university of groningen

faculty of science and engineering

An Investigation Into the Normalisation Channel for Measuring the Branching Fraction of $B_c^+ \to \tau^+ \nu_\tau$ at the LHCb

Master Thesis

June 24, 2024

Student: Max A. Wiegertjes

Supervisor: Mick Mulder, PhD

"Die richtige Methode der Philosophie wäre eigentlich die: Nichts zu sagen, als was sich sagen lässt, also Sätze der Naturwissenschaft also etwas, was mit Philosophie nichts zu tun hat, und dann immer, wenn ein anderer etwas Metaphysisches sagen wollte, ihm nachzuweisen, dass er gewissen Zeichen in seinen Sätzen keine Bedeutung gegeben hat. Diese Methode wäre für den anderen unbefriedigender hätte nicht das Gefühl, dass wir ihn Philosophie lehrten aber sie wäre die einzig streng richtige. (...) Wovon man nicht sprechen kann, darüber muss man schweigen."

- Ludwig Wittgenstein, (1921) *Tractus Logico-Philosophicus*

Abstract

In the Standard Model all lepton flavours share the same coupling constants. Violation of this property of lepton flavour universality may point towards New Physics. Previously, measurements of the branching fraction of decays of B mesons, such as reported in the $\mathcal{R}(D^{(*)})$ measurements by LHCb and several other research groups, have hinted towards such lepton flavour universality violations. If lepton flavour universality is violated, a measurement of the branching fraction $B_c^+ \to \tau^+\nu_\tau$ could provide evidence in favour of such phenomena. However, the measurement of this decay requires the use of a novel reconstruction method where direct hits in the Vertex Locator of the LHCb are used to gain information on the kinematics of the decay. To better understand this technique, called *B-tracking*, and to ultimately minimize uncertainty in the measurement values, the research in this thesis intends to investigate potential normalisation modes for the decay $B_c^+ \to \tau^+\nu_\tau$. By estimating factors such as the measurement yield, reconstruction efficiency, and several uncertainties associated to the candidate normalisation channels, the best normalisation channel can be determined. By analysing simulation samples for several candidate decays, a good normalisation channel for a measurement of $B_c^+ \to \tau^+\nu_\tau$ at LHCb was determined to be $B^+ \to D^-\pi^+\pi^+$. Using this normalisation channel a first measurement of the branching fraction of $B_c^+ \to \tau^+\nu_\tau$ can be performed, which may alleviate or strengthen tensions in the lepton flavour universality property of the Standard Model.

Samenvatting

De hedendaagse natuurkunde heeft ontdekt dat de wereld op het meest fundamentele niveau beschreven kan worden als het samenspel tussen verscheidene zogeheten velden. Het universum is in dezen een soort kosmische dans. Deze velden nemen een waarde (of meerdere waarden) aan voor elk punt in de ruimtetijd. De manier waarop dit gebeurt, volgt enkele fundamentele symmetrische principes. Deeltjes, de kleinste eenheden die bestaan in deze wereld, zijn trillingen in deze velden. Het Standaardmodel is de theorie die dit samenspel beschrijft. Al decennialang is dit een enorm succesvolle theorie die haast alle experimentele resultaten kan verklaren.

Natuurkundigen proberen op allerlei manieren gaten in het Standaardmodel te prikken door verschillende voorspellingen experimenteel te testen. Een onderdeel van het Standaardmodel is zogeheten Leptonsmaakuniversaliteit. Dit is het idee dat alle deeltjes van een bepaald type (leptonen) zich gelijk gedragen in de kosmische dans. In de afgelopen jaren zijn er bewijzen gevonden dat dit principe misschien niet altijd waar is. Echter, deze bewijzen zijn niet eenduidig. Om deze reden worden er experimenten bedacht om dit nog beter te testen. Het idee is dat een meting van het verval van bepaalde deeltjes hierover uitsluitsel zou kunnen geven.

Het verval waarin we geïnteresseerd zijn, is het verval van de B_c^+ meson (bestaande uit een *beauty*- en een *charm*-quark) naar een τ^+ (het zwaardere broertje van het elektron) en een corresponderend neutrino. Om dit verval te identificeren in de LHCb-detector is een nieuwe methode bedacht, genaamd *B-tracking*. Om de meetefficiëntie van deze methode in te kunnen schatten en om de systematische onzekerheden voor deze meting te minimaliseren, wordt er gekozen om de meting uit te voeren ten opzichte van een ander verval, de zogeheten normalisatiemodus. Het doel van het onderzoek, beschreven in deze scriptie, is om erachter te komen welk verval dat dan zou moeten zijn en hoe dat het experiment verder gaat helpen. Hiervoor heb ik vooral naar simulatiedata gekeken om te begrijpen wat het effect is van B-tracking voor verscheidene vervallen.

Contents

1 Introduction

In modern physics, the existence of an intricate connection between symmetry and the fundamental physics of the universe is well understood. The three fundamental forces of particle physics each have an associated gauge symmetry, expressed as a Lie Group. $SU(3)$ for the strong interaction, $SU(2)$ for the weak interaction, and $U(1)$ for electromagnetism. Under high-energy conditions the weak and electromagnetic interactions are unified into an electroweak interaction, with a symmetry group of $SU(2) \times U(1)$. It is further hypothesized that in the very early universe, all three forces were unified under a single larger symmetry, a so-called Grand Unified Interaction.

The theoretical framework that derives these interactions from symmetric principles, and describes the way that they govern the behaviour of matter particles is called the Standard Model of particle physics (SM). The SM provides a near-complete description of reality on the smallest scales that can be effectively probed with modern technology. The SM has successfully predicted the existence of several particles.

However, a number of prima facie shortcomings of the SM have been noted. The SM provides no particle for the gravitational force nor has it any definitive dark matter particles. Furthermore, under the assumption of a zero baryon number early universe, it cannot explain the apparent matter-antimatter asymmetry observed. This provides a motivation for theoreticians to look beyond the SM, formulate extensions to it, or to construct even more fundamental theories. An extensive number of such theories exist, with many differing corresponding predictions.

It is up to experimental physicists then, to perform high precision measurements of physical parameters, in order to test them against the SM. Strong evidence in disagreement with the SM can be the starting point towards physics beyond the SM. Colloquially, such finding are called New Physics (NP). It is the search for NP then, that forms the motivation for the research that this thesis intends to contribute to.

An important instrument in testing the SM is CERN's Large Hadron Collider (LHC). This is the largest particle accelerator ever built, and according to some metrics the largest man-made machine ever built. The LHC consists of a highly controlled 27 km long beam pipe, with a number of detectors placed along it. One of these detectors belongs to the LHCb research group, which is specialised in measuring some of the heaviest particles produced in the LHC. Decays involving b quarks are of particular interest. The LHCb group is a large international collaboration of more than 1700 people spanning 100 scientific institutes and 22 countries [1]. The bfys (b physics) group of the Dutch National Institute for Subatomic Physics (Nikhef) is the seat of the Dutch contribution towards the LHCb, and is itself a collaboration between the University of Amsterdam (UvA), the Vrije Universiteit of Amsterdam (VU), Maastricht University (UM), and the Rijksuniversiteit Groningen (RUG).

The Groningen LHCb group is currently working towards a first measurement of the decay $B_c^+ \to \tau^+\nu_\tau$[1] in order to test the Standard Model of particle physics, and in hopes of finding New Physics. In order to reduce systematic uncertainty, this decay will be measured relative to some other decay: the normalisation channel. The goal of this thesis is **to investigate possible normalisation modes for the measurement of the branching fraction of $B_c^+ \to \tau^+\nu_\tau$, and to select the best candidate.**

[1] In this thesis the charge conjugate of all particles will always be implicitly included, unless stated otherwise.

Thesis Outline

The contents of this thesis have been divided into several sections. While each section addresses its own topic, the idea is that all sections are related to answering the question of which normalisation mode to select. This thesis is concerned with different kinds of simulated data, as well as with real data collected at the LHCb detector. The following is an outline of how this thesis is structured:

First, the context of this research within the current state of physics is sketched in Section 2. The SM will be introduced (Sec. 2.1), and the concept of Lepton Flavour Universality and how this relates to the decay $B_c^+ \to \tau^+\nu_\tau$ will be explained (Sec. 2.2). This will be followed by a description of the LHCb detector, and how data is processed at the LHCb (Sec. 3). Details on how and what types of simulated data samples are produced and their context within this research are discussed in Section 3.3. Next, Section 4 will cover B-tracking, the novel method that is utilised to perform the first measurement of the branching fraction of $B_c^+ \to \tau^+\nu_\tau$.

After all the preliminaries have been detailed, the main analyses of this thesis will be discussed. Section 5 is concerned with selecting the potential normalisation modes for further analysis. A list of candidates is compiled based on specific selection criteria. Previous data gathered at LHCb, as well as simulated data samples are discussed to come to the conclusion that there are two realistic normalisation mode channels: $B^+ \to D^-\pi^+\pi^+$ and $B^+ \to \tau^+\nu_\tau$.

These two decays are then further investigated in Section 6. Simulated data samples originating from two different simulation methods are compared, and for both methods several efficiencies are determined (Sec. 6.2). Furthermore, the trigger lines (algorithms for selecting decays) for these decays are discussed and tested (Sec. 6.3), and several relevant efficiencies are calculated (Sec. 6.4). Based on these results, estimates of the uncertainties in one of the final measurement values can be calculated.

The trigger lines for these decays were implemented early 2024 and are now actively selecting data. In Section 7 a small data sample of collected data was briefly analysed to validate their performance. Some early observations based on this data could already be made. Beyond these sections, the conclusions that can be made based on these analyses are briefly summarized (Sec. 8), and some possibilities for further research are listed (Sec. 9).

2 Background

To provide the motivation for the research of this thesis, the context of this research in modern particle physics is sketched in this section. The theoretical framework which it intends to test (the SM) is briefly introduced, as well as some previous research relevant to this thesis.

2.1 The Standard Model

The Standard Model of particle physics is the leading theoretical framework in modern particle physics. Formulated in quantum field-theoretic terms, the dynamics and kinematics of the model are controlled by the SM Lagrangian. This single equation effectively contains the entirety of particle physics, and can be derived from symmetry principles.

It is generally understood that the laws of physics should be frame-independent. I.e. the laws of physics are not dependent on space, time or orientation. In mathematics such invariance is described by symmetry groups. The Poincaré symmetry group contains all translations, rotations, and boosts in Minkowski space-time, and describes all the aforementioned invariant actions in four-dimensional space-time. The mathematical objects that obey these symmetries are representations of this group. These mathematical objects turn out to be fields. To the symmetries and their representations particles can be associated, which appear as excitations of their corresponding field. By further imposing local symmetries called gauge symmetries, vector fields arise naturally out of this theory. Many properties of these particles are determined by their associated symmetry groups. This is the way in which the Standard Model is constructed.

The Standard Model Lagrangian can be constructed by starting from the Yang-Mills Lagrangian. Which in general form is given by:

$$\mathcal{L}_{YM} = \bar{\psi}(i\not{D})\psi - \frac{1}{4}(F^i_{\mu\nu})^2 - m\psi\bar{\psi} \tag{1}$$

Here ψ and $\bar{\psi}$ are some Dirac field and its Dirac adjoint, $\not{D}(= \gamma^\mu(\partial_\mu - igA^a_\mu t^a_r))$ is the covariant derivative for the relevant gauge symmetry, $F^i_{\mu\nu}(= \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + gf^{ijk}A^j_\mu A^k_\nu)$ is the field tensor in representation i with A^a_ν being the vector potential, g the coupling strength, and f^{ijk} the structure constants of the associated Lie Algebra. By considering massless Dirac fields for the SM gauge symmetry $SU(3) \times SU(2) \times U(1)$ [2], the following Lagrangian can be found:

$$\begin{aligned}\mathcal{L}_{\text{gauge}} = i\bar{\psi}_L\gamma^\mu(\partial_\mu + ig_s\lambda_a G^a_\mu + ig\sigma_i W^i_\mu + iY_L g' B_\mu)\psi_L + \\ i\bar{\psi}_R\gamma^\mu(\partial_\mu + ig_s\lambda_a G^a_\mu + iY_R g' B_\mu)\psi_R - \frac{1}{4}F^{\mu\nu}_a F_{a\,\mu\nu}.\end{aligned} \tag{2}$$

The Dirac field has been split into its left-handed (doublet) and right-handed (singlet) components $\psi = \psi_L + \psi_R$. These *Fermion* fields now couple to the gluon fields G^a_μ associated to the $SU(3)$ symmetry and to the B_μ field associated with $U(1)$ symmetry. Due to parity violation associated with the $SU(2)$ gauge symmetry the fields W^i_μ couple only to the left-handed fields. After electroweak symmetry breaking, W^3 and B mix to give rise to the familiar Z^0 and photon neutral boson fields, where W^1 and W^2 mix to form the $W^\pm$ charged boson fields. g_s, g, and g' are the coupling constants of the fermions to these bosons. The weak hypercharges Y depend on the type of field that is considered.

[2] The SM gauge symmetry group is often denoted as $G_{\text{SM}} = SU(3)_c \times SU(2)_L \times U(1)_Y$ instead, where c indicates that the gauge field corresponding to the $SU(3)$ symmetry acts on particles with a color charge, the L indicates that the $SU(2)$ symmetry only acts on left-handed fields, and the Y indicates that the $U(1)$ is associated with the weak hypercharge.

	Particle	Y_L	Y_R
Leptons	e	$+1$	-2
	μ	$+1$	-2
	τ	$+1$	-2
	ν_e	-1	
	ν_μ	-1	
	ν_τ	-1	
Quarks	u	$+\frac{1}{3}$	$+\frac{4}{3}$
	d	$+\frac{1}{3}$	$-\frac{2}{3}$
	c	$+\frac{1}{3}$	$+\frac{4}{3}$
	s	$+\frac{1}{3}$	$-\frac{2}{3}$
	t	$+\frac{1}{3}$	$+\frac{4}{3}$
	b	$+\frac{1}{3}$	$-\frac{2}{3}$

Table 1: Weak hypercharge values for all lepton fields.

The left-handed lepton doublets become $\psi_L = (\nu_{l\,L}, l_L)^T$ with $l \in \{e, \mu, \tau\}$, and quark doublets become $\psi_L = (q_{u\,L}, q_{d\,L})^T$ with $q_u \in \{u, c, t\}$ and $q_d \in \{d, s, b\}$. Furthermore, there are right-handed singlets for all charged leptons $(e_R, \mu_R, \tau_R, u_R, c_R, t_R, d_R, s_R, b_R)$. In the full SM, after electroweak symmetry breaking these fields are given mass through the Yukawa interactions with the Higgs field (not discussed here). Table 1 shows the values of the hypercharge for each of these fields.

