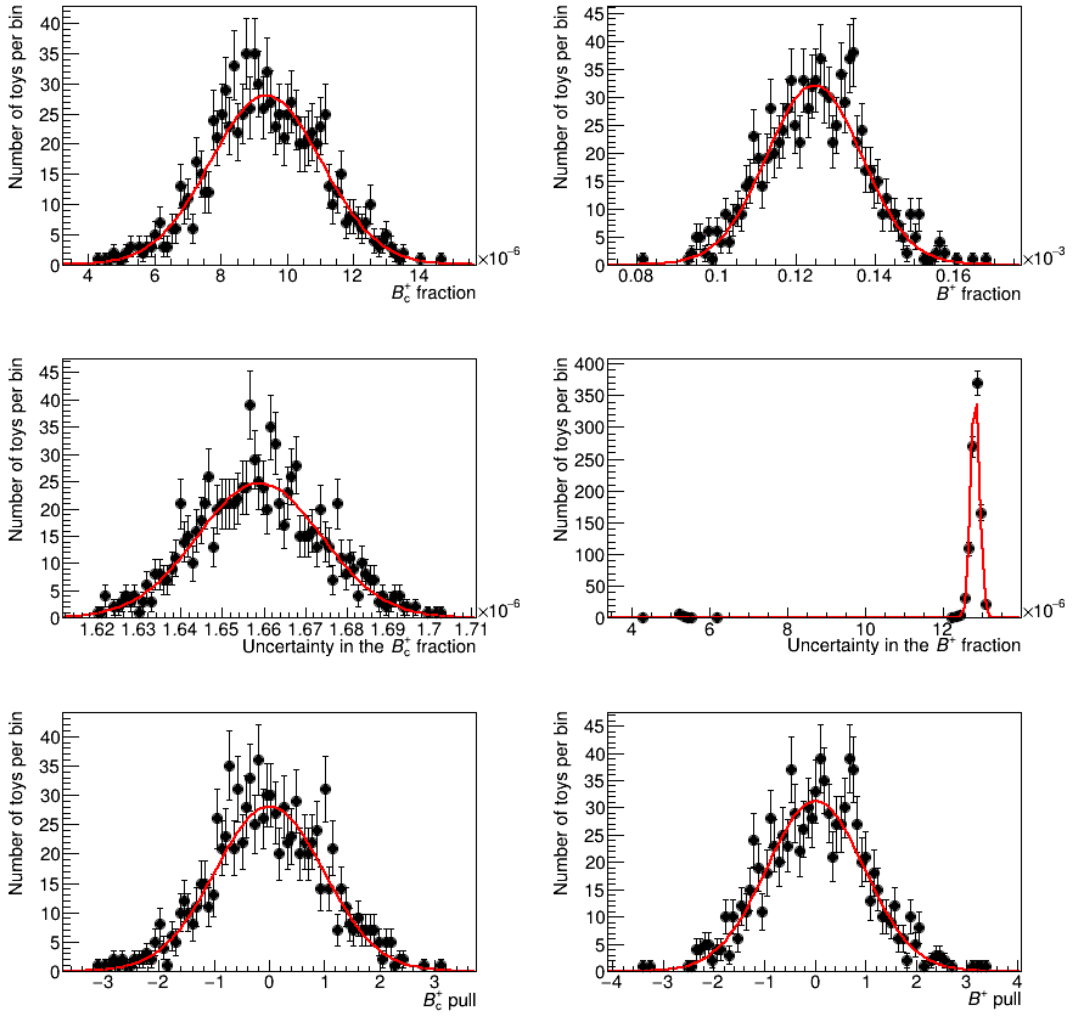

Figure 7.1: Results on the signal fractions, errors and pulls from the likelihood fit in the $(\text{BDT1}, m_{corr})$ space for $\epsilon_{rec} = 88\%$. Gaussian fits to the data are shown in red. The background yield that is used in the fit assumes that $\mathcal{BR}(B^+ \to D^0\,\pi^+\pi^+\pi^-) = 1$.

[1] A fit using the BDT2-score is not performed and can be done in further studies.

7.2.1 Likelihood fit using BDT1 and m_{corr}

In this section, the results on the likelihood fits in the (BDT1, m_{corr}) space is discussed. Fits are done for the reconstruction efficiencies listed in table 7.1. An example of a toy study for $\epsilon_{rec} = 88\%$ is shown in figure 7.1. For both signals, the number of toys per bin is shown versus the signal fraction, uncertainty and pull. The data can be well described by Gaussian fits, which are shown in red.

For a fit to be reliable, the mean and standard deviation of the pull should be compatible with 0 and 1, respectively. Results on the pull are listed in table B.1 of appendix B. Most results are compatible within 1.5σ, which is considered reliable in this study. The B_c^+-result from the fit is not reliable for $\epsilon_{rec} = 84\%$; the B^+-result from the fit is not reliable for $\epsilon_{rec} = 84\%$ and $\epsilon_{rec} = 99\%$. For the reliable fits, the best estimate of the fit is obtained by using the average values of the Gaussian fit of the fraction and uncertainty.

Figure 7.2 shows the projection of the fit for $\epsilon_{rec} = 88\%$, thus corresponding to the toy study in figure 7.1. It shows the absolute yield in bins of corrected mass corresponding to an integrated luminosity of 10 fb^{-1}. Pseudo-data is generated based on the total PDF, after which the fit is performed and based on the pseudo-data. Fit projections are very similar for other efficiencies. The only difference is the total background yield, as only ϵ_{rec} is different.