In addition to the weak hypercharge, weak isospins (I_3) are also associated to each particle field. The value that this constant takes for the left-handed doublets is $+1/2$ for all neutrinos ($\nu_{l\,L}$) and up-type quarks ($q_{u\,L}$), while the value equals $-1/2$ for all charged leptons (e_L, μ_L, τ_L) and down-type quarks ($q_{d\,L}$). For all right-handed singlets the weak isospin is 0. The electromagnetic charge, which should be familiar to all readers, can then be calculated according to the simple equation:

$$Q = I_3 + \frac{1}{2}Y \tag{3}$$

Based on these considerations, a full classification of elementary particles in the SM can now be made. Firstly, bosons with an integer spin can be distinguished from fermions with a half-integer spin. There are four vector bosons which mediate the fundamental interactions, and one scalar boson. The gluon mediates the strong interaction associated to $SU(3)$, the $W^\pm$ and Z^0 bosons mediate the weak interaction's charged and neutral currents associated to $SU(2)$, and the photon γ mediates the electromagnetic interaction associated to $U(1)$. The Higgs scalar-boson gives mass to these other particles.

Additional classifications can be made in the fermion fields based on their interaction with each of the fundamental forces. Quarks can be distinguished from leptons, in that they couple to the gluon fields, where leptons do not. Up-type quarks and neutrinos can be distinguished from down-type quarks and charged leptons based on their (hyper)charge. Neutrinos have a neutral charge and do therefore not couple to the electromagnetic field. These four different types of fermions are represented by the rows in Figure 1.

In addition to these interaction-based classifications, additional flavour-based classifications can also be made. For each type of fermion three flavours can be distinguished. All in order of increasing mass: for the up-type quarks there are up, charm, and top quarks. For the down-type quarks there down, strange, and bottom (or beauty) quarks. For the charged leptons there are electrons, muons, and taus. Finally, for the neutrinos there are the corresponding electron-, muon, and tau- neutrinos. Based on the Standard Model, the different flavours of leptons differ only in their mass. The SM has the property of lepton flavour universality (LFU). That means that besides the effect of the different masses of these leptons in the Yukawa interaction and flavour-mixing effects, all three lepton flavours behave the same.

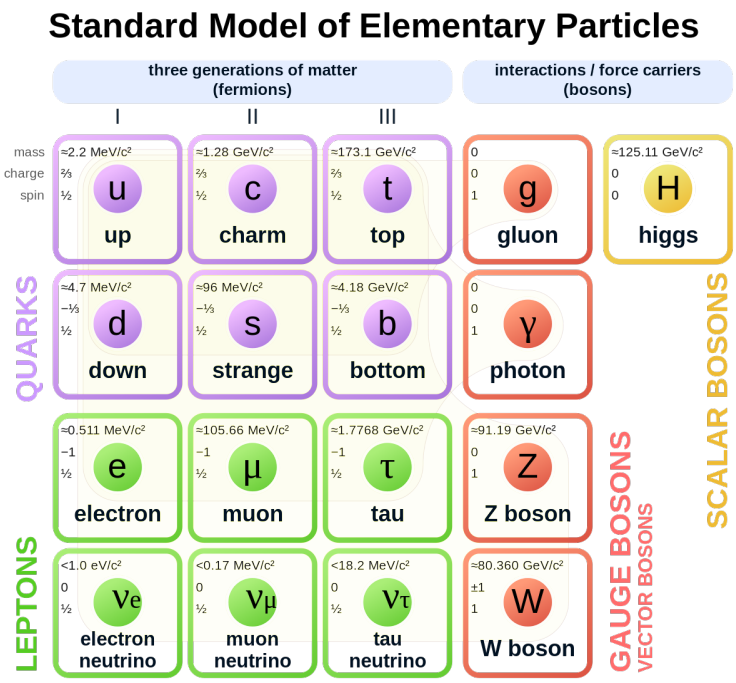


Figure 1: Table showing the SM elementary particles grouped by particle type.

So, while the coupling strengths for the interactions with the different bosons may differ between each type of fermion (due to the differences in their hypercharges for example), they do not further depend on the lepton flavours. That is to say the coupling constants g_s, g', and g are universal.

2.2 Lepton Flavour Universality

In recent years potential violations in LFU have been observed in the semileptonic decays of B^0-mesons (bound state of a $\bar{b}$ and a d quark) into $D^{(*)+3}$ mesons (c and $\bar{d}$). The branching fraction of such decays involving a τ^+ have been measured relative to those involving an e or a μ. The ratio of these branching fractions is written as:

$$\mathcal{R}(D^{(*)}) = \frac{\mathcal{B}(B^0 \to D^{(*)-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{(*)-}l^+\nu_l)} \tag{4}$$

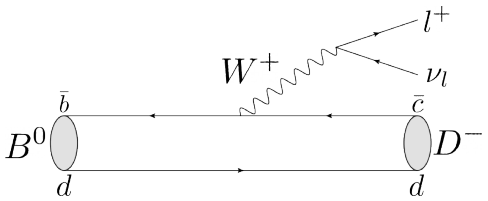
,

where $l \in \{e, \mu\}$. Colloquially, these measurements are therefore known as $R(D^{(*)})$. The experimental average of this ratio deviates from the theoretical average of SM predictions by three standard deviations (3σ) [2]. While this is not considered to be a discovery (this requires a 5σ deviation), it is considered to be a hint of New Physics. As a result potential further violations in LFU have gathered a lot of attention from physicists, in the experimental realm, as well as in the theoretical realm.

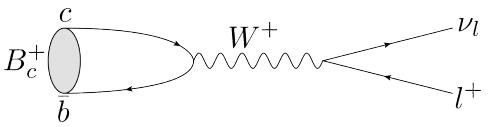
In Figure 2a the Feynman diagram of the decay $B^0 \to D^- l^+ \nu_l$, whose branching fraction was measured in the $R(D)$ measurement, is shown. On the quark level, i.e. ignoring the spectator d quark, this is the semileptonic decay of a b quark into a c quark by radiating off a charged lepton and its corresponding neutrino through the mediation of a W^+ boson. The Feynman diagram of this transition is equal to

[3]The asterisk here implies that the meson is in an excited state. So in fact the branching fractions of two different decays are considered here: $B^0 \to D^- l^+ \nu_l$ and $B^0 \to D^{*-} l^+ \nu_l$

the fully leptonic decay of a b and c quark into those same leptons. Hadronised, these quarks form a B_c^+ meson. So any any NP effects present in this part of the $B^0 \to D^- l^+ \nu_l$ could also be found in the fully leptonic decay of this meson $B_c^+ \to l^+ \nu_l$ (2b). Based on the calculation of NP Wilson coefficients, if these effects enter through pseudoscalar interactions, the branching fraction for the tauonic version of this decay could be increased by $\sim$ 17-30 times relative to SM predictions.



(a) $B^0 \to D^- l^+ \nu_l$



(b) $B_c^+ \to l^+ \nu_l$

Figure 2: Two Feynman diagrams showing how the quark-level process ignoring the spectator d quark is the same in both decays. Here l is a lepton: $l \in (e, \mu, \tau)$.

No strong evidence has been found for LFU violations between the e and μ flavours. As a result theoretical models have arisen that attempt to split the third (τ) generation from the other two. Some models introduce *leptoquark* interactions in addition to the SM interactions. Bachelor student Jakub Kwasniak [3] has investigated such theoretical models. These supposed leptoquarks carry properties of both leptons and quarks. Often times these theories are constructed by supposing a greater gauge symmetry which the SM gauge symmetry is a sub-group of. In some cases leptoquarks arise naturally out of such models. The measurement of the branching fraction of $\mathcal{B}(B_c^+ \to \tau^+ \nu_\tau)$ could place constraints on such NP models, and be a test for LFU.

3 The LHC*b* experiment

To understand the results provided in this thesis, it is necessary to have a broad understanding of the LHC*b* detector, and how data is collected and processed in this experiment. This section will provide an overview of the LHC*b* and its subsystems, the data flow of this detector, and how simulated data is used and generated for the experiment.

3.1 Detector Description

The LHC contains four major experiments: ATLAS, ALICE, CMS, and LHC*b*. The design of each of these detectors is suited for a specialised set of research goals. Three of these experiments are designed to maximise their solid-angle coverage around the collision event (colloquially referred to as 4π detectors for this reason). The largest of these is ATLAS, a general-purpose detector able to investigate a wide range of topics. ATLAS is primarily responsible for the first measurement of the Higgs boson. CMS is also a general-purpose detector with broadly the same research goals as ATLAS, but with a different technical implementation. These two general-purpose detectors are therefore able to perform independent measurements of the same physical phenomena. The third 4π detector, ALICE, is specialised in heavy ion collisions to better understand the strong interaction.

The LHC*b* detector in contrast, is the only forward-facing spectrometer on the LHC. Its detectors are concentrated in a small solid-angle, close to the beam line. As a result, it is specialised in detecting particles whose momentum is primarily along the beam line. This makes it uniquely well suited for detection of decays of heavy flavour particles. Of main interest are hadrons containing a b quark, hence the name of the experiment. Relevant for this thesis are two types of heavy mesons containing a b quark: the B^+ meson and the B_c^+ meson. These particles are bound states of the b quark in combination with either a u quark or a c quark respectively.

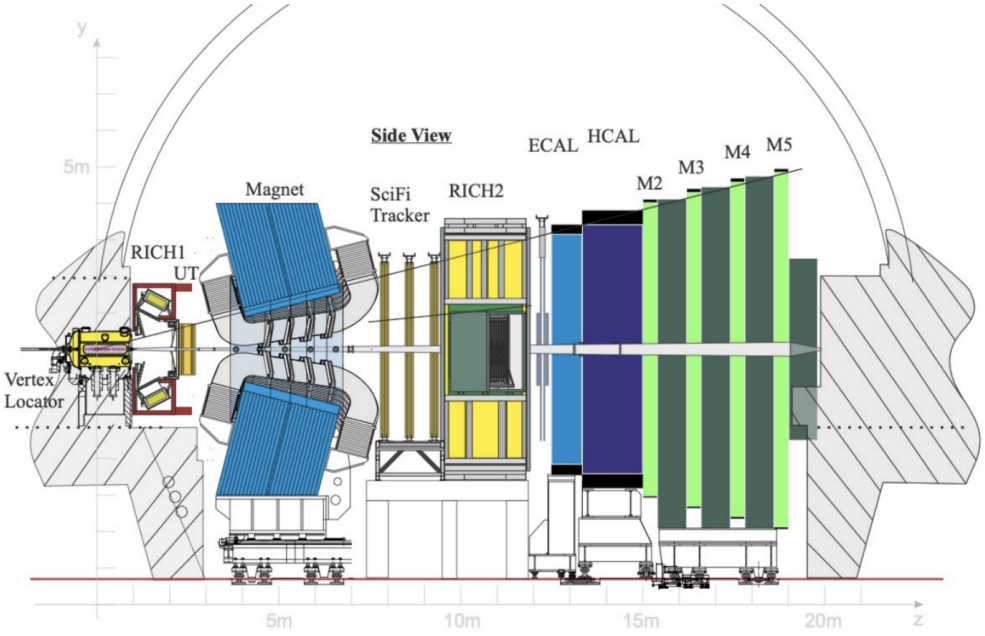


Figure 3: Schematic showing the different components of the LHC*b*.

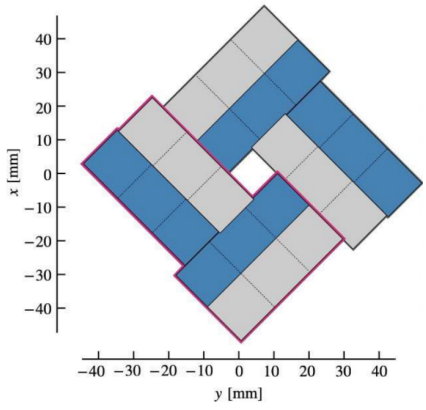
Due to the geometry of the detector, only particles with specific kinematic properties will be detectable in the LHC*b*. Specifically, the particles momentum needs to have a large longitudinal component. The quantity used to describe this is called the pseudorapidity, and can be expressed as:

$$\eta = -\ln(\tan(\frac{\theta}{2})) = \operatorname{arctanh}\frac{P_Z}{|\vec{p}|} \tag{5}$$

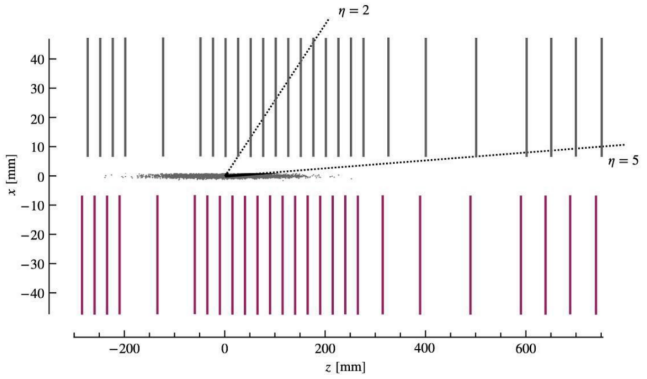
Here θ is the angle between the momentum of a particle and the beam-line, and $\vec{p}$ and P_z respectively are the particle's three-momentum and the component of that momentum longitudinal to the beam line (z axis). The range in which particles are detectable is called the *Acceptance* of the detector. The LHC*b* detector acceptance covers the pseudorapidity range $2 < \eta < 5$ [4]. In this thesis this will be referred to as the *LHCb acceptance*.

Figure 3 shows a schematic of the LHC*b* detector. The detector consists of several sub-systems, each performing an important task for the gathering of data for the experiment. Starting from the beam crossing (left in the schematic), we first encounter the Vertex Locator (VELO), which is designed to determine the location of the primary decay vertices that correspond to pp collisions. A few meters behind the VELO, a large magnet is placed that is necessary to determine the momenta of charged particles. This magnet can be polarised either up (magup) or down (magdown). The polarity of this magnet is switched occasionally[4] to avoid bias in the data taking. A Ring-Imaging Cherenkov detector (RICH1) and the Upstream Tracker (UT) are placed upstream of this magnet (closer to the collision point). Downstream of the magnet the RICH2 and a Scintillating-Fibre tracker (SciFi) are positioned. This is then followed by an electromagnetic calorimeter (ECAL), specialised in detecting light particles such as electrons, photons and neutral pions, as well as a hadronic calorimeter (HCAL) specialised in the detection of heavier (hadronic) particles. Lastly, at the end of the detector several Muon Stations (M2-M5) are placed for the identification of muons.

A sub-system of the LHC*b* that is of particular importance for this thesis is the VELO. This detector is placed around the collision point of the beam. The VELO consists of two sets of modules that can be moved to close around the beam line. These are called the A-side and C-side respectively. The modules of the two sides of the VELO are off-set by a distance of 12.5 mm along to the beam line. Each module consists of 12 ASICs (application specific integrated circuits), arranged in four 3×1 arrays. Each ASIC consists of 256×256 pixels (each of size $55\,\mu\text{m} \times 55\,\mu\text{m}$) [5].



(a) Front view of the VELO modules. The gray ASICs are facing towards the viewer, while the blue ASICs are facing away from the viewer. The C-side of the module is highlighted in purple.



(b) Side view of the VELO modules. Here the C-side modules are shown in purple, and the A-side modules are shown in grey. For visualisation, a scatter plot of primary vertices is shown, along with two lines showing the accpetance range of the LHC*b*.

Figure 4: Schematics of the front and side view of the VELO. Taken from [6].

[4]Typically the polarity of the magnets is switched every few months.

3.2 LHC*b* data flow

The LHC*b* detector in Run 3 was upgraded to be able to handle up to 40 million bunch-crossings per second. This results in a raw detector output data rate of ~ 5 TB/s. This is far too much to all be saved. Therefore, as data is collected, decisions already need to be made on what data is to be stored for analysis. These decisions are made by automated systems called triggers. In Run 1 and Run 2, the first trigger was a hardware trigger (L0) designed to make low-level selections on what data to process further. This was followed by a software implemented high-level trigger (HLT), which was further sub-divided into two steps: HLT1 and HLT2. Starting Run 3, the hardware trigger is no longer used, and all selection is handled by the software triggers. HLT1 partially reconstructs the decay of the particles, and reduces the data by a factor of $\sim 20 - 70$. The data passing HLT1 is temporarily saved on tape for up to a few weeks. HLT2 performs a full reconstruction of the decays that it is assigned to reconstruct, and reduces the amount of data further by another factor of $\sim 7 - 20$. That data is then saved for off-site analysis. HLT2 is instructed to reconstruct specific decays that analysts are interested in by implementing a *trigger line* for that decay. Such a trigger line acts as a sort of recipe for building a decay. If all the ingredients are present, the decay passes the trigger line and can be analysed. After all this is done, only ~ 10 GB/s remains, which is low enough to be stored long-term [7].

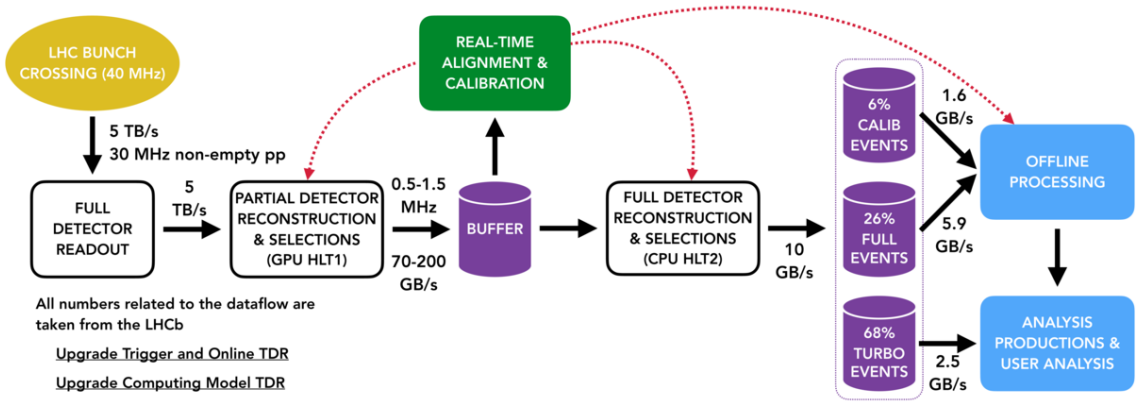


Figure 5: Schematic showing the data flow of the LHC*b* from detector readout to analysis[7].

The main tool for designing and testing a trigger-line is a software package called Moore. To test the trigger-line, it is applied to simulated data of that specific decay using Moore. If it is found to be sufficiently efficient, then it can be applied to real data. A trigger-line for the decay $B^+ \to \tau^+\nu_\tau$ has already been written, and analysing the resulting data is one part of this thesis. The main tool for this analysis is the software package DaVinci. DaVinci can be used to produce `.ROOT` files containing specific properties related to the events and reconstructed objects in those events.

Both Moore and DaVinci use objects called functors. As of Run 3, the ThOr (throughput oriented) framework is used. A functor is essentially an instruction to perform some calculation. The Gaudi framework, on which both Moore and DaVinci are based, then compiles these instructions into working C++ code as the application is run. For example, if one wants to know the mass of some particle the user only needs to specify to which particle he would like the functor `F.mass()` to be applied to. Following this request, C++ code is then compiled in order to collect this information. Using DaVinci, the masses of the specified particles can then be saved in a `.ROOT` file.

3.3 Simulation

In this thesis, simulated data is analysed for the purposes of investigating candidate decays for the normalisation channel of $B_c^+ \to \tau^+\nu_\tau$. Two different sources of simulated data are considered. LHC*b* Monte Carlo Simulations, and RapidSim. Both of these are Monte Carlo (MC) simulations. This refers to the method of generating the data samples. To generate an event, first the starting parameters are randomly chosen according to some probability distribution. Then, at each time step a set of possible processes is calculated. Each of these solutions is given a weight based on their likelihood. One option is then chosen, and this continues for a number of time steps. Different MC simulations can differ greatly in their accuracy, based on the physical processes considered, the amount of time-steps calculated, and the assignment of weights to the solutions.

LHCb Monte Carlo

The more detailed, and computationally heavy of the two simulation methods considered in this thesis is the LHC*b* MC simulation. Figure 6 shows the data flow in the LHC*b* experiment, including the simulation of data, encircled in blue. In Run 3, after the trigger step, sprucing and analysis with DaVinci follow.

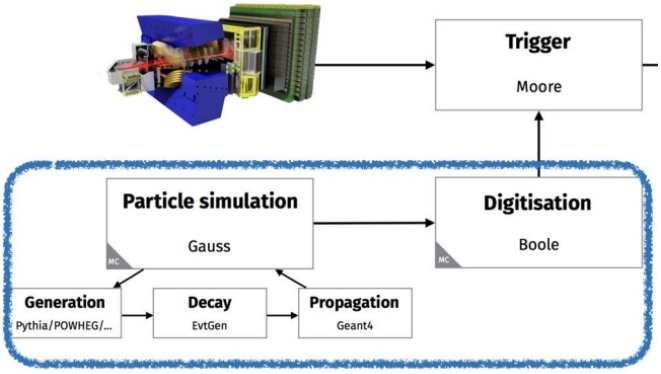


Figure 6: Schematic showing the data flow of the LHC*b* experiment, along with the relevant software packages at each step. The steps highlighted in blue are done in simulation. Taken from [8].

The simulation of events for the LHC*b* consists of two main steps: Particle Simulation and Digitisation [8]. Simulation is handled by the Gauss framework, which consists of three stages. First, the event is generated, the software most often used for this is Pythia. In this stage, all particles that are created at the collision and their physics (hard processes, hadronisation, the underlying process, etc.) are generated. The second stage is the simulation of the decays of these particles using the EvtGen software. This is done based on known decays and their experimental or theoretical branching fractions. At this stage a user-specified decay can be specified to occur in the event. In the case of the research described in this thesis, either a $B^+ \to \tau^+\nu_\tau$ or a $B^+ \to D^-\pi^+\pi^+$ decay is required to be present in the event. Only if one of the required decays is produced, does the simulation process continue to the next step. The last step in the simulation of the event is propagation, which is handled by Geant4. This is the most computationally expensive step in the simulation. At every step all possible processes (including interaction with detector material) are determined, and selected based on likelihood.

Before the last step in the simulation process, certain *generator level cuts* are made. In the case of the simulations mentioned in this thesis, this entails that all *long-lived* particles (this includes pions, muons and kaons) are within the LHC*b* detector acceptance. However, the acceptance here is less strict than the LHC*b* acceptance that was stated in Section 3.1. The generator level-cut for the simulations used in this thesis require that the pseudorapidity of the *daughter* particles (final state decay products) falls within the range $1.595979 < \eta < 5.298309$. This will be referred to henceforth as the *MC Acceptance*. Furthermore, because the LHC*b* detector is one-sided, we are not interested in processes in the -z direction. For these events, the z coordinate is simply flipped, and the propagation is still simulated. In the end, Gauss provides the hits and energy deposits in the LHC*b*.

After the simulation is complete, the next step is digitisation. This is handled by the software package Boole. Using this software the reaction of the different detectors to the results from Gauss are simulated. At the end of this step, the simulation data should be indistinguishable from real LHC*b* data. Typically the simulated tracks are (partly) saved, in order to estimate the efficiency of the reconstruction. These saved tracks we will call the *true* MC data.

For the simulation of events and their digitisation, it is important to have accurate information of the detector. At LHC*b*, such information is stored in the detector description database (dddb). The detector information used in the simulation is given a tag which can be called to access it. Furthermore, the simulation conditions are saved in a database (conddb), and are also given a tag. These tags need to be specified when running over MC data in Moore or DaVinci. The relevant tags for the simulations investigated this thesis are given in Table 2.

	$B^+ \to \tau^+\nu$	$B^+ \to D^-\pi^+\pi^+$
`conddb_tag`	`sim-20210617-vc-md100` (magdown) `sim-20210617-vc-mu100` (magup)	`sim-20231017-vc-md100` (magdown) `sim-20231017-vc-mu100` (magup)
`dddb_tag`	`dddb-20210617`	`dddb-20231017`

Table 2: Table showing the conddb and dddb tags for the MC data samples used in this thesis.

RapidSim

In addition to the LHC*b* Monte Carlo Simulation, the simulation software RapidSim was used. RapidSim simulations are far less detailed than the full LHC*b* MC simulation, but they are also much faster, which allows the software to be run locally. Where just the propagation step of the LHC*b* MC simulation running on a large CERN computer cluster is likely to take more than a minute per event [8], RapidSim running locally is able to generate a million events in the same time. A large contributor to the efficiency of RapidSim is that it entirely skips simulating the detector response. Instead it applies a simple Gaussian smear to the momenta of the particles.

In RapidSim events can be generated in a wide acceptance. For the RapidSim data samples used in this thesis, events were generated in the ranges $0 < \eta < 8$ and $0 < P_T < 100$ GeV. These data samples were cut on the pseudorapidity of the daughter particles to align with the MC acceptance.

For this thesis, hits in the VELO modules were modelled using a script developed by M.D. Galati [9]. In this model the hit efficiency for the relevant particles was assumed to be 100%.

4 Reconstructing $B_c^+ \to \tau^+\nu_\tau$

To choose a normalisation channel for $B_c^+ \to \tau^+\nu_\tau$ it is important to understand this decay. In this section the topology and the reconstruction method of the $B_c^+ \to \tau^+\nu_\tau$ decay will be explained. This decay has not been measured before, due to its unique characteristics that make it difficult to measure. However, the decay could be measured by using a new reconstruction technique developed specifically for this decay called *B-tracking*. Nick Sundstrom recently wrote his Master's thesis *Triggering Invisible Decays* [10] on triggering $B_c^+ \to \tau^+\nu_\tau$ with B-tracking.

4.1 Kinematic Properties

The τ^+ coming out of the decay is short-lived. Consequently, it leaves no hits downstream in the LHCb detector. It is not directly identified, but reconstructed based on some secondary decay. The sub-decay $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$ was chosen for this purpose. The total signal decay then becomes $B_c^+ \to (\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau)\nu_\tau$.

A major difficulty in reconstructing this decay lies in the fact that two neutrinos are produced in the process. Due to their weakly-interacting nature, these cannot be detected in the LHCb detector. This means that the only visible decay particles are the three pions coming out of the sub-decay $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$. Essential information is therefore missing. The momenta carried by the neutrinos is unknown. This makes it impossible to fully reconstruct the decay.

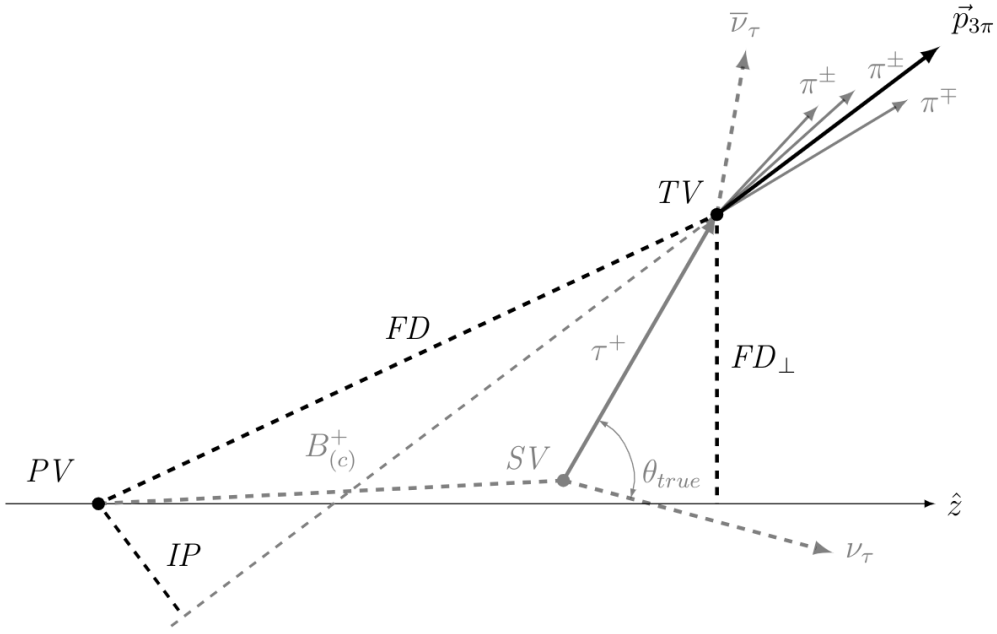


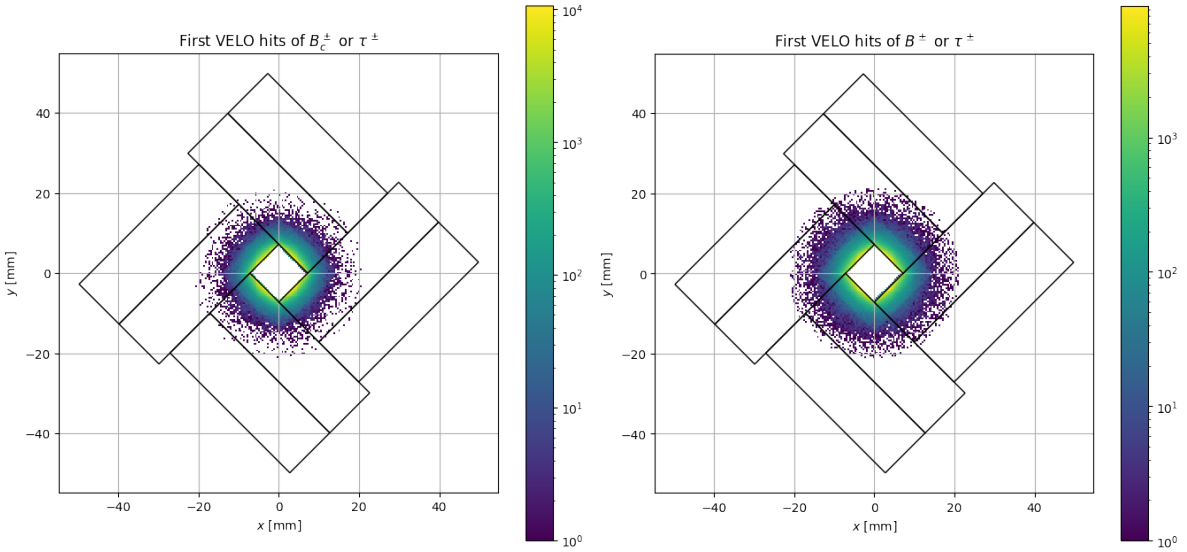
Figure 7: Kinematics of the $B_c^+ \to \tau^+\nu_\tau$ decay. Taken from Jelte de Jong's Master's thesis [11].