The procedure described above for $\epsilon_{rec} = 88\%$ is done for every efficiency listed in table 7.1. The correlation between $\mathcal{P}_{B_c^+}$ and $\mathcal{P}_{B^+}$ is approximately 27% for all fits and therefore low enough to avoid minimization problems. The

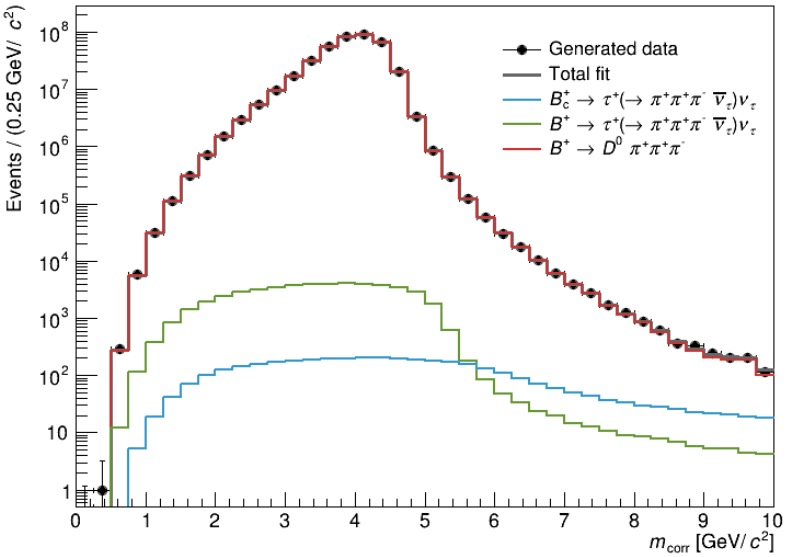

Figure 7.2: The projection of the two-dimensional likelihood fit on the corrected mass for $\epsilon_{rec} = 88\%$. The absolute background yield corresponds to an integrated luminosity of 10 fb^{-1} and $\mathcal{BR}(B^+ \to D^0\,\pi^+\pi^+\pi^-) = 1$.

Table 7.2: Signal fractions and their uncertainties as a function of reconstruction efficiency ϵ_{rec}, obtained from pseudoexperiment fits in the (BDT1, m_{corr}) space.

ϵ_{rec}	$f_{B_c^+} \pm \sigma_{B_c^+}$	$f_{B^+} \pm \sigma_{B^+}$	$f_{B_c^+}/\sigma_{B_c^+}$	f_{B^+}/σ_{B^+}
84 %	$(4.72 \pm 1.40) \cdot 10^{-6}$	$(9.32 \pm 1.10) \cdot 10^{-5}$	3.3	8.4
88 %	$(9.38 \pm 1.66) \cdot 10^{-6}$	$(1.25 \pm 0.128) \cdot 10^{-4}$	5.6	9.6
92 %	$(1.20 \pm 0.207) \cdot 10^{-5}$	$(1.73 \pm 0.158) \cdot 10^{-4}$	5.7	10.8
96 %	$(2.11 \pm 0.307) \cdot 10^{-5}$	$(3.91 \pm 0.228) \cdot 10^{-4}$	6.8	17.0
98 %	$(4.59 \pm 0.474) \cdot 10^{-5}$	$(7.20 \pm 0.328) \cdot 10^{-4}$	9.8	21.8
99 %	$(9.68 \pm 0.760) \cdot 10^{-5}$	$(1.46 \pm 0.0476) \cdot 10^{-3}$	12.7	30.4
99.5 %	$(2.23 \pm 0.128) \cdot 10^{-4}$	$(2.89 \pm 0.0715) \cdot 10^{-3}$	17.1	40.7

resulting best estimates of the signal fractions ($f_{B_{(c)}^+} \pm \sigma_{B_{(c)}^+}$) are shown in table 7.2. The corresponding significance of each, $f_{B_{(c)}^+}/\sigma_{B_{(c)}^+}$, is obtained and shown as well. As the toy studies can be well approximated by Gaussian fits, $\sigma_{B_{(c)}^+}$ corresponds to a 1σ deviation to good approximation. For that reason, the significance is well estimated in this way.

For both signals, table 7.2 shows larger significance for increasing ϵ_{rec}. This means that for larger ϵ_{rec}, the likelihood fit is able to determine the signal fractions more precisely. This behaviour is expected as the relative amount of signal is larger in that case. The behaviour of the significance versus the reconstruction efficiency following table 7.2 is shown in figure 7.3 for both signals. The background yield corresponding to the reconstruction efficiency is denoted as well. Fits that were considered not reliable are denoted with small markers.

The significance of the signals shows similar behaviour. A sharp increase in the significance is observed for larger ϵ_{rec}. Even for $\epsilon_{rec} \geq 88\%$ the significance is larger than 5σ. In fact, the only toy study that shows a significance below 5σ is the B_c^+ result for $\epsilon_{rec} = 84\%$. Note however that this result was considered to be unreliable.

7.2.2 Likelihood fit as crosscheck

This section discusses the comparison of two likelihood fits, in order to understand if MVA3 adds information to the fit with respect to MVA1. For this purpose, the first fit is done using (BDT1, m_{corr}); the second fit is done using (BDT3, m_{corr}). In both fits only the B_c^+ and B^+ components are used. The comparison is done for $\epsilon_{rec} = 88\%$.

The likelihood fit in section 7.2.1 used signal fractions as free parameters, because the fit did not converge when using absolute yields. In this case, fit convergence is achieved when using absolute yields. This may be because in these *crosscheck* fits, the number of components is two, one less compared to the previous fit. Besides using absolute yields as free parameters, the fit procedure is exactly the same as in section 7.2.1.

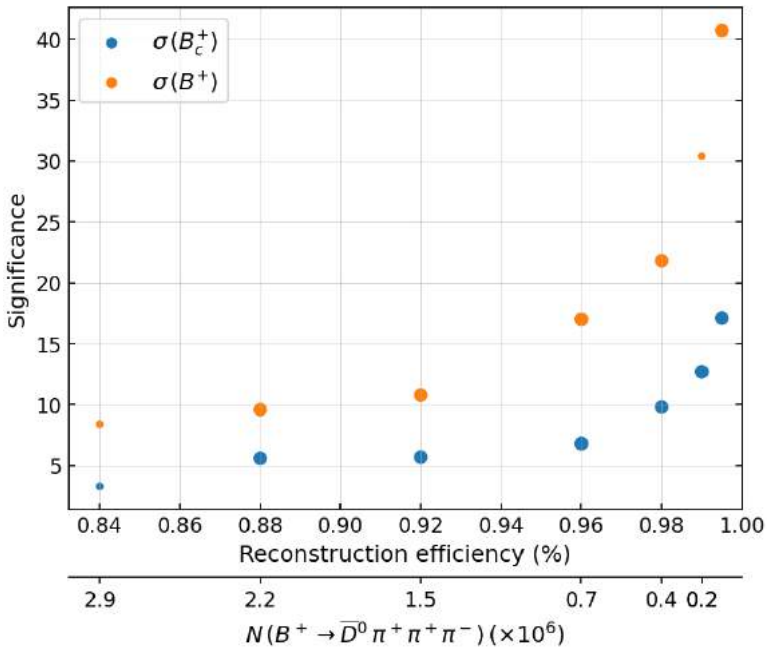

Figure 7.3: The significance of observing $B_c^+ \to \tau^+\nu_\tau$ (blue) or $B^+ \to \tau^+\nu_\tau$ (orange) versus the reconstruction efficiency in the worst-case scenario, i.e. $\mathcal{BR}(B^+ \to \overline{D}^0\pi^+\pi^+\pi^-) = 1$. The background yield corresponding to the particular ϵ_{rec} is also shown. Fits that were considered not reliable are denoted with small markers.