Figure 7 shows the general shape of the decay, along with the sub-decay. The B_c^+ meson is produced at the primary vertex (PV), and decays at the secondary vertex (SV). At the SV a τ^+ and corresponding neutrino are produced. This τ^+ then decays at the tertiary vertex or tau-vertex (TV). The flight distance (FD) then is the distance covered by the B_c^+ and the τ^+ combined. It is the length of the line segment going from the PV to the TV. The transverse flight distance ($FD_\perp$) is then the length of the component of the FD transverse to the beam line (z-axis). The distance of closest approach between the trajectory of visible final-state particles (pions) and the PV is then defined as the impact parameter (IP).

Another important physical parameter of the decay is the angle between the two particles produced at the SV (θ_{true}), because this holds a piece of information about the momentum of the invisible neutrino. If the TV can be reconstructed, then $FD_{(\perp)}$, IP, and $\vec{p}_{3\pi}$ can all be determined. However, due to the missing momentum carried by the invisible neutrinos, and due to a large abundance of pions in the LHCb it becomes practically impossible to identify any three-pion vertex as coming from the signal decay. There are several selections that can be made on the reconstructed vertex that increase the probability of it belonging to the signal decay, such as a cut based on the combined 3π mass. This would however not be enough to reconstruct the full decay. This appears to be a dead end, and generally this is where analysis of this decay would end.

4.2 B-tracking

To reconstruct this decay then, becomes an exercise in what can be done with the information that *is* available. There is one piece of information that we have not yet talked about which can help reconstruct this decay. While the B_c^+ and τ^+ are both very short-lived, and most of these particles produced in the LHCb will not make it into a VELO module, it is certain that some will. These particles can then leave direct hits in the VELO. It is these VELO-hits[5] that the B-tracking algorithm uses in its reconstruction of the decay. Figure 8 shows the locations of VELO-hits caused by the signal particles as simulated by RapidSim [9].



(a) $B_c^+ \to \tau^+\nu_\tau$ events

(b) $B^+ \to \tau^+\nu_\tau$ events

Figure 8: Front view of hits in the VELO from data simulated in RapidSim. Taken from [9].

To select signal taus a good estimate of the reconstructed τ^+ momentum would be useful. However, the missing momentum of neutrinos prevent this. The B-tracking algorithm is designed to be able to estimate the mass of the τ^+, which can then be used to distinguish signal from background. Starting by taking the combined pion mass as an estimate for the τ^+ mass given by:

$$m_{3\pi} = \sqrt{E_{3\pi}^2 - \vec{p}_{3\pi}^2}, \tag{6}$$

[5]These are also referred to as *B-hits*. This term encompasses hits in the VELO from any relevant particle, and not just the B^+. The author therefore prefers to use the term VELO-hit.

This estimate can be improved if a VELO-hit can be associated to the decay. The first VELO hit (FH) associated to the decay is assumed to be caused by the B_c^+. From this, the direction of the B_c^+ momentum can be determined. By considering the transverse component of the three-pion momentum ($\vec{p}_{3\pi\,\perp}$) to this PV-FH line some of the missing information can be retrieved such that a corrected mass can be defined as:

$$m_{\text{corr}} = \sqrt{m_{3\pi^2} + |\vec{p}_{3\pi\,\perp}|^2} + |\vec{p}_{3\pi\,\perp}| \tag{7}$$

The way that the B-tracking algorithm intends to find VELO-hits corresponding to this decay, is by searching in a cylinder with a radius of 0.5 mm around the PV-TV line. If a VELO hit is found, then the angle θ_{corr} between the combined three-pion direction of flight, and the PV-FH line can be calculated. This is an approximation of the true angle θ_{true} that is shown in figure 2. Using this value then, $\vec{p}_{3\pi\,\perp}$ can be determined, and the corrected mass can be determined.

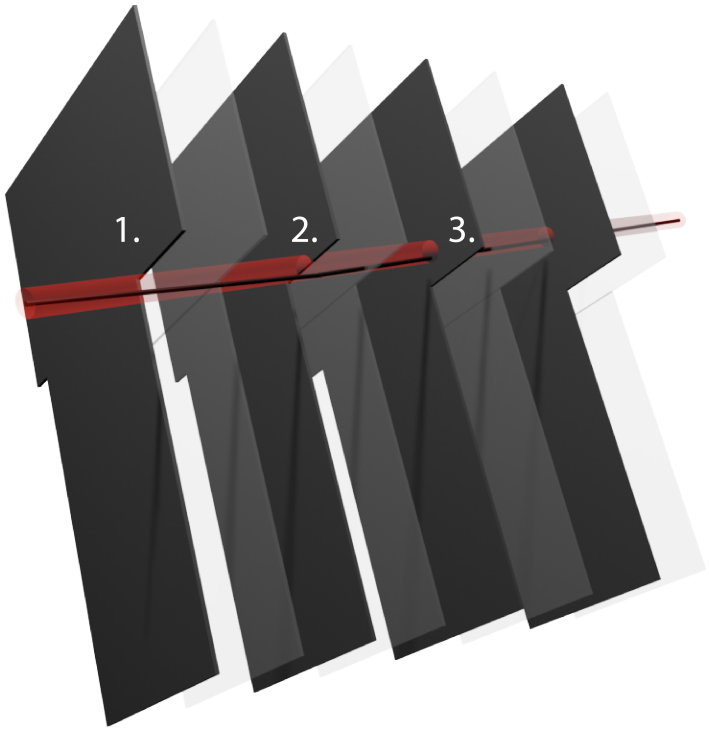


Figure 9: 3D rendering of a section of the C-side of the VELO. Half of the modules are made transparent for better visibility. The search window is represented by a red cylinder. The black line represents the track of the B^+ and τ^+ particles. The size of the search-window is exaggerated for clarity.

Figure 9 shows three distinct cases to consider. In events that only contain Case 1., the search-window does not cross a a VELO sensor, and neither do the particle tracks. In events containing Case 2., the search window does cross the sensor, but the particle track does not. Consequently, there is no VELO-hit. Just one of these sensor-crossings is enough for the event to pass the HLT2 line. These events are saved for offline analysis. If the event contains Case 3., also the particle track crosses a sensor. In this case there is a VELO-hit. Only events with such a VELO-hit can be reconstructed with the method explained above.

5 Choosing a Normalisation Mode

The research goal of this thesis is to select a normalisation mode for the measurement of the decay $B_c^+ \to \tau^+\nu_\tau$. This section will clarify what this entails, and what the criteria are for selecting a decay for this purpose. A set of various normalisation mode candidates is compiled, and simulation data is used to formulate conclusions on the feasibility of these decays being used for normalisation. Based on the results of this analysis, two decays are chosen to investigate further.

5.1 Considerations

The overarching goal to which this thesis intends to contribute, is to report a first measurement of the branching fraction of the decay $B_c^+ \to \tau^+\nu_\tau$. This quantity is a measure of the chance of B_c^+ mesons decaying through this specific decay mode. However, what can be measured is not this branching fraction, but the number of reconstructed signal events. In principle, the branching fraction can be calculated from this by using the following formula:

$$\mathcal{B}(B_c^+ \to \tau^+\nu_\tau) = [\mathcal{L}\sigma_{B_c^+}\epsilon_{B_c^+\to\tau^+\nu}]^{-1} \times \mathcal{N}(B_c^+ \to \tau^+\nu_\tau) \tag{8}$$

Here $\mathcal{N}$ is the number of signal events reconstructed, $\mathcal{L}$ is the integrated luminosity, σ_{B_c} is the production cross-section of the B_c^+ meson within the LHCb acceptance, $\epsilon_{B_c^+\to\tau^+\nu}$ is the total detection efficiency of the decay mode, and $\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ is the branching fraction of the decay.

To get an accurate estimate of the branching fraction the total number of B_c^+ mesons produced ($\mathcal{N}_{\text{total}} = \sigma_{B_c^+} \times \mathcal{L}$) should be known accurately. However, this number is poorly known as the production cross-section, as well as the exact integrated luminosity are not well known. The solution is to not measure just a branching fraction for $B_c^+ \to \tau^+\nu_\tau$, but to measure it relative to some normalisation mode. In the resulting ratio of branching fractions these quantities drop out, so that they do not need to be determined. Moreover, a ratio of efficiencies can often be determined more precisely than just the efficiency of the single decay. This results in less systematic error. This is especially true for very small efficiencies. As a result, while the normalisation mode also contributes additional uncertainty in other ways (e.g. via measurement yield), a measurement relative to a good normalisation channel is likely to be very beneficial for the statistical significance of the measurement of $B_c^+ \to \tau^+\nu_\tau$. This is expected to have a very small efficiency due to the unlikelihood of a VELO-hit occurring. Such a ratio can be written as the following equation:

$$\mathcal{R} = \frac{\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)}{\mathcal{B}(\text{norm.})} = \frac{\sigma_N}{\sigma_{B_c^+}} \frac{\epsilon_N}{\epsilon_{B_c^+\to\tau\nu_\tau}} \frac{\mathcal{N}(B_c^+ \to \tau^+\nu_\tau)}{\mathcal{N}(\text{norm.})} \tag{9}$$

In this equation $\mathcal{N}(\text{norm.})$, σ_N and ϵ_N respectively are the signal yield of the normalisation channel, the production cross-section of the normalisation channel parent particle, and the total reconstruction efficiency of the normalisation channel. To minimise the uncertainty in the value of the ratio $\mathcal{R}$, the uncertainty of the three contributing factors (cross-section, efficiency, and measurement yield ratios) need to be separately minimised. If the branching fraction of the normalisation channel is well known, this then also minimises the uncertainty in the branching fraction of $B_c^+ \to \tau^+\nu_\tau$.

Cross-sections

For the cross-section ratio in Equation 9 there are two realistic options, a normalisation mode with a B_c^+ as the parent particle, or one with a B^+. In the first case the ratio disappears entirely, thus eliminating any uncertainty introduced from the number of B_c^+ mesons produced. In the second case, the production fraction of B_c^+ mesons (f_c) relative to B^+ (f_u) and B^0 (f_d) mesons has to be used. The following value was as been measured for 13 TeV [12]:

$$\frac{f_c}{f_u + f_d} = (3.78 \pm 0.04 \pm 0.15 \pm 0.89) \times 10^{-3} \tag{10}$$

In this value, the first error is statistical, the second is systematic, and the third error in this value is due to a large statistical error in the theoretical average of the branching fraction $B_c^+ \to J/\psi\mu^+\nu_\mu$ which was used for the calculation of this value.

These production fractions can be related to the production cross-sections in the following way:

$$\sigma_{B^+_{(c)}} = \sigma(pp \to B^+_{(c)} X_b) = f_{u/c} \times \frac{\sigma(pp \to b\bar{b})}{2} \tag{11}$$

Under the assumption that b quarks are equally likely to hadronise with u quarks as with d quarks ($f_u = f_d$), the cross-section ratio is simply the production fraction ratio in Eq. 10 multiplied by 2. The uncertainty in this cross-section ratio results in a large uncertainty in the final value $\mathcal{R}$ whenever a B^+ decay is chosen for the normalisation channel. Whether this penalty can be overcome depends on the remaining uncertainties associated to the introduction of normalisation mode.

Measurement Yield

One of these other uncertainties would be the uncertainty in the measurement yield. The error for this is taken to be the Poissonian standard-deviation $\delta\mathcal{N} = \sqrt{\mathcal{N}}$. As such, the relative error in the measurement yield can be simply reduced by measuring more decays. This can either be achieved by having a higher efficiency, or by having the decay occur more often. The LHC, and the LHCb detector are continuously being upgraded in order to create a higher integrated luminosity, and as the LHC keeps running more and more data will be collected, increasing the potential measurement yield of all decays. In Run 1 and 2 a total integrated luminosity of $9\,\text{fb}^{-1}$ was achieved. An additional $41\,\text{fb}^{-1}$ is expected in Run 3 and 4 combined, and later runs will contribute more still, up to a total integrated luminosity of $\sim 300\,\text{fb}^{-1}$. As for the question of choosing the right normalisation mode, decays with a large branching fraction will occur more often. As such, a decay mode with a high measurement efficiency and branching fraction will have a low uncertainty in the measurement yield.

Efficiencies

The efficiency ratio can be determined by simulating the decays in the LHCb experiment. The total efficiency for each decay can be split into several sub-efficiencies. When a decay is chosen that is similar to $B_c^+ \to \tau^+\nu_\tau$, then similar sub-efficiencies for both decays can be defined. In this case certain sub-efficiencies can (partly) cancel, decreasing the uncertainty. Moreover, the contribution to the total uncertainty by small efficiencies specifically can be reduced if similar efficiencies are found.

Consequently, it is favourable to have a normalisation mode that can be reconstructed in a way similar to $B_c^+ \to \tau^+\nu_\tau$. I.e. the B-tracking algorithm can be used for the normalisation mode as well. The most favourable option would be a $B^+_{(c)}$ meson decaying into a secondary semi-stable particle with a similar lifetime and mass to the τ^+. Table 3 lists some relevant particles with their corresponding average lifetimes and masses.

Calibration

Another consideration is calibration. Because the method used in this measurement (B-tracking) is new its efficiency is still poorly understood. It may be useful to consider a decay mode which is already well known through regular reconstruction methods. If B-tracking can also be used on this decay mode, then the efficiency of this method can be experimentally determined. This is not really a consideration for the normalisation mode specifically, but can still be sufficient reason in order to investigate some decay further as calibration is necessary anyway.

	Lifetime (s)	Mass (MeV/c^2)
τ^+	$(2.903 \pm 0.005) \times 10^{-13}$	1776.86 ± 0.12
D^+	$(1.040 \pm 0.007) \times 10^{-12}$	1869.62 ± 0.20
D_s^+	$(5.00 \pm 0.07) \times 10^{-13}$	1968.47 ± 0.33
B^+	$(1.638 \pm 0.004) \times 10^{-12}$	5279.34 ± 0.12
B_c^+	$(5.10 \pm 0.09) \times 10^{-13}$	6274.9 ± 0.8

Table 3: PDG values for the lifetimes and masses of several relevant particles [13].