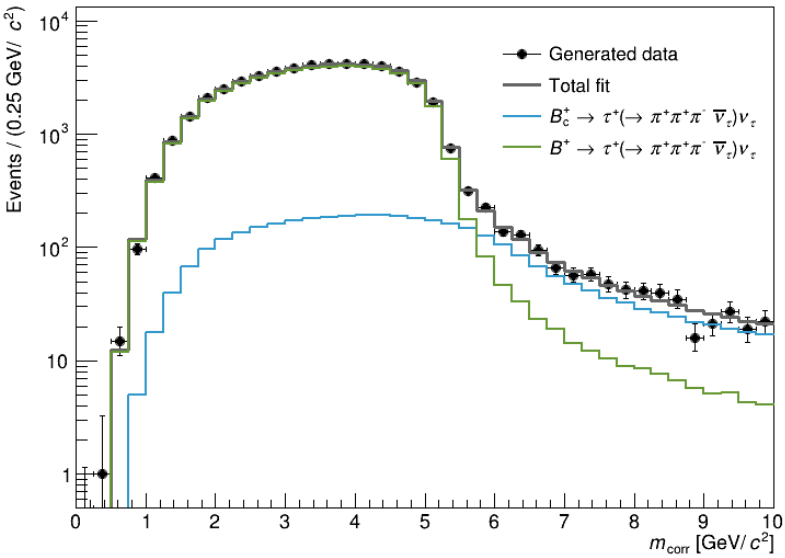

Figure 7.4: The projection of the two-dimensional likelihood fit on the corrected mass for $\epsilon_{rec} = 88\%$, using B_c^+ and B^+ components. The absolute background yield corresponds to an integrated luminosity of 10 fb^{-1}.

Table 7.3: Signal yields and their uncertainties obtained from two different likelihood fits, for a reconstruction efficiency of $\epsilon_{rec} = 88\%$.

Crosscheck fit	$N_{B_c^+} \pm \sigma_{B_c^+}$	$N_{B^+} \pm \sigma_{B^+}$	$\sigma_{B_c^+}/N_{B_c^+}$ (%)	σ_{B^+}/N_{B^+} (%)
$(\mathrm{BDT1}, m_{corr})$	3381 ± 96	48503 ± 233	2.8	0.48
$(\mathrm{BDT3}, m_{corr})$	3520 ± 95	48401 ± 232	2.7	0.48

The projection of the fit on the corrected mass is shown in figure 7.4 for $\epsilon_{rec} = 88\%$. Again, the absolute yield in bins of corrected mass for an integrated luminosity of 10 fb^{-1} is shown. Pseudo-data is generated based on the total PDF, $\mathcal{P}_{B_c^+} + \mathcal{P}_{B^+}$. The best fit estimate is searched for using this pseudo-data. The correlation between $\mathcal{P}_{B_c^+}$ and $\mathcal{P}_{B^+}$ is lower than 26% for both fits and therefore low enough to avoid minimization problems.

Results on the pull are listed in table B.2 of appendix B. The B_c^+-result is reliable for both fits; the B^+-result is reliable for the $(\mathrm{BDT1}, m_{corr})$ fit but not reliable for the $(\mathrm{BDT3}, m_{corr})$ fit. For the reliable fits, the best estimate of the fit is obtained by using the average values of the Gaussian fit of the yield and uncertainty.

The fit results together with their relative uncertainty $\sigma_{B_{(c)}^+}/N_{B_{(c)}^+}$ are shown in table 7.3. Only conclusions can be drawn on the B_c^+-results, as the B^+-results for the $(\mathrm{BDT3}, m_{corr})$ fit was shown to be not reliable. For that reason, the following only holds for the B_c^+ results. The two fits show very comparable results, meaning that MVA1 is able to separate B_c^+ from B^+ similarly to MVA3. For that reason, a likelihood fit in the $(\mathrm{BDT1}, \mathrm{BDT3}, m_{corr})$ space is expected to give similar results as the fit in the $(\mathrm{BDT1}, m_{corr})$ space.

Chapter 8

Discussion

This feasibility study aimed to identify significant background decays mimicking the signals through their final state. For that purpose, a lot of simplifying assumptions have been made. This is reasonable for a feasibility study but a revision should be reconsidered once it turns into a real analysis. The thesis especially built towards the results presented in chapter 6 and 7. This chapter gives a discussion on these results and proposes different improvements for further research.

8.1 Selection

The B-tracking efficiency is an important factor in estimating the yields and therefore relevant to examine in more detail. In this thesis, the B-tracking tool is approximated by a simplistic view of the VELO-detector, as explained in section 4.3. As figure 2.4 shows, this approximation is only true for the region close to the PV. The model for determining the number of VELO hits and thus estimating the B-tracking efficiency, can be improved by using a more realistic distribution of VELO modules. This may lead to different systematic effects for different background types, leading to more accurate results.

8.2 Backgrounds

Dangerous backgrounds were considered to meet two requirements: to have high yields and to have a close match of the corrected mass distribution compared to the signal. There may be backgrounds with equally large yields as $B^+ \to \overline{D}^0 \pi^+\pi^+\pi^-$, but with even more correlation between the signal shape. For backgrounds that have more correlation, the likelihood fit may result in lower significance to observe the signals. Therefore, for completeness, some backgrounds are listed below that can be looked into more closely.

Firstly, there can be looked into direct charmonium decays of the form $B^+ \to J/\psi(1S)\,\pi^+\pi^+\pi^- X$ where X and $J/\psi(1S)$ are not reconstructed. B-mesons may also decay into baryon modes; in particular the Λ_c^+-baryon may be of interest as

it has a similar mass and decay width as the τ-lepton [23]. This study has not considered these modes as typical branching fractions are smaller with respect to $B^+ \to \overline{D}^0 \pi^+\pi^+\pi^-$. However, they may still contribute significantly to the yield, depending on their B-tracking efficiency.