List of Normalisation Mode Candidates

Considering the requirements above. The following list of decays contains potential normalisation modes that could respect these requirements:

1. $B^+ \to D^- \pi^+ \pi^+$
2. $B^+ \to \tau^+ \nu_\tau$
3. $B^+ \to J/\psi K^+$
4. $B_c^+ \to J/\psi D_s^+$
5. $B_c^+ \to J/\psi \pi^+$
6. $B_c^+ \to J/\psi \mu^+ \bar{\nu_\mu}$

In order to choose the best normalisation channel, it is useful to have an idea of their potential measurement yields, branching fractions, and detection efficiencies. These quantities will be estimated in the following sections.

5.2 Previous Results

By looking at data that was previously collected in Run 1 of the LHC*b*, expected yields for Run 3 and Run 4 could be estimated by extrapolating these numbers using the expected integrated luminosity and energy for Run 3 and 4. The integrated luminosity and collision energy for Run 1 through Run 4 are provided Table 4.

	Int. Lum. (fb^{-1})	Energy (TeV)
Run 1 (2011)	1	7
Run 1 (2012)	2	8
Run 2	6	13
Run 3 + 4	41	13.6

Table 4: Approximate integrated luminsoties and *pp* collision energy of the LHC*b* for different runs.

	Run 1 (2011)	Run 1 (2012)	Run 1 (total)	Run 3 + 4
$B^+ \to J/\psi K^+$	107783 [14]	243119 [14]	350902 [14]	8491671
$B^+ \to D^- \pi^+ \pi^+$	-	-	48606 [15]	1176245
$B_c^+ \to J/\psi \mu^+ \bar{\nu_\mu}$	3537 [16]	-	19140 [17]	463188
$B_c^+ \to J/\psi \pi^+$	839 [16]	-	3009 [18]	72818
$B_c^+ \to J/\psi D_s^+$	-	-	28.9 [18]	699

Table 5: Table showing what would be the yields of some normalisation mode candidates based on extrapolating measurement yields of Run 1 to Run 3 and 4 conditions.

The ratio between the total amount of times that a particular decay occurred in two distinct runs can be estimated using these values ($Ratio = \frac{\mathcal{L}_1 E_1}{\mathcal{L}_2 E_2}$). It can be calculated that Run 2 had a ~ 3.4 greater yield than Run 1 ($2011+2012$), and Run 3 and 4 have a combined expected yield that is ~ 24.2 times greater than the yield of Run 1, and ~ 7.1 times greater than the yield of Run 2, or ~ 5.5 times the combined yield of Run 1 and Run 2. Using these numbers, expected yields for Run 3 and Run 4 combined could be crudely estimated. Extrapolated measurement yields for Run 3 and 4 based on Run 1 yields are shown in Table 5.

To estimate the total efficiency of the reconstruction in these previous runs we need to have an estimate for the total number of decays occurring within acceptance, such that this efficiency is simply the ratio $\mathcal{N}_{\text{reconstructed}}/\mathcal{N}_{\text{total}}$. An estimate for $\mathcal{N}_{\text{total}}$ can be made by simply calculating:

$$\mathcal{N}_{\text{total}} = \mathcal{L} \times \sigma \times \mathcal{B}. \tag{12}$$

A measurement of the production cross-section $\sigma(pp \to B^{\pm}X)$ in the acceptance $2.0 < y < 4.5$[6] and $0 < P_T < 40$ GeV has been measured to be $86.6 \pm 0.5 \pm 5.4 \pm 3.4\,\mu$b at 13 TeV.

By correcting this cross-section for the right (MC) acceptance range[7] and by using the cross-section ratio from [12], the production cross-sections for both the B^+ meson and the B_c^+ meson in the relevant acceptance can be calculated to be:

- $\sigma_{B_c} = 1.59\,\mu$b
- $\sigma_{B^+} = 179\,\mu$b

Multiplying these values for the cross-sections by a combined integrated luminosity of 41fb^{-1} for Run 3 and 4, and the branching fraction (including sub-decays) of each normalisation mode candidate, results in the total yields and reconstruction efficiency estimates for each decay. The relevant branching fraction values can be found in Appendix A.1. The results of these estimates are stated in Table 6. The "Total" column shows the results from this calculation. It is an estimate of the total number of times that the decay mode occurred. The "Rec. Yield" column shows the results from the extrapolation above. No measurements of $B_c^+ \to \tau^+\bar{\nu_\tau}$ or $B^+ \to \tau^+\bar{\nu_\tau}$ have been done yet at LHCb, therefore the table is left empty for those decays. The "$\epsilon_{rec.}$" column shows the ratio of these numbers, and is an estimate of the total reconstruction efficiencies for the measurement of the decay modes in Run 1.

Decay Mode	Total	Rec. Yield	$\epsilon_{rec.}$
$B^+ \to D^-\pi^+\pi^+$	1.81×10^8	1.18×10^6	6.52×10^{-3}
$B^+ \to J/\psi K^+$	1.10×10^8	8.49×10^5	7.72×10^{-3}
$B_c^+ \to J/\psi\mu^+\bar{\nu_\mu}$	1.80×10^7	4.63×10^5	0.026
$B_c^+ \to J/\psi\pi^+$	1.23×10^6	7.28×10^4	0.059
$B_c^+ \to J/\psi D_s^+$	1.11×10^5	699	6.30×10^{-3}
$B_c^+ \to \tau^+\nu_\tau$	2.89×10^7	-	-
$B^+ \to \tau^+\nu_\tau$	1.82×10^7	-	-

Table 6: Reconstruction efficiency estimates of the normalisation mode candidates based on estimates of the total occurrence of the decay, and expected measurement yield.

This table gives an idea of what decays are feasible to choose as a normalisation mode. However, these numbers are calculated based on regular reconstruction methods used in Run 1. Using B-tracking will significantly reduce the efficiency,

[6]Here the rapidity y is used instead of the pseudorapidity η. These two quantities are however related such that η appears as the first-order contribution of the Maclaurin expansion of y expressed in terms of m/P_T.

[7]Thanks to Mick Mulder for calculating this.

5.3 Estimating VELO Efficiencies with RapidSim

Data samples generated using RapidSim were used to calculate VELO-hit efficiencies for each of the normalisation mode candidates. For each decay ten million events were generated according to the method specified in Section 3.3. The generated data sample was then VELO-hit filtered[8]. An event passes this filter if either the parent particle (B^+ or B_c^+), or a short-lived charged particle (τ^+, D^+ or D_s^+), or both leave at least one hit in the VELO module. The efficiency of this hit-filtering was calculated as follows:

$$\epsilon_{velo} = \frac{\text{Number of events passing hit-filtering script}}{\text{Number of events with all daughters in MC acceptance}} \tag{13}$$

By multiplying these efficiencies with the values calculated in the "Total" column of Table 6 estimates can be made of the amount of decays that are expected to leave a VELO-hit. These numbers are stated in the "VELO-hit Yield" column in Table 7 below. This does not take into account any additional cuts that are inevitably necessary for the reconstruction of these decays. Hence, the actual measurement yields for all these decays are expected to be lower.

Decay Mode	ϵ_{velo}	VELO-hit Yield
$B^+ \to D^-\pi^+\pi^+$	0.0073	1321300
$B^+ \to J/\psi K^+$	0.00194	209000
$B_c^+ \to J/\psi\mu^+\nu_\mu$	1.34×10^{-5}	241
$B_c^+ \to J/\psi\pi^+$	1.24×10^{-5}	15
$B_c^+ \to J/\psi D_s^+$	2.18×10^{-4}	24
$B_c^+ \to \tau^+\nu_\tau$	1.17×10^{-4}	3381
$B^+ \to \tau^+\nu_\tau$	0.00254	45500

Table 7: Table showing the VELO-hit efficiency for each normalisation mode candidate, as well as for $B_c^+ \to \tau^+\nu_\tau$ as calculated using the RapidSim-generated data samples. Expected yields were calculated using these efficiencies.

The effects of the differences in lifetime of the different particles (see Table 3) is reflected clearly in the VELO-efficiency for each decay. For example, the decays $B_c^+ \to J/\psi\mu^+\bar{\nu_\mu}$ and $B_c^+ \to J/\psi\pi^+$ have the lowest efficiencies. This is because for these decays only the short-lived B_c^+ can leave a hit. For the decay $B_c^+ \to J/\psi D_s^+$ the D_s^+ can also leave a hit, and therefore the efficiency is increased. The difference between the efficiency of this decay, and $B_c^+ \to \tau^+\nu_\tau$ can be explained by the difference in lifetime between the D_s^+ and the τ^+. All decays with the B^+ as their parent particle have much higher efficiencies, and additional short-lived charged particles increase their efficiencies even more.

[8] A modified version of M.D. Galati's hit-filtering script was used for this[9].

5.4 Conclusions on Normalisation Mode Candidates

The different normalisation mode candidates each have specific advantages and disadvantages. Each candidate will now be discussed briefly.

$B^+ \to \tau^+ \nu_\tau$

Being extremely similar to $B_c^+ \to \tau\nu$, $B^+ \to \tau^+\nu_\tau$ would be an obvious choice as a normalisation channel. Due to the lifetime, mass, and quark contents of the $B_{(c)}^+$ meson being the only differences between the decays, the reconstruction of these decays are also essentially the same, the only difference being the likelihood of a VELO-hit occurring. The total rate of occurrence of these two decays is also very similar. The higher abundance of the B^+ meson relative to the B_c^+ meson is approximately cancelled by the CKM matrix element V_{cb} (~ 0.04) being larger than V_{ub} (~ 0.004). As a result the difference in the expected yield is predominantly caused by the difference in VELO-hit efficiency between these two decays.

This decay has been previously measured at Belle [19] and BaBar [20]. However, data is limited, and the decay has not been measured by LHC*b*. In the absence of a reconstruction method being available for this decay without B-tracking, this decay can not be used for calibration purposes. Furthermore, this decay itself is sensitive to lepton flavour universality violations. Therefore, any NP effects in the measurement could partly cancel in the branching fraction ratio. As a result, a third decay is required for independent normalisation of both decays, as well as for calibration of the B-tracking algorithm.

$B^+ \to D^- \pi^+ \pi^+$

Of the normalisation channel candidates, this decay has the highest expected yield with B-tracking. The reason for this is that both the B^+ and D^- are relatively long-lived. Furthermore, this decay has already been measured previously at LHC*b* with a high yield [15], and its branching fraction has been independently measured at the B-factories [21][22]. This makes it a good decay for calibration purposes. This decay is one of the strongest candidates, and for this reason this decay will be further investigated in this thesis.

$B^+ \to J/\psi K^+$

Of the normalisation mode candidates considered, this decay has the highest yield without B-tracking. Furthermore, it also has a large yield when requiring a VELO-hit. This makes it a good candidate for calibration. However, in contrast to the previous decays mentioned, this decay does not contain a second short-lived charged particle. Therefore, measurement of this decay is a good method to isolate and investigate the effect of the VELO-hit of just the B^+ meson. The different decay topology of this decay however, also means that its reconstruction method is different, making it a sub-optimal normalisation mode. This thesis will not go into an in-depth analysis of this decay, but further research into this decay would be useful. Daniel Martinez, a PhD student of the Groningen LHC*b* group, is currently researching this decay.

$B_c^+ \to J/\psi D_s^+$

This is the only candidate considered which has both a B_c^+ meson as its parent particle and has a second charged short-lived particle present. These properties would make it a prime-candidate as a normalisation mode for $B_c^+ \to \tau\nu$. However, its measurement yield without B-tracking is estimated to be low, and the VELO-hit requirement decreases it even further. These low expected measurement yields at $41\,\text{fb}^{-1}$ integrated luminosity make it impossible for this decay to be used as a normalisation channel. As such, this decay can be dismissed.

$B_c^+ \to J/\psi \mu^+ \nu_\mu$

This is probably the best normalisation mode candidate with a B_c^+ as its parent particle. It has a decently large expected measurement yield when not using the B-tracking algorithm, and it also has the highest expected measurement yield of the B_c^+ decays when B-tracking is used. However, the B-tracking measurement yield is most likely too low to be a good normalisation channel.

Still, this is an interesting decay to investigate. It is a semileptonic decay with the same quark-level process as the $B_c^+ \to \tau^+ \nu_\tau$ decay, with an additional spectator c quark being present. Therefore, if the hadronic form-factors are well understood, this decay might provide one of the most direct lepton flavour universality tests in B_c^+ decays. As such, this normalisation could be used in future measurements of $B_c^+ \to \tau^+ \nu_\tau$. In fact, Amhis et al. [23] suggest this decay as a good normalisation channel for measurement at future Z-factories.

$B_c^+ \to J/\psi \pi^+$

Lastly, this decay was also considered as a possible normalisation channel. An advantage of this decay relative to the muonic decay above, is that it can be fully reconstructed, where the previous decay cannot be fully reconstructed due to the invisible neutrino. However, the branching fraction of this decay is lower. As such, the expected measurement yield of this decay is even lower still.

For this project, this seems to be an unsuitable decay. B-tracking cannot be used for this decay, as the expected yield is too low. However, this decay could also be used in future measurements, if higher yields can be achieved.

6 Analysis of Simulation Samples

Based on the results of the previous section, the decays $B^+ \to D^-\pi^+\pi^+$ and $B^+ \to \tau^+\nu_\tau$ were chosen to be potential normalisation mode candidates for the decay $B_c^+ \to \tau^+\nu_\tau$. In this section MC data samples for these decays will be analysed. First, in Section 6.1 the trigger lines for these decays will be discussed. Then, in Section 6.2, data of these decay simulated with RapidSim is compared to LHCb MC data. Several sub-efficiencies relevant to the reconstruction of these decays are determined in 6.4. Using these efficiencies and corresponding uncertainties obtained in these sections, an estimate of the total uncertainty on the final measurement value for both normalisation mode candidates can be made.

6.1 The Chosen Normalisation Mode Candidates

Before showing the simulation data samples, it is important to better understand the two normalisation mode candidates and their HLT2 lines. In section 4 the reconstruction method of $B_c^+ \to \tau^+\nu_\tau$ was already discussed. In order to minimise uncertainty in the measurement value, the chosen normalisation mode should have a similar reconstruction method. In this section, the reconstruction methods of the two remaining normalisation candidates will be discussed.

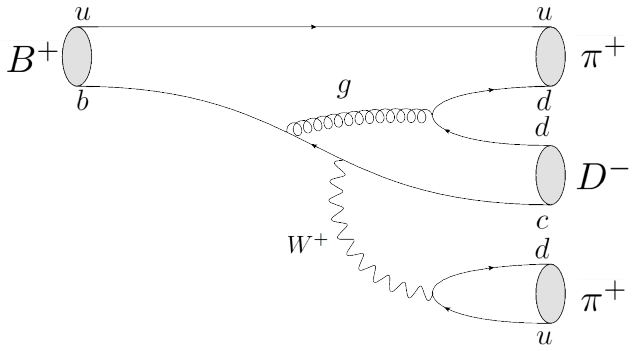
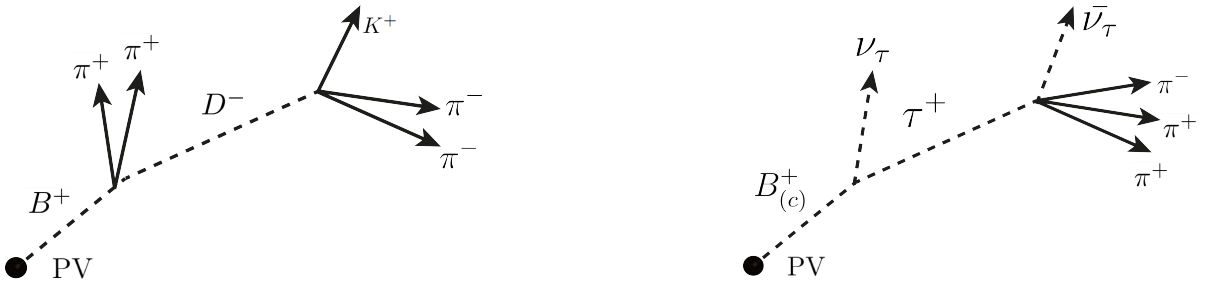


Figure 10: Highest contributing Feynman for $B^+ \to D^-\pi^+\pi^+$

Figure 10 shows the Feynman diagram of $B^+ \to D^-\pi^+\pi^+$ with the highest contribution to this decay process. The D^- meson produced in this decay is short-lived (see Table 3 on p. 20). For the signal sub-decay of this D^- meson the decay $D^- \to K^+\pi^-\pi^-$ was chosen. This sub-decay was chosen because it has a high branching fraction ($\sim 9\%$ [24]) and because it contains three decay-products, similar to the $\tau \to 3\pi\nu$ decay.



(a) $B \to D^-\pi^+\pi^+$ (b) $B_{(c)}^+ \to \tau^+\nu_\tau$

Figure 11: The two decays compared

The overall kinematics (Figure 11) of the $B^+ \to D^-\pi^+\pi^+$ and $B^+_{(c)} \to \tau^+\nu_\tau$ are then rather similar in terms of their kinematics. Both can be reconstructed using the B-tracking algorithm (Sec. 4.2). A major difference is that the $B^+_c \to \tau^+\nu_\tau$ decays contain the invisible neutrinos, while $B^+ \to D^-\pi^+\pi^+$ can be fully reconstructed. This makes $B^+ \to D^-\pi^+\pi^+$ a good normalisation mode because it allows for calibration of the B-tracking algorithm.

In order to calculate the efficiency of the reconstruction of the two potential normalisation channels, their HLT2 lines were run over simulated data samples for both decays. In the process of reconstructing these decays several cuts are made. In order to understand how the different parts of the reconstruction contribute to the total efficiency, the total efficiency can be subdivided into several sub-efficiencies. We define:

$$\epsilon_{total} = \epsilon_{Acc} \times \epsilon_{\tau/D-cuts} \times \epsilon_{HLT2} \times \epsilon_{sensor} \times \epsilon_{hit} \tag{14}$$

Here:

- ϵ_{Acc} is the efficiency of the decay products of the decay ending up in the LHCb acceptance. This efficiency will not be investigated in this thesis as no events outside of this acceptance are considered in the LHCb MC data.
- $\epsilon_{\tau/D-cuts}$ is the efficiency of several cuts made on the τ^+ or D^- to make sure that these particles can end up in the VELO in both respective decays.
- ϵ_{HLT2} is the efficiency of the total HLT2 line without B-tracking. This contains several cuts that differ between both decays.
- ϵ_{sensor} is the efficiency of a VELO sensor being in the B-tracking search-window. This allows for the second case of Figure 9 (p. 17).
- ϵ_{hit} is the efficiency of a VELO hit occurring within the search-window (only Case 3. from p. 17).

The HLT2 trigger line used for the reconstruction of the decay $B^+ \to \tau^+\nu_\tau$ is called `make_b2taunu_tau2pipipi_with_btracking`. In order to determine the last three of these efficiencies this line is run with three different settings: without B-tracking, requiring a VELO sensor in the search window, and requiring a VELO hit. Table 8 shows every cut made in this line. It shows which cuts belong to what sub-efficiencies defined above, and on what range each variable is cut, as well as a short description of that variable.

A similar table can be made for the `make_b2dpipi_d2kpipi_with_btracking` line. Table 9 shows the cuts made in the HLT2 line for $B^+ \to D^-\pi^+\pi^+$. The main difference here is the possibility to reconstruct the D^- and B^+ masses, so cuts can be made on these variables. Furthermore, an additional cut based on the direction angle is made.

The τ/D-cut constitutes a cut on the transverse momentum and on the transverse flight distance of the τ^+ or D^-. This combined cut favours shorter-lived particles. This will remove a portion of background from long-lived charged particles. In order to end up in the VELO, long-lived particles require less transverse momentum, and are hence more likely to be cut away. The cuts made on the τ^+ and the D^- are the same, however. Consequently, more long-lived low P_T D^- mesons will be cut than taus.

Efficiency	Variable	Range	Description
In-acceptance	η	(1.595979, 5.298309)	MC generator-level cut. Long-lived final-state particles (πs) in this range.
τ-cuts	$P_T(\tau^+)$ (GeV)	$(5, \infty)$	High P_T required to make it to the VELO
	FD_T (mm)	$(4, \infty)$	VELO hits cannot occur for smaller FD_T
HLT	$\text{PID}_{K^+}(\pi^+)$	$(\text{-}\infty, 4)$	Log-likelihood of the particle being a K^+ rather than a π^+ (to identify a π^+)
	$m_{3\pi}$ (MeV)	$(500, 1825)$	Mass requirement for the τ vertex
	$m_{2\pi}$ (MeV)	$(0, 1670)$	Mass requirement for each two π combination
	m_{corr} (GeV)	$(2, \infty)$	Corrected mass cut on the τ^+
	$\text{DOCA}_{\text{combi}}$ (mm)	$(0, 0.2)$	Minimal distance of approach between two tracks
	χ^2_{vertex}	$[0, 16]$	Quality of the τ vertex
	$\chi^2_{\text{IP}}(\tau^+)$	$[9, \infty)$	χ^2_{PV} contribution from associating the reconstructed τ^+ to the PV based on IP
Sensor	$N_{\text{sensor-crossing}}$	$[1, \infty)$	Number of sensors in the search window
Hit	$N_{\text{velo-hit}}$	$[1, \infty)$	Number of VELO-hits

Table 8: Cuts made for reconstructing $B^+ \to \tau^+\nu_\tau$. The cuts are grouped according to which sub-efficiency they belong. The passing ranges, and a short description for each cut are provided.

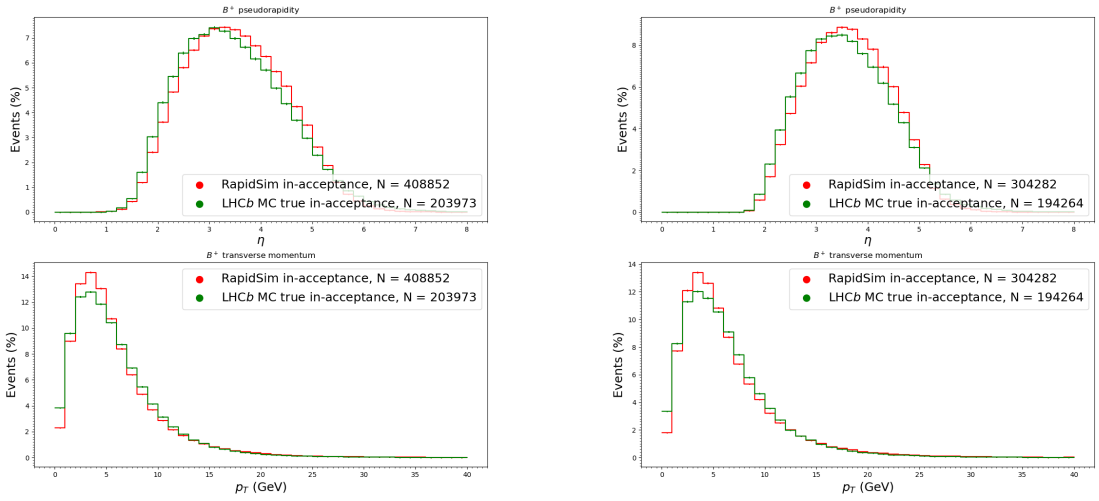
Efficiency	Variable	Range	Description
In-acceptance	η	(1.595979, 5.298309)	MC generator-level cut. Long-lived final-state particles in this range.
D-cuts	$P_T(D^-)$ (GeV)	$(5, \infty)$	High P_T required to make it to the VELO
	$FD_T(D^-)$ (mm)	$(4, \infty)$	Distance from beam line to VELO ≈ 5 mm
HLT	$\text{PID}_{K^+}(\pi^+)$	$(\text{-}\infty, 4)$	Log-likelihood of the particle track being K^+ rather than a π^+ (to identify a π^+).
	$\text{PID}_{K^+}(K^+)$	$(4, \infty)$	Log-likelihood of the particle track being K^+ rather than a π^+ (to identify a K^+).
	m_{D^-} (MeV)	$(1830, 1910)$	Mass requirement for the D^-
	$m_{2\pi}$ (GeV)	$(0, 3.5)$	Mass requirement for 2π combinations
	m_{B^+} (GeV)	$(5, 5.7)$	Mass requirement for the B^+
	DOCA(2π) (mm)	$(0, 0.1)$	Distance of closest approach of two pion tracks
	DIRA($2\pi K$) (cos(rad))	$(0.9995, 1)$	Direction Angle: angle between the PV-TV line and the $2\pi K$ momentum
	χ^2_{vertex}	$(0, 16)$	Quality of the D vertex
	$\chi^2_{IP}(D^-)$	$(9, \infty)$	χ^2_{PV} contribution from associating the reconstructed D^- to the PV based on IP
Sensor	$N_{\text{sensor-crossing}}$	$[1, \infty)$	Number of sensors in the search window
Hit	$N_{\text{velo-hit}}$	$[1, \infty)$	Numer of VELO-hits

Table 9: Cuts made for reconstructing $B^+ \to D^-\pi^+\pi^+$. The cuts are grouped according to which sub-efficiency they belong. The passing ranges, and a short description for each cut are provided.

6.2 Comparing RapidSim and Monte Carlo Simulation

In this section the LHC*b* MC simulation samples for the two normalisation mode candidates $B^+ \to \tau^+\nu_\tau$ and $B^+ \to D^-\pi^+\pi^+$ are analysed. The purpose of this is twofold: to verify the VELO-hit efficiencies calculated with RapidSim in Section 5.3, and to determine $\epsilon_{D/\tau}$-cuts for both decays (as defined in Sec. 6.1). The primary source of simulated data is the LHC*b* Monte Carlo (MC) samples generated with Gauss and Boole (Sec. 3.3). The data samples produced by these simulations are compared to samples produced with RapidSim, where all samples are cut to be within the MC acceptance. Specifically, this section contains an analysis of the non-reconstructed, or *true* MC data. That means directly accessing the particle tracks and hits in the LHC*b* detector that are present in the MC data. Other aspects may still be reconstructed. For RapidSim, *true* data means that the values are as they were produced by the simulation of the decay, and not smeared. Several kinematic properties of the particles in these simulation samples are compared to those samples as simulated by RapidSim.

The transverse momentum and pseudorapidity are compared between the LHC*b* MC and the RapidSim data for both decays. Plots of these quantities are shown below. Note that the number of events is normalised in order to properly compare the shapes. 2D density plots of these quantities for all data samples are provided in Appendix A.3. Figure 12 shows the MC and RapidSim in-acceptance data samples without any additional constraints. It can be observed that these kinematic properties as simulated by RapidSim and the MC simulation data are largely in agreement. For $B^+ \to \tau^+\nu_\tau$ the transverse momentum is slightly higher around its distribution peak in the RapidSim data, and the pseudorapidity distribution is slightly shifted towards larger values as well. The $B \to D^-\pi^+\pi^+$ plots shows almost the same distributions as the $B^+ \to \tau^+\nu_\tau$ plots, as one would expect. The B^+ meson should exhibit the same kinematics, independent of decay. However, a slight difference can be observed. This difference can be explained by taking into consideration the generator level cut. The decay particles from the $B^+ \to D^-\pi^+\pi^+$ decay cover on average a slightly larger pseudorapidity range. Therefore, more events with the B^+ pseudorapidity on the extreme ends of their distribution will be cut away, resulting in a slightly slimmer pseudorapidity distribution.



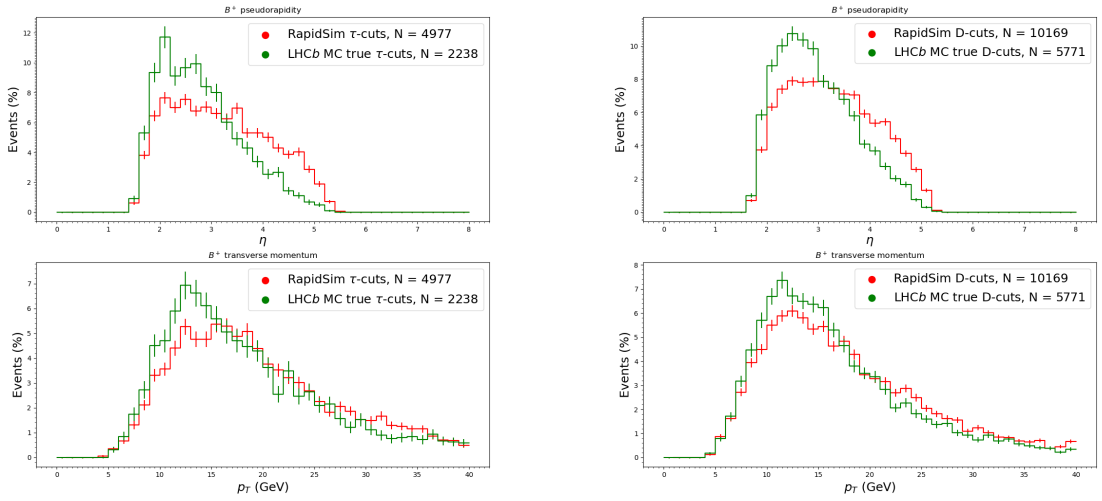
(a) $B^+ \to \tau^+\nu_\tau$ (b) $B \to D^-\pi^+\pi^+$

Figure 12: Pseudorapidity and transverse momentum of in-acceptance B^+ mesons in both decays from RapidSim (red) and LHC*b* MC (green) data.

In the previous section (6.1) the cuts in the efficiency $\epsilon_{D/\tau}$-cuts were defined to make sure that it is possible for the decay to leave a hit in the VELO. The true MC and RapidSim data samples for both decays were filtered based on the cuts that this efficiency consists of. After these cuts were done the distributions of the (true) transverse momentum and pseudorapidity were again compared for RapidSim and the MC simulation in Figure 13.