Furthermore, the background list is not limited to using the B^+-meson as mother particle. In fact, following the FCC-ee study [15], similar backgrounds exist where the B^+ is replaced by the neutral b-hadron $\in \{B^0, B_s^0, \Lambda_b^0\}$. Although these neutral particles do not leave any VELO-hits, they still can contribute to the background yield. This is because the VELO is not able to identify the particle that left a hit. Hits from decay products of B, like a τ^+ or π^+, may appear for the trigger as a hit from a B-candidate.

Finally, obvious candidates for potentially dangerous backgrounds are decays where the B_c^+ mother is involved. Examples are $B_c^+ \to J/\psi(1S)\,\tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$ and $B_c^+ \to J/\psi(1S)\,\pi^+\pi^+\pi^-$. The branching fractions of both are of comparable size with the signal decay [44, 45] and therefore of interest for further studies.

8.3 Multivariate analysis

The MVA presented in chapter 6 uses the observables mentioned in section 4.3. Although in the end the fit result is most important, a short comment is made about possible improvements on the MVA. Firstly, the invariant mass of the pions can be a distinguishing feature for certain backgrounds. For the signal decays, the invariant mass is bound from above by the τ-mass, while this is not the case for *direct B-decays*. For that reason, a cut in this variable may reduce certain backgrounds significantly. In addition, Dalitz plots can be used to identify signal decays based on their resonance structures. Extending the RapidSim framework with TAUOLA was outside the scope of this project, but may lead to a gain in MVA performance in further work.

8.4 Likelihood fit

The likelihood fit using $B^+ \to \overline{D}^0 \pi^+\pi^+\pi^-$ in the worst-case scenario as presented in chapter 7, shows positive results. For the reliable fits, the significance to observe $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$, assuming SM branching fractions, has shown to be larger than 5.6σ and 9.6σ, respectively. The associated relative uncertainties[1] are 17.7% and 10.2%, respectively. Comparing this with the relative uncertainty of 23% obtained from Belle and Babar, this is very promising.

The reconstruction efficiencies from section 4.2.2 for charged and neutral particles were estimated to be 95% and 90%, respectively. Considering the fact that the significance of both signals reported above correspond to $\epsilon_{rec} = 88\%$, it

[1] In the current assumptions, the relative uncertainty is the inverse of the significance.

is argued that an observation of $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ is feasible, given the current assumptions.

The large significances that are obtained are remarkable, especially considering the enormous amount of background yield that is used in the fit. Although some fits were considered unreliable, a clear trend can be seen in figure 7.3. The absence of clear outliers and the similarity between the B_c^+ and B^+ results, suggests that the overall behaviour of the significance is accurate. However, before hard conclusions can be drawn on the fit results, we must first understand where the large significance comes from.

The amount of $B^+ \to \overline{D}^0\,\pi^+\pi^+\pi^-$ background in the fit is well illustrated in figure 7.2. As this is a projection of the PDFs on m_{corr}, it does not accurately resemble the true signal-to-background ratio in each bin. In fact, in certain bins the amount of signal can actually become significant compared to the amount of background. This ratio would naturally increase for larger ϵ_{rec}, as is shown in figure 8.1 for $\epsilon_{rec} = 99.5\%$. This figure shows that especially for large m_{corr}, signal-to-background ratios become larger. Possibly, the optimization of the fit may be mainly based on the bins having a large ratio. This may explain the large significance that is observed for both signals in table 7.2, especially for large ϵ_{rec}. Further research can look at the effect of the corrected mass range on the fit results and show whether the above reasoning is correct.

In section 7.2, the reliability of the fits was checked using the pull distribution. A fit result was considered reliable if its mean and standard deviation were compatible within 1.5σ. Besides using the pull, the fit result should be compatible

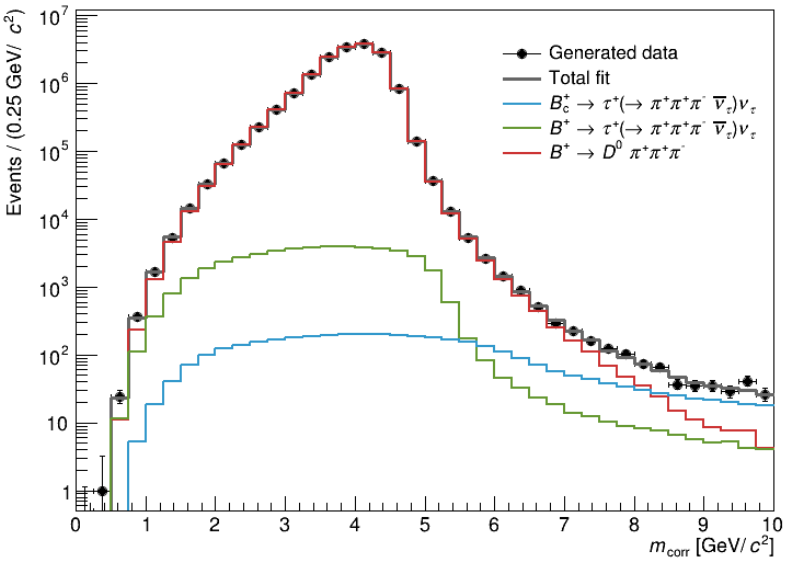

Figure 8.1: The projection of the two-dimensional likelihood fit on the corrected mass for $\epsilon_{rec} = 99.5\%$. The absolute background yield corresponds to an integrated luminosity of 10 fb^{-1} and $\mathcal{BR}(B^+ \to D^0\,\pi^+\pi^+\pi^-) = 1$.

Table 8.1: The compatibility of the fit results with the estimated (input) signal fractions in terms of standard deviations, for different reconstruction efficiencies. Not compatible results are shown in red.

ϵ_{rec}	$B_c^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$	$B^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$
84 %	1.28σ	0.10σ
88 %	0.42σ	0.18σ
92 %	0.50σ	0.70σ
96 %	1.62σ	1.00σ
98 %	1.31σ	0.49σ
99 %	0.96σ	0.23σ
99.5 %	1.18σ	0.66σ

as well with the SM prediction on the signal fractions, which were shown in table 7.1. The compatibility (also within 1.5σ) is reported in table 8.1. The B_c^+ fraction is not compatible with its estimated fraction for $\epsilon_{rec} = 96\%$; the B^+ fraction is compatible with its estimated fraction for all efficiencies.

For the purpose of a feasibility study, it is no problem that the observed discrepancies are compatible within 1.5σ, as opposed to 1σ. It is however notable that quite a number of fits show a discrepancy larger than 1σ. Especially, if we consider the fact that only two free parameters are used and that the PDFs are histograms, i.e. independent of fit parameters.