Due to the cuts on P_T and FD_T the P_T distributions for all data samples have shifted towards higher values. The peak now occurs at $\sim$ 10 - 15 GeV, as apposed to $\sim$ 5 GeV for the in-acceptance samples. The distributions of the transverse momentum of both decays remain largely in agreement between the RapidSim samples and the LHCb MC samples. Still, the P_T distributions from the RapidSim samples tend slightly towards the tail end relative to the LHCb MC sample, resulting in a flatter distribution.

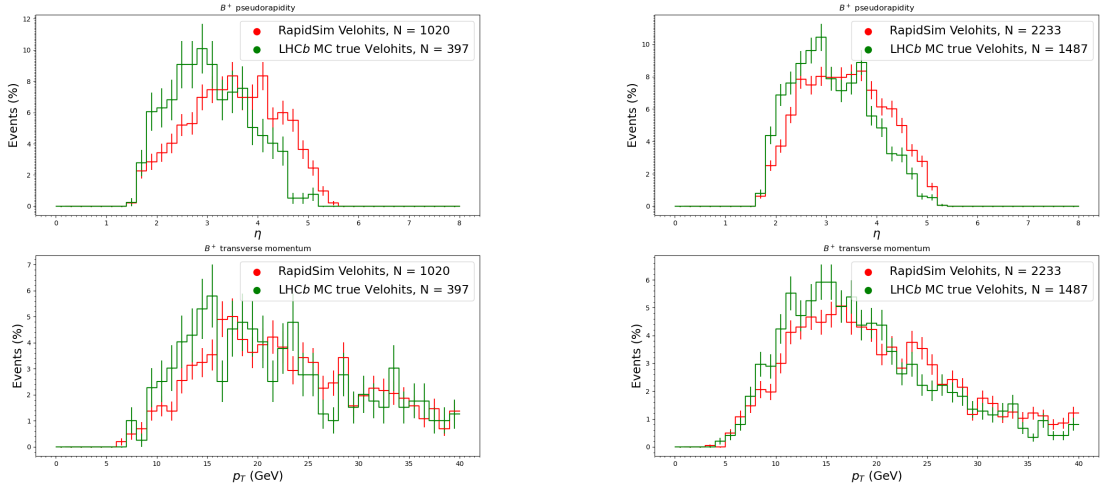
To pass the D/τ cuts, and hence in order to make it to the VELO, large angles (of the τ direction-of-flight relative to the beam line) and therefore small η values are favoured. This is reflected in the pseudorapidity distributions shown in Figure 13. For both decays and simulation methods the distribution shift towards smaller values. However, this shift is much more significant for the LHCb data sample. A possible explanation for this is that in RapidSim the pseudorapidity and transverse momenta are approximated to be independent parameters, while in truth they should be correlated [25]. Therefore, after the D/τ cuts are done, the RapidSim data no longer contains realistic pseudorapidity distributions.



(a) $B^+ \to \tau^+\nu_\tau$ (b) $B^+ \to D^-\pi^+\pi^+$

Figure 13: Pseudorapidity and transverse momentum of D/τ-cut B^+ mesons in both decays from RapidSim (red) and LHCb MC (green) data compared.

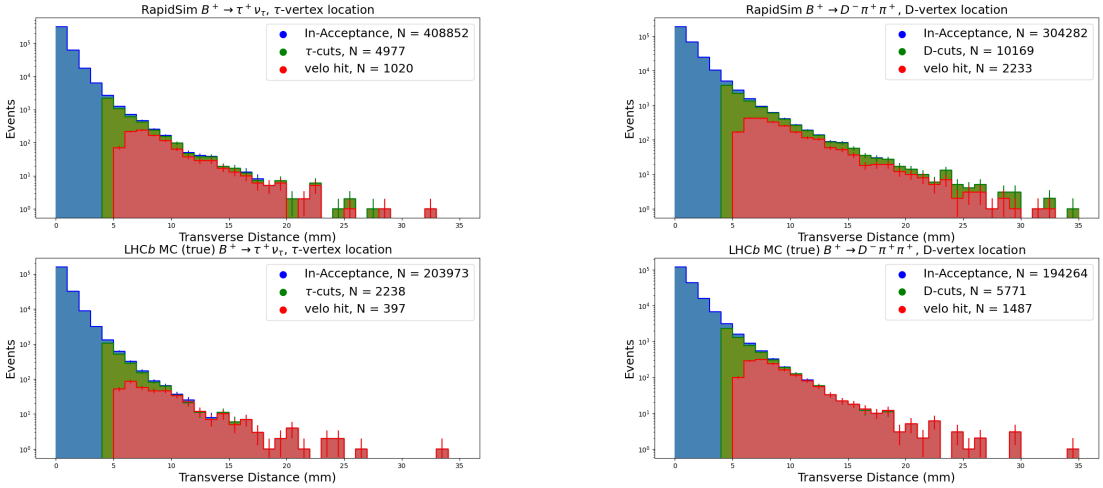
To further compare the LHCb and RapidSim data samples, the data samples were also filtered on leaving a hit in the VELO in order to compare the RapidSim VELO-hit efficiencies from Section 5.3 with those calculated from the LHCb MC data sample. Distributions of the transverse momentum and pseudorapidity are again shown in Figure 14. The P_T distributions remain largely in agreement, while distributions in η differ more strongly. The statistic for $B^+ \to \tau^+\nu_\tau$ is rather low after these cuts.



(a) $B^+ \to \tau^+\nu_\tau$ (b) $B^+ \to D^-\pi^+\pi^+$

Figure 14: Pseudorapidity and transverse momentum of the VELO-hit filtered B^+ mesons in both decays from RapidSim (red) and LHCb MC (green) data compared.

The τ/D cut contains a cut on the transverse distance of the tertiary vertex from the primary vertex. This is equal to the combined transverse flight distance of the B^+ along with the τ^+ or D^-. As a result is may also be useful to plot the transverse flight distance of the TV. Plots of their distributions with a logarithmic y-axis for both simulation samples and for both decays are shown in Figure 15. Additionally, 2D histograms of this transverse flight distance, as well as the transverse momentum of the τ^+ or D^- are shown in Appendix A.4.



(a) $B^+ \to \tau^+\nu_\tau$ (b) $B \to D^-\pi^+\pi^+$

Figure 15: PV-TV transverse distance with the events shown non-normalised with a logarithmic axis. The same samples are shown as were shown in Figure 12, 13 and 14.

As expected based on the lifetime of the D^- meson relative to the τ^+, the transverse flight distance of the D-vertex in the $B^+ \to D^-\pi^+\pi^+$ decay achieves higher values than the τ-vertex in the $B^+ \to \tau^+\nu_\tau$ decay. Furthermore, no VELO hits occur below 5 mm transverse flight distance. This suggests that the transverse flight distance cut can be safely increased to 5 mm.

A subset of events leave a VELO-hit, but do not pass the τ/D-cut. This is to be expected. A long-lived τ^+ or D^- with a low transverse momentum will not pass the cut. This effect can be observed clearly in the 2D histograms in Appendix A.4. Since the D^- has a longer lifetime than the τ^+, and is consequently more likely to be cut away, the amount of τ/D-cut events relative to the VELO-hit filtered events is expected to be lower for $B^+ \to D^-\pi^+\pi^+$ than for $B^+ \to \tau^+\nu_\tau$. In the RapidSim samples this is not observed. For both decays the amount of VELO-hit events is $\sim 1/5$ of the τ/D-cut events. However, such a difference is observed in the LHCb MC samples. Here, for $B^+ \to \tau^+\nu_\tau$ this ratio is $\sim 1/5.5$, while for $B^+ \to D^-\pi^+\pi^+$ this ratio is $\sim 1/4$.

A final observation that can be made from these plots is a difference between the RapidSim and LHCb MC data samples. In the LHCb MC data sample it can be seen that events with a transverse distance of the TV above 10 mm are almost guaranteed to leave a VELO-hit, while the percentage of events leaving a hit in the RapidSim data sample is much lower. This suggest an inaccuracy in the VELO-hit filtering script used to determine the VELO-hit efficiencies.

Table 10 shows the efficiencies of the D/τ cuts and of the VELO hit, both relative to the total in-acceptance yield. Uncertainties in these values are calculated according to the derivation in Appendix A.2. The resulting values from the LHCb simulation and RapidSim are compared for both decays. Although these efficiencies agree relatively well, it can be seen that the efficiency calculations based on the RapidSim data samples tend to be slightly optimistic, resulting in efficiencies that are up to $\sim 25\%$ higher than what was calculated from the LHCb MC samples. Based on these comparisons, the efficiencies calculated in Section 5.3 are taken to be realistic estimates, and the conclusions made based on those results remain valid, making the two decays discussed in this section good normalisation mode candidates.

	$B^+ \to \tau^+\nu_\tau$		$B^+ \to D^-\pi^+\pi^+$	
	RapidSim	LHCb MC	RapidSim	LHCb MC
τ/D cuts	$1.217 \pm 0.017\,\%$	$1.097 \pm 0.023\,\%$	$3.342 \pm 0.034\,\%$	$2.971 \pm 0.040\,\%$
VELO-hit	$20.49 \pm 0.70\,\%$	$17.74 \pm 0.97\,\%$	$21.96 \pm 0.51\,\%$	$25.77 \pm 0.75\,\%$
Cumulative	$0.2494 \pm 0.0078\,\%$	$0.195 \pm 0.098\,\%$	$0.734 \pm 0.016\,\%$	$0.765 \pm 0.020\,\%$

Table 10: Table showing the D/τ-cut and VELO-hit efficiencies.

To conclude this comparison between RapidSim and the LHCb MC simulations. RapidSim approximates pseudorapidity and transverse momentum to be independent parameters, which the LHCb MC simulation does not. Therefore, cuts on one of these parameters can produce unrealistic spectra for the other. Fortunately, this effect is not yet obvious on the generator-level cut. The initial in-acceptance kinematics of the B^+ mesons agree well between both simulation methods. However, once the D/τ cuts and VELO-hit cuts are made, a discrepancy becomes obvious. This could cause further discrepancies in any calculations done with RapidSim involving these parameters. Still, the effect of this difference does not seem to significantly skew the VELO-hit efficiency estimates. As such, the VELO-hit efficiency values from the Section 5.3 will be assumed to be sufficient.

6.3 Reconstructed Data Samples

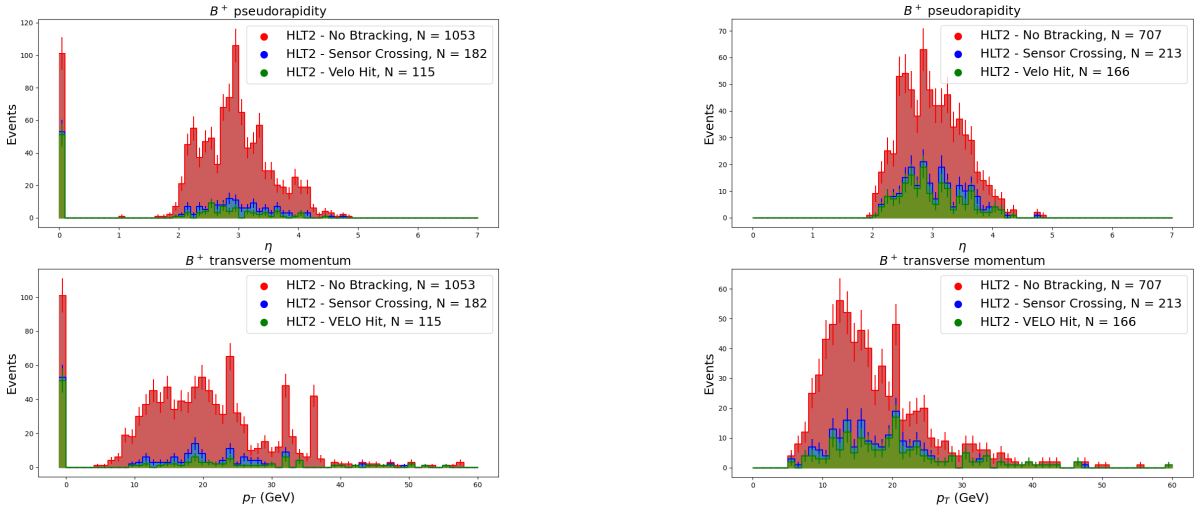
To determine the τ/D-cut and VELO-hit efficiencies *true* MC data samples were needed. However, to determine the HLT2 efficiencies, data samples containing reconstructed events passing the HLT2 trigger lines are necessary. These will be discussed in this section.

To reconstruct $B^+ \to \tau^+ \nu_\tau$ the HLT2 line `make_b2taunu_tau2pipipi_with_btracking` was used to run over the same LHC*b* MC data sample that was used in Section 6.2. Based on the sub-efficiencies defined in Section 6.1, this line was run three times. Once without B-tracking enabled, once requiring a sensor-crossing in the search-window, and once requiring a VELO hit.

To prevent the issue of background[9] appearing in these reconstructed data samples, the events were filtered based on the true PID (particle identification) of the reconstructed particles. This means that MC data was used to match the reconstructed particles to their true identity. However, due to an issue where the wrong MC particles were matched to the reconstructed particles[10], this was done in a more inclusive way. The filter applied to the reconstructed data samples was for the true PID of the final-state particles to correspond to any particles that occur in the decay. For $B^+ \to \tau^+ \nu_\tau$ that implies that $\texttt{MC\_PID}(F) \in \{B^\pm, \tau^\pm, \pi^\pm\}$, where F is a reconstructed final state particle ($\pi^\pm$).

Similarly, the `make_b2dpipi_d2kpipi_with_btracking` line was run with the same three different settings in order to determine the different reconstruction efficiencies for $B^+ \to D^- \pi^+ \pi^+$. This data was filtered on the `MC_PID` of the reconstructed final state particles in a similar way to $B^+ \to \tau^+ \nu_\tau$. In the case of this decay the requirement was $\texttt{MC\_PID}(F) \in \{B^\pm, D^\pm, \pi^\pm, K^\pm\}$, where F is a reconstructed final state particle ($\pi^\pm$, $K^\pm$).

The pseudorapidity and transverse momenta were again plotted for these data-samples (Fig. 16). In Appendix A.3 2D histograms of these quantities are also shown. It appears that in a portion of the $B^+ \to \tau^+ \nu_\tau$ events the pseudorapidity and transverse momentum of the B^+ meson could not be reconstructed. In these events the pseudorapidity of the B^+ defaults to a value of 0 or -1, which is represented by the left-most bars in Histogram 16a.



(a) $B^+ \to \tau^+ \nu_\tau$ (b) $B \to D^- \pi^+ \pi^+$

Figure 16: Transverse momentum and pseudorapidity of the reconstructed B^+ meson for the three different settings of the trigger line for both normalisation mode candidates.

[9] Even though this is simulated data, because these data samples are reconstructed it is still possible for particles belonging to different non-signal decays present in the data to be reconstructed as being a part of the signal decay.