An additional likelihood fit may be done using MVA2, the MVA trained on $B^+ \to \tau^+\nu_\tau$ and background data. Additional MVAs can be trained using different backgrounds, for example with the significant charm contribution. Then, higher dimensional fits can be performed, using the corresponding BDT-scores with m_{corr}. Crosschecks may be done to understand if certain BDT-scores do add any value to the existing fit.

Chapter 9

Conclusion

Our understanding of the building blocks of the Universe is summarized in the Standard Model of particle physics. The absence of gravity and hints of the existence of different types of matter lead to the general consensus that the framework is far from complete. The model predicts that Nature treats the charged lepton flavours identically, accounting for their mass difference and couplings. Hints to deviations of this lepton flavour universality are observed in various measurements. The yet unobserved $B_c^+ \to \tau^+\nu_\tau$ decay may provide additional insight from an independent point of view. This work extends on the study of Ref. [9], where the novel B-tracking tool has been proven worthwile for the analysis of $B_c^+ \to \tau^+\nu_\tau$ and of $B^+ \to \tau^+\nu_\tau$.

In particular, this thesis has looked into the feasibility of measuring the branching fractions of $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ at the LHCb detector at CERN. Both signals are reconstructed using the $\tau^+ \to \pi^+\pi^+\pi^-\nu_\tau$ mode. Besides the signal decays, there exist many decay modes that produce three pions as well but do not originate from the signal decay chain. A measurement is only feasible if the signal decays are not dominated by unwanted backgrounds. For that purpose, an exhaustive list of beauty backgrounds is presented.

A simplified but worst-case scenario is considered where the potentially most dangerous background is assumed to occur in 100% of the cases. This background is identified using its expected yield and corrected mass distribution, which is simulated using the RapidSim framework. Chapter 5 concluded that $B^+ \to \overline{D}^0 \pi^+\pi^+\pi^-$ is the most dangerous of the considered backgrounds. Charm backgrounds show similar yields but are much more distinct form signal in the corrected mass variable.

The feasibility to measure both signals is examined in this worst-case scenario. In chapter 6, different multivariate analyses have been trained to classify signal from background events. MVA1 was trained on B_c^+-signal versus background; MVA2 was trained on B^+-signal versus background; MVA3 was trained on B_c^+- versus B^+-signal. All MVAs were shown to have high discriminatory power between both datasets, especially for MVA1. The IP and $p_{T_{3\pi}}$ were found to be the most distinguishing variables between B_c^+-signal and background.

In chapter 7, a likelihood fit was performed using BDT1 together with m_{corr}. There is clear trend in the significance of the fit results versus the background yield. For all reliable fits, the significance to observe $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ was shown to be larger than 5.6σ and 9.6σ, respectively. Both exceed the 5σ limit, meaning that **measurements on the branching fractions of both signals are feasible in the worst-case**. A crosscheck fit has shown that adding information of BDT3 is unlikely to improve the significance obtained.

The fit results are promising, however we should be careful in drawing firm conclusions. The high significance for both signals may be due to the high signal-to-background ratios in certain bins, suggesting that the fit results are highly dependent on the corrected mass range used. Also, it is noteworthy that some fits, even for lower background yields, show discrepancies larger than 1σ in the pull distributions. Especially when you consider that only two free parameters are used in the fit. Future research benefits from searching for additional backgrounds, which may reveal corrected mass distributions that are even more signal-like. In addition, considering the resonance structures of the τ-decay may improve the signal classification considerably.

Appendix A

Potentially dangerous backgrounds and their yields

In chapter 5, various backgrounds of the form $B \to D\,X$ are considered, where $D \in \{D_s^+, D^+, D^0\}$ decays into at least three charged pions and where X is not reconstructed. Table A.1, A.2 and A.3 compare the relative contributions of the most occurring D_s^+, D^+ and D^0 modes, respectively. The values for the efficiencies are assumed following the discussion of section 4.2.2 and 4.2.3. B-tracking efficiencies are assumed to be equal for the considered modes. To simplify the analysis, only the largest contribution (shown in red) is assumed to occur.

Table A.1: Listed are possible decay modes of the intermediate D_s^+ into at least three charged pions, where not reconstructed mesons are denoted in brackets. These modes contribute to **intermediate charm** backgrounds, as defined in chapter 5. Relative contributions are compared using the reconstruction efficiency ϵ_{rec}, misidentification efficiency ϵ_{mis} and branching fraction $\mathcal{BR}$. The largest contribution is shown in red.

D_s^+ channel	$\mathcal{BR}$	$1-\epsilon_{rec}(\%)$	$\epsilon_{mis}(\%)$	$\mathcal{BR}\cdot(1-\epsilon_{rec})\cdot\epsilon_{mis}$
Partially reconstructed				
$D_s^+ \to 2\pi^+\pi^-(\pi^+\pi^-\pi^0)$	$4.9\cdot10^{-2}$	5	-	$2.5\cdot10^{-3}$
$D_s^+ \to 2\pi^+\pi^-(K^+K^-)$	$8.3\cdot10^{-3}$	5	-	$4.3\cdot10^{-4}$
$D_s^+ \to 2\pi^+\pi^-(\pi^+\pi^-)$	$7.9\cdot10^{-3}$	5	-	$4.0\cdot10^{-4}$
$D_s^+ \to 2\pi^+\pi^-(K_S^0)$	$2.8\cdot10^{-3}$	10	-	$2.8\cdot10^{-4}$
$D_s^+ \to 2\pi^+\pi^-(2K_S^0)$	$8.4\cdot10^{-4}$	10	-	$8.4\cdot10^{-5}$
Misidentified				
$D_s^+ \to K^+\pi^+\pi^-$	$6.5\cdot10^{-3}$	-	5	$2.0\cdot10^{-4}$
$D_s^+ \to K^+K^-\pi^+$	$5.4\cdot10^{-2}$	-	0.25	$1.4\cdot10^{-4}$
Other[1]				
$D_s^+ \to \tau^+\nu_\tau$	$5.1\cdot10^{-3}$	-	-	$5.1\cdot10^{-3}$

[1] Here, $\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau$ is assumed.

Table A.2: Listed are possible decay modes of the intermediate D^+ into at least three charged pions, where not reconstructed mesons are denoted in brackets. These modes contribute to **Semileptonic, $(q\bar{q}) \rightarrow 3\pi$** backgrounds, as defined in chapter 5. Relative contributions are compared using the reconstruction efficiency ϵ_{rec}, misidentification efficiency ϵ_{mis} and branching fraction $\mathcal{BR}$. The largest contribution is shown in red.

D^+ channel	$\mathcal{BR}$	$1-\epsilon_{rec}(\%)$	$\epsilon_{mis}(\%)$	$\mathcal{BR}\cdot(1-\epsilon_{rec})\cdot\epsilon_{mis}$
Partially reconstructed				
$D^+ \rightarrow 2\pi^+\pi^-(K^0_S)$	$3.1\cdot10^{-2}$	10	-	$3.1\cdot10^{-3}$
$D^+ \rightarrow 2\pi^+\pi^-(\pi^0)$	$1.2\cdot10^{-2}$	10	-	$1.2\cdot10^{-3}$
$D^+ \rightarrow 2\pi^+\pi^-(K^-\pi^+)$	$5.7\cdot10^{-3}$	5	-	$2.9\cdot10^{-4}$
$D^+ \rightarrow 2\pi^+\pi^-(\pi^+\pi^-)$	$1.7\cdot10^{-3}$	5	-	$8.3\cdot10^{-5}$
$D^+ \rightarrow 2\pi^+\pi^-(K^+K^-)$	$2.3\cdot10^{-4}$	5	-	$1.2\cdot10^{-5}$
Misidentified				
$D^+ \rightarrow K^-\pi^+\pi^+$	$9.4\cdot10^{-2}$	-	5	$4.7\cdot10^{-3}$
$D^+ \rightarrow K^+K^-\pi^+$	$9.7\cdot10^{-3}$	-	0.25	$2.4\cdot10^{-5}$
Other[1]				
$D^+ \rightarrow \tau^+\nu_\tau$	$1.1\cdot10^{-4}$	-	-	$1.1\cdot10^{-4}$

[1] Here, $\tau^+ \rightarrow \pi^+\pi^+\pi^-\overline{\nu}_\tau$ is assumed.

Table A.3: Listed are possible decay modes of the intermediate D^+_s into at least three charged pions, where not reconstructed mesons are denoted in brackets. These modes contribute to **Semileptonic, $(q\bar{q}) \rightarrow 3\pi$** backgrounds, as defined in chapter 5. Relative contributions are compared using the reconstruction efficiency ϵ_{rec}, misidentification efficiency ϵ_{mis} and branching fraction $\mathcal{BR}$. The largest contribution is shown in red.

D^0 channel	$\mathcal{BR}$	$1-\epsilon_{rec}(\%)$	$\epsilon_{mis}(\%)$	$\mathcal{BR}\cdot(1-\epsilon_{rec})\cdot\epsilon_{mis}$
Partially reconstructed				
$D^0 \rightarrow 2\pi^+\pi^-(K^-)$	$8.2\cdot10^{-2}$	5	-	$4.1\cdot10^{-3}$
$D^0 \rightarrow 2\pi^+\pi^-(K^-\pi^0)$	$4.3\cdot10^{-2}$	5	-	$2.2\cdot10^{-3}$
$D^0 \rightarrow 2\pi^+\pi^-(\pi^-)$	$7.6\cdot10^{-3}$	5	-	$3.8\cdot10^{-4}$
$D^0 \rightarrow 2\pi^+\pi^-(\pi^-\pi^0)$	$4.2\cdot10^{-3}$	5	-	$2.1\cdot10^{-4}$
$D^0 \rightarrow 2\pi^+\pi^-(K^0_S\pi^-)$	$2.7\cdot10^{-3}$	5	-	$1.3\cdot10^{-4}$
Misidentified				
$D^0 \rightarrow K^-\pi^+\pi^-(e^+\nu_e)$	$2.8\cdot10^{-4}$	5	5	$7.0\cdot10^{-7}$
$D^0 \rightarrow K^+K^-\pi^+(\pi^-)$	$2.5\cdot10^{-3}$	5	0.25	$3.1\cdot10^{-7}$

The signals together with the largest decay modes in each background category as defined in chapter 5 are listed in table A.4. The yields are estimated using equation 5.1 where the used values for the efficiencies follow the discussion in chapter 4.

Table A.4: Listed are the signal decays and potentially dangerous background channels considered in this study. The estimated yield N/fb^{-1} is calculated following equation 5.1 using the B-meson production cross section σ_B, branching fraction $\mathcal{BR}$, reconstruction efficiency ϵ_{rec} and B-tracking efficiency $\epsilon_{tracking}$.

Channel[1]	σ_B ($\cdot 10^9$ fb)	$\mathcal{BR}$	$1-\epsilon_{\text{rec}}(\%)$	$\epsilon_{tracking}(\%)$	N/fb^{-1}
		Signals			
$B_c^+ \to \tau^+\nu_\tau$	0.65	$1.82 \cdot 10^{-3}$	-	0.0290	$3.43 \cdot 10^2$
$B^+ \to \tau^+\nu_\tau$	87	$1.015 \cdot 10^{-5}$	-	0.549	$4.85 \cdot 10^3$
		Charm			
$D^+ \to \tau^+\nu_\tau$	834	$1.12 \cdot 10^{-4}$	-	0.185	$1.73 \cdot 10^5$
$D_s^+ \to \tau^+\nu_\tau$	353	$5.10 \cdot 10^{-3}$	-	0.0190	$3.42 \cdot 10^5$
		Semileptonic $\tau \to 3\pi$			
$B^+ \to \tau^+\nu_\tau\, \overline{D}^0$	87	$7.17 \cdot 10^{-4}$	10	0.497	$3.10 \cdot 10^4$
$B^+ \to \tau^+\nu_\tau\, \overline{D}^{*0}$	87	$1.75 \cdot 10^{-3}$	10	0.490	$7.46 \cdot 10^4$
$B^+ \to \tau^+\nu_\tau\, \overline{D}^{**0}$	87	$1.75 \cdot 10^{-3}$	10	0.480	$7.31 \cdot 10^4$
		Intermediate charm[2]			
$B^+ \to D_s^+\overline{D}^0$	87	$4.592 \cdot 10^{-5}$	10	0.871	$3.48 \cdot 10^3$
$B^+ \to D_s^{*+}(\to D_s^+\gamma)\, \overline{D}^0$	87	$3.88 \cdot 10^{-5}$	10	0.831	$2.81 \cdot 10^3$
$B^+ \to D_s^+\overline{D}^{*0}$	87	$4.182 \cdot 10^{-5}$	10	0.860	$3.13 \cdot 10^3$
$B^+ \to D_s^{*+}(\to D_s^+\gamma)\, \overline{D}^{*0}$	87	$8.72 \cdot 10^{-5}$	10	0.812	$6.16 \cdot 10^3$
$B^+ \to D_s^+(\to D_s^+\gamma)\overline{D}^{**0}$	87	$1.38 \cdot 10^{-4} \times 1/3$	10	0.797	$3.19 \cdot 10^3$
$B^+ \to D_s^{*+}(\to D_s^+\gamma)\overline{D}^{**0}$	87	$1.38 \cdot 10^{-4} \times 2/3$	10	0.766	$6.13 \cdot 10^3$
		Direct B-decay			
$B^+ \to \overline{D}^0\, \pi^+\pi^+\pi^-$	87	$5.6 \cdot 10^{-3}$	10	0.378	$1.84 \cdot 10^5$
$B^+ \to \overline{D}^{*0}\, \pi^+\pi^+\pi^-$	87	$1.03 \cdot 10^{-2}$	10	0.377	$3.38 \cdot 10^5$
$B^+ \to \overline{D}^{*0}\, \pi^+\pi^+\pi^-\, \pi^0$	87	$1.8 \cdot 10^{-2}$	10	0.378	$5.92 \cdot 10^5$
		Semileptonic[3] $(q\overline{q}) \to 3\pi$			
$B^+ \to \overline{D}^0\, e^+\nu_e$	87	$1.68 \cdot 10^{-4}$	5	0.378	$2.76 \cdot 10^3$
$B^+ \to \overline{D}^0\, \mu^+\nu_\mu$	87	$1.68 \cdot 10^{-4}$	5	0.375	$2.74 \cdot 10^3$
$B^+ \to \overline{D}^{*0}(\to \overline{D}^0\gamma)\, e^+\nu_e$	87	$4.04 \cdot 10^{-4}$	5	0.380	$6.68 \cdot 10^3$
$B^+ \to \overline{D}^{*0}(\to \overline{D}^0\gamma)\, \mu^+\nu_\mu$	87	$4.04 \cdot 10^{-4}$	5	0.373	$6.56 \cdot 10^3$
$B^+ \to \overline{D}_2^*(2460)^0 e^+\, \nu_e$	87	$1.54 \cdot 10^{-4}$	5	1.18	$7.90 \cdot 10^3$
$B^+ \to \overline{D}_2^*(2460)^0\, \mu^+\nu_\mu$	87	$1.54 \cdot 10^{-4}$	5	1.19	$7.97 \cdot 10^3$

[1] All tau-leptons are assumed to decay via $\tau^+ \to \pi^+\pi^+\pi^-\overline{\nu}_\tau$

[2] Signal-like pions are reconstructed via $D_s^+ \to \tau^+(\to \pi^+\pi^+\pi^-\overline{\nu}_\tau)\nu_\tau$

[3] Signal-like pions are reconstructed via the $\overline{D}^0 \to K^+\pi^-\pi^-\pi^+$ and $\overline{D}_2^*(2460)^0 \to D^-(\to K_S^0\,\pi^-\pi^-\pi^+)\,\pi^+$ decays.

Appendix B

Reliability of the likelihood fits

This appendix reports on the reliability of the likelihood fits that are done in chapter 7. In section 7.2.1, a two-dimensional likelihood fit is performed using BDT1 and m_{corr} for different reconstruction efficiencies. A fit is only considered reliable if the mean and standard deviation of the resulting pull distribution is respectively compatible with 0 and 1, within 1.5σ. The results are shown in table B.1. Non-reliable results are shown in red.

Table B.1: The compatibility of the mean and standard deviation of the pull distribution with 0 and 1, corresponding to the likelihood fit using BDT1 and m_{corr} in section 7.2.1. The results are shown in terms of standard deviations, for different reconstruction efficiencies. Non-reliable fits, exceeding 1.5σ, are shown in red.

	B_c^+-result		B^+-result	
ϵ_{rec}	pull mean	pull sigma	pull mean	pull sigma
84 %	0.17σ	3.81σ	3.05σ	0.96σ
88 %	0.85σ	0.52σ	0.75σ	0.17σ
92 %	1.19σ	0.09σ	0.99σ	0.91σ
96 %	1.11σ	0.32σ	0.22σ	0σ
98 %	1.10σ	1.36σ	0.50σ	0.50σ
99 %	1.01σ	0.65σ	1.15σ	2.42σ
99.5 %	1.06σ	0.27σ	0.76σ	0.18σ

Table B.2: The compatibility of the mean and standard deviation of the pull distribution with 0 and 1, corresponding to two different likelihood fits, for a reconstruction efficiency of $\epsilon_{rec} = 88\%$. The results are shown in terms of standard deviations, for different reconstruction efficiencies. Non-reliable fits, exceeding 1.5σ, are shown in red.

	B_c^+-result		B^+-result	
Crosscheck fit	pull mean	pull sigma	pull mean	pull sigma
(BDT1, m_{corr})	1.13σ	0.65σ	0.39σ	1.43σ
(BDT3, m_{corr})	0.13σ	0.27σ	2.15σ	2.04σ

In section 7.2.2 a crosscheck fit using the BDT3-score is done for $\epsilon_{rec} = 88\%$, to understand if this test-statistic adds any information to the separation power of BDT1. Table B.2 shows the results on the compatibility of the corresponding pull distribution.

References

[1] S. K. Bose, *The atomic hypothesis*, 2015. doi: 10.48550/ARXIV.1503.01792.

[2] LHCb, R. Aaij *et al.*, *Search for lepton-universality violation in* $B^+ \to K^+\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **122** (2019), no. 19 191801, arXiv:1903.09252.

[3] LHCb, R. Aaij *et al.*, *Test of lepton universality with* $B^0 \to K^{*0}\ell^+\ell^-$ *decays*, JHEP **08** (2017) 055, arXiv:1705.05802.

[4] ATLAS, CMS, LHCb, E. Graverini, *Flavour anomalies: a review*, J. Phys. Conf. Ser. **1137** (2019), no. 1 012025, arXiv:1807.11373.

[5] LHCb, R. Aaij *et al.*, *Measurement of the ratio of branching fractions* $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$, Phys. Rev. Lett. **120** (2018), no. 12 121801, arXiv:1711.05623.

[6] K. Freese, *Status of Dark Matter in the Universe*, Int. J. Mod. Phys. **1** (2017), no. 06 325, arXiv:1701.01840.

[7] Belle, B. Kronenbitter *et al.*, *Measurement of the branching fraction of* $B^+ \to \tau^+\nu_\tau$ *decays with the semileptonic tagging method*, Phys. Rev. D **92** (2015), no. 5 051102, arXiv:1503.05613.

[8] BaBar, J. P. Lees *et al.*, *Evidence of* $B^+ \to \tau^+\nu$ *decays with hadronic B tags*, Phys. Rev. D **88** (2013), no. 3 031102, arXiv:1207.0698.

[9] J. I. Rol, *b-Meson Tracking at LHCb: a Feasibility Study*, Master's thesis, University of Groningen, the Netherlands, March 2021.

[10] Super-Kamiokande, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.

[11] SNO, Q. R. Ahmad *et al.*, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv:nucl-ex/0204008.

[12] D. Galbraith and C. Burgard, *Example: Standard model of physics*, https://texample.net/tikz/examples/model-physics/. Accessed: 2022-06-07.

[13] LHCb, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, Nature Phys. **18** (2022), no. 3 277, arXiv:2103.11769.

[14] Belle-II, T. Abe *et al.*, *Belle II Technical Design Report*, arXiv:1011.0352.

[15] Y. Amhis *et al.*, *Prospects for $B_c^+ \to \tau^+\nu$ at FCC-ee*, JHEP **12** (2021) 133, arXiv:2105.13330.

[16] T. Zheng *et al.*, *Analysis of $B_c \to \tau\nu_\tau$ at CEPC*, Chin. Phys. C **45** (2021), no. 2 023001, arXiv:2007.08234.

[17] *LHC Machine*, JINST **3** (2008) S08001.

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

[19] LHCb, *LHCb Tracker Upgrade Technical Design Report*, .

[20] LHCb, I. Bediaga, *LHCb VELO Upgrade Technical Design Report*, .

[21] LHCb Velo, V. Coco *et al.*, *Velo Upgrade Module Nomenclature*, .

[22] R. Aaij *et al.*, *Allen: A high level trigger on GPUs for LHCb*, Comput. Softw. Big Sci. **4** (2020), no. 1 7, arXiv:1912.09161.

[23] Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020), no. 8 083C01.

[24] C. Méaux, *A phenomenological study of $D^+ \to \pi^+\ell^+\ell^-$ and the search for $B^0_{(s)} \to \tau^+\tau^-$ at LHCb*, PhD thesis, Marseille, CPPM, 2019.

[25] J. Serrano, *Study of rare leptonic B meson decays with the LHCb experiment*, PhD thesis, Marseille, CPT, 2019.

[26] P. Lichard, *Resonance $a_1(1420)$ and the Three-Pion Decays of the Tauon*, arXiv:1703.06315.

[27] G. A. Cowan, D. C. Craik, and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **214** (2017) 239, arXiv:1612.07489.

[28] S. Jadach, J. H. Kuhn, and Z. Was, *TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons*, Comput. Phys. Commun. **64** (1990) 275.

[29] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[30] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (2015), no. 07 1530022, arXiv:1412.6352.

[31] O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, .

[32] K. G. Sobczak, *Study of charmless three-body decays of neutral B mesons with the LHCb spectrometer*, PhD thesis, U. Blaise Pascal, Clermont-Ferrand, 2011.

[33] LHCb, R. Aaij *et al.*, *Measurement of the $B^{\pm}$ production cross-section in pp collisions at $\sqrt{s} = 7$ and 13 TeV*, JHEP **12** (2017) 026, arXiv:1710.04921.

[34] LHCb, R. Aaij *et al.*, *Measurement of the B_c^- meson production fraction and asymmetry in 7 and 13 TeV pp collisions*, Phys. Rev. D **100** (2019), no. 11 112006, arXiv:1910.13404.

[35] LHCb Collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2015) 159. 46 p, arXiv:1510.01707.

[36] CLEO, S. Ahmed *et al.*, *Measurements of $B \to D_s^{(*)+} D^{*(*)}$ branching fractions*, Phys. Rev. D **62** (2000) 112003, arXiv:hep-ex/0008015.

[37] Belle, G. Caria *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, Phys. Rev. Lett. **124** (2020), no. 16 161803, arXiv:1910.05864.

[38] F. Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python*, J. Machine Learning Res. **12** (2011) 2825, arXiv:1201.0490.

[39] G. Cowan, *Statistical data analysis*, 1998.

[40] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*, arXiv:1210.7141.

[41] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81.

[42] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, .

[43] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[44] W. Cheng *et al.*, *helicity form factors and the decays **, Chin. Phys. C **46** (2022), no. 5 053103, arXiv:2107.08405.

[45] CMS, V. Khachatryan *et al.*, *Measurement of the ratio of the production cross sections times branching fractions of $B_c^{\pm} \to J/\psi\pi^{\pm}$ and $B^{\pm} \to J/\psi K^{\pm}$ and $\mathcal{B}(B_c^{\pm} \to J/\psi\pi^{\pm}\pi^{\pm}\pi^{\mp})/\mathcal{B}(B_c^{\pm} \to J/\psi\pi^{\pm})$ in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **01** (2015) 063, arXiv:1410.5729.

Acknowledgements

In the following I would like to thank all the people who personally supported me during my master project. My enthusiasm for experimental particle physics has grown because of all of you, all people who are genuinely motivated to do science.

Kristof and Maarten, thank you for always being available for help, regarding the project but also for guiding me to the next step in my professional career. Kristof, your rapid and extensive feedback has helped me enormously, especially in the writing phase of my thesis. Maarten, I enjoyed our discussions and your passionate clarifications around the whiteboard as we thought about the next steps to take.

MD, we joined the group around the same time, which helped me a lot at the start of my project. It was a pleasure to work with you from the start, despite the largely online contact due to the corona measures. Your unique choice of using certain emoticons to digitally convey your expression often brought a big smile to my face.

Mick, two weeks after the start of my project I was about to participate in an online Spanish course. What are the odds that I will meet a fellow physicist in that same course, whom I have only known for two weeks? It was a pleasant experience. Thank you for your always genuine helpfulness. *Ha sido un placer trabajar contigo.*

Gerco, thank you for asking interesting questions during the group meetings. These have always led to lively discussions and helped me to make further progress.