[10] This is possibly caused by the miss-identified MC particle being responsible for all VELO-hits associated to the reconstructed particle track.

In accordance with its higher VELO-hit efficiency, it can be seen that the data samples for $B \to D^- \pi^+ \pi^+$ show higher yields than for $B^+ \to \tau^+ \nu_\tau$. It can be seen that the requirement of a VELO sensor in the search-window (HLT2 - Sensor Crossing data sample) significantly reduces the number of events. The requirement of a VELO-hit reduces this number even further. As a result the statistics are small for these data samples for both decays. The transverse distance of the TV from the PV for these samples is shown in Figure 17. For both decays no events occurred with VELO-hits under 5 mm. This is in agreement with the conclusions made in Section 6.2.

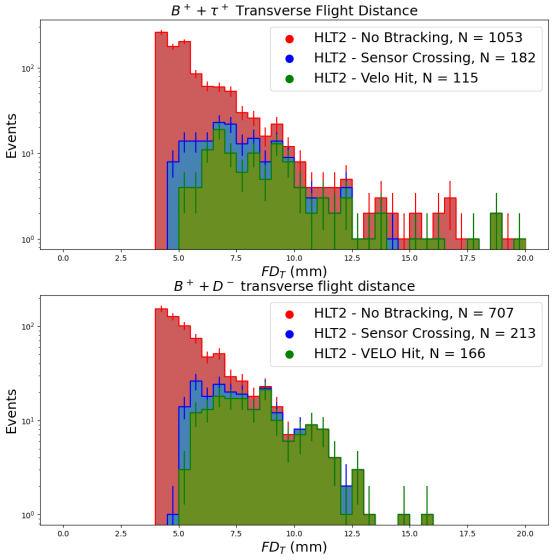


Figure 17: Transverse flight distance of the tertiary decay vertex for the three different reconstructed data samples for $B^+ \to \tau^+ \nu_\tau$ (upper plot) and $B^+ \to D^- \pi^+ \pi^+$ (lower plot).

6.4 Efficiencies

Based on the yield in each data sample, the different sub-efficiencies defined in Section 6.1 can be calculated. The in-acceptance and τ/D-cut data yields from the LHCb true data samples are used, and the HLT - No Btracking, HLT2 - Sensor Crossing, and HLT2 - VELO-hit yields from the reconstructed data samples are used for these calculations. The uncertainty in these efficiencies are calculated in the same way as in Section 6.2, where a Poissonian statistic with the larger sample approximation is used (Appendix A.2). This results in the efficiencies for $B^+ \to \tau^+ \nu_\tau$ shown in Table 11. The yield in the MC sample after each cut is shown, the efficiency of that cut, as well as the cumulative efficiency including previous cuts. These efficiencies are calculated relative to the total in-acceptance yield. The same sub-efficiencies as well as the cumulative efficiencies were also calculated for $B^+ \to D^- \pi^+ \pi^+$. This is shown in Table 12.

	Yield	Sub-Efficiency	Cumulative
In Acceptance	203973	-	-
τ cuts	2238	0.01097 ± 0.00023	0.01097 ± 0.00023
HLT2	1053	0.471 ± 0.018	$(5.16 \pm 0.16) \times 10^{-3}$
HLT2 + Sensor	182	0.173 ± 0.014	$(8.92 \pm 0.66) \times 10^{-4}$
VELO Hit	115	0.631 ± 0.075	$(5.63 \pm 0.53) \times 10^{-4}$

Table 11: Sub-efficiencies for $B^+ \to \tau^+ \nu_\tau$ based on the LHCb MC data sample.

	Yield	Sub-Efficiency	Cumulative
In Acceptance	194264	-	-
D cuts	5771	0.02971 ± 0.00040	0.02971 ± 0.00040
HLT2	707	0.1225 ± 0.0049	$(3.64 \pm 0.14) \times 10^{-3}$
HLT2 + Sensor	213	0.301 ± 0.024	$(1.096 \pm 0.075) \times 10^{-3}$
VELO Hit	166	0.780 ± 0.081	$(8.54 \pm 0.66) \times 10^{-4}$

Table 12: Sub-efficiencies for $B^+ \to D^- \pi^+ \pi^+$ based on the LHCb MC data sample.

The uncertainty in the last two sub-efficiencies is rather high. Especially, for the VELO-hit efficiency it is difficult to formulate qualitative statements on the agreement between the two decays. For the other sub-efficiencies the values differ quite significantly. A larger MC sample would reduce uncertainty and make it possible to formulate more conclusive statements.

The total cumulative efficiency for the reconstruction of these decays is much lower than their VELO-hit efficiency based on the true MC data that was calculated in Section 6.2 (7.65×10^{-3} and 1.95×10^{-3} for $B^+ \to D^- \pi^+ \pi^+$ and $B^+ \to \tau^+ \nu_\tau$ respectively). This implies that the cuts made in the HLT2 lines for these decays are more stringent than the probability of just a hit occurring in the VELO. For $B^+ \to D^- \pi^+ \pi^+$ the yields is reduced by a factor of $\sim 1/9$ due to the HLT2 cuts, while for $B^+ \to \tau^+ \nu_\tau$ this is a factor of $\sim 1/3.5$. Furthermore it can be seen that the total cumulative reconstruction efficiency for the two decays are closer to each other than is suggested by just the VELO-hit efficiency in the true data. This suggests that the HLT2 line for $B^+ \to \tau^+ \nu_\tau$ is more efficient than the $B^+ \to D^- \pi^+ \pi^+$ line. This is also reflected in the cumulative efficiency for the HLT2 line without B-tracking. This is counter-intuitive at first because $B^+ \to \tau^+ \nu_\tau$ cannot be fully reconstructed, while $B^+ \to D^- \pi^+ \pi^+$ can be fully reconstructed. However, due to the additional cuts made in the reconstruction of $B^+ \to D^- \pi^+ \pi^+$ (such as on B^+ mass and direction angle), as well as due to the larger amount of events cut due to the P_T and FD_T cuts, the HLT2 line for this decay is expected to be more stringent than the $B^+ \to \tau^+ \nu_\tau$ line.

Using these values for the total reconstruction efficiency of both decays, estimates of the expected yields can be made in a similar fashion as in Section 5.3. Again an integrated luminosity of 41 fb^{-1} is assumed. These values are likely closer to the true measurement yields than the previous results based on just the VELO-hit, because these new values take into account the effect of the HLT2 reconstruction. Because no LHCb Monte Carlo data was available for $B_c^+ \to \tau^+\nu$, the total efficiency was estimated to be the same as for $B^+ \to \tau^+\nu$ except for a lower VELO-hit efficiency. The total reconstruction efficiency of $B_c^+ \to \tau^+\nu$ was multiplied by the ratio of the VELO hit efficiencies for both decays that were calculated from the RapidSim simulations to get an estimate the total reconstruction efficiency for $B_c^+ \to \tau^+\nu$. This value was then used to estimate a measurement yield for this decay as well. Table 13 shows the results of these calculations. Because these values only serve as crude estimates uncertainties were ignored.

	Total Amount Produced	Total Efficiency	Measurement Yield
$B^+ \to D^-\pi^+\pi^+$	1.81×10^8	8.54×10^{-4}	~ 154574
$B^+ \to \tau^+\nu_\tau$	2.89×10^7	5.63×10^{-4}	~ 16271
$B_c^+ \to \tau^+\nu_\tau$	1.82×10^7	2.59×10^{-5}	~ 471

Table 13: Estimates of measurement yields for three decays based on estimated of their total reconstruction efficiency. These are crude estimates, and uncertainties are assumed to be very large.

The total relative uncertainty of the branching fraction ratio can be calculated as the square sum of the relative uncertainties of the different factors present in the branching fraction ratio:

$$\frac{\delta\mathcal{R}}{\mathcal{R}} = \sqrt{\frac{\delta(\frac{\sigma_{B^+}}{\sigma_{B_c^+}})^2}{(\frac{\sigma_{B^+}}{\sigma_{B_c^+}})^2} + \frac{\delta\epsilon^2_{B_c^+\to\tau^+\nu}}{\epsilon^2_{B_c^+\to\tau^+\nu}} + \frac{\delta\epsilon^2_{\text{norm.}}}{\epsilon^2_{\text{norm.}}} + \frac{\delta\mathcal{N}(B_c^+\to\tau^+\nu)^2}{\mathcal{N}(B_c^+\to\tau^+\nu)^2} + \frac{\delta\mathcal{N}(\text{norm.})^2}{\mathcal{N}(\text{norm.})^2}} \tag{15}$$

In this thesis the normalisation candidates $B^+ \to \tau^+\nu$ and $B^+ \to D^-\pi^+\pi^+$ have been considered. It is likely that the efficiencies of certain cuts that appear in both the reconstruction of $B_c^+ \to \tau^+\nu$ as well as the chosen normalisation mode will cancel out. For example the efficiency of the B_c^+ or VELO-hit will cancel up to the difference in lifetime of these particle. The same holds for the τ^+ and D^-. In this work the effect of such cancelling sub-efficiencies have not been determined.

Estimates of the total uncertainty of $\mathcal{R}$ can however still be made based on the values determined in this thesis. For a lower estimate of the relative uncertainty of $\mathcal{R}$ the contributions from the efficiencies are ignored. In this case the only contributions come from the uncertainty in the production fraction ratio of the B^+ and B_c^+, and from the statistical uncertainty in the measurement yield. Recall that the uncertainty in the branching fraction ratio is very high (23.9% relative uncertainty) due to the use of poorly known theoretical calculations for the branching fraction of $B_c^+ \to J/\psi\mu^+\nu_\mu$. The values for the measurement yield are taken from Table 13, and their uncertainties are calculated by using the Poissonian standard deviation. For a higher estimate none of the uncertainties on the efficiencies are assumed to cancel, in this case the values in Table 11 and Table 12 are used. Where the relative uncertainty for $B_c^+ \to \tau^+\nu$ is estimated to be the same for $B^+ \to \tau^+\nu$. All the values used for the calculation of this estimate are summarized in Table 14. The uncertainties on the efficiencies can be reduced if larger MC datasets are analysed.

Factor	Value	Uncertainty	Relative
$\sigma_{B^+}/\sigma_{B_c^+}$	7.56×10^{-3}	1.81×10^{-3}	23.9%
$\epsilon_{B^+ \to \tau^+ \nu_\tau}$	5.63×10^{-4}	0.53×10^{-4}	9.4%
$\epsilon_{B^+ \to D^- \pi^+ \pi^+}$	8.54×10^{-4}	0.66×10^{-4}	7.7%
$\epsilon_{B_c^+ \to \tau^+ \nu_\tau}$	2.59×10^{-5}	0.24×10^{-5}	9.4%
$\mathcal{N}(B^+ \to \tau^+ \nu_\tau)$	154574	393	0.25%
$\mathcal{N}(B^+ \to D^- \pi^+ \pi^+)$	16271	128	0.79%
$\mathcal{N}(B_c^+ \to \tau^+ \nu_\tau)$	427	21	4.9%

Table 14: Values used in the calculation of $\delta\mathcal{R}/\mathcal{R}$

.

This then results in an estimate for the relative uncertainty for the branching fraction ratio with $B^+ \to D^- \pi^+ \pi^+$ as the normalisation of $\delta\mathcal{R}/\mathcal{R}$ = 24.3-27.2%. If $B^+ \to \tau^+ \nu$ is instead used as the normalisation channel, the resulting relative uncertainty becomes $\delta\mathcal{R}/\mathcal{R}$ = 24.4-27.8%. However, when $B^+ \to \tau^+ \nu$ is used as a normalisation channel the B-tracking algorithm needs to be desperately calibrated, which is likely to introduce additional uncertainties.

It is clear from these estimates that the uncertainty in the production cross-section ratio is the primary contributor to uncertainty in the final value. It should be remembered that these values are rough estimates of the relative uncertainty. It does not take into account the analysis of background decays, or the cancelling of efficiencies, that will inevitably be present in the measurement yield.

7 Real Data

In early 2023, due to a combination of unfortunate events, the part of the LHCb that encloses the VELO volume, the so-called RF foil, was deformed. An issue with the Gas Injection System which adjusts the pressure of the beam vacuum, in combination with a power failure switching off the cooling systems, followed by a failure of the safety systems, all lead to a large overpressure within the VELO volume relative to the beam vacuum. As a consequence the RF foils were significantly deformed, resulting in displacements of up to ~ 15 mm. Due to this issue the VELO module could not fully close around the beam line for the rest of 2023. However, the foil has since been replaced, and starting April 2024 data taking has resumed. The trigger lines for $B^+ \to \tau^+\nu_\tau$ and $B^+ \to D^-\pi^+\pi^+$ have been implemented and data for these decays is being collected since then. Consequently, a first look at the real data for these decays can be useful to validate these lines are working correctly.

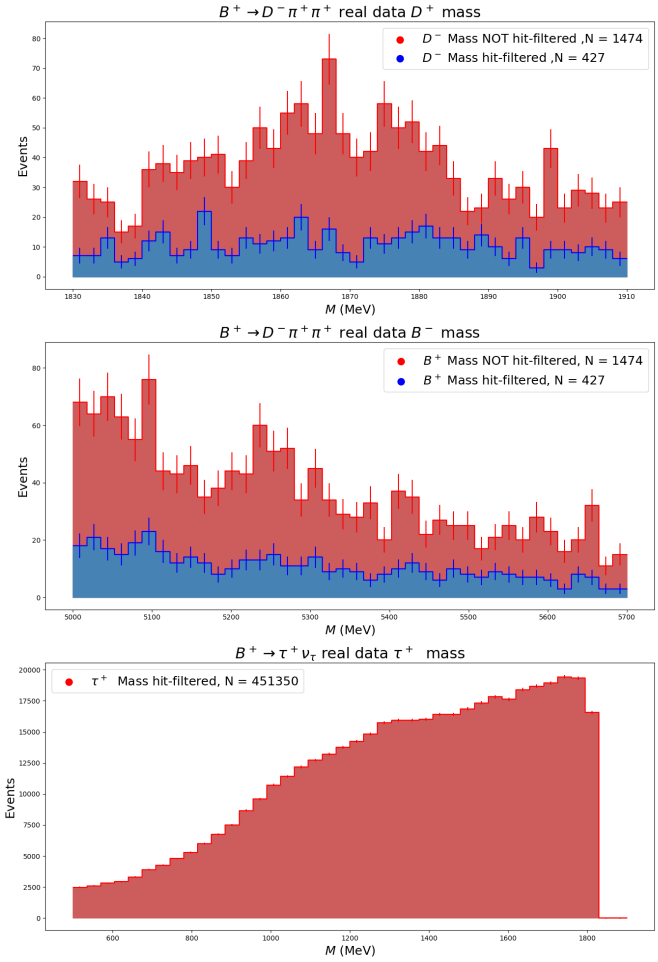


Figure 18: Masses of reconstructed particles from the real data sample.

For this thesis a relatively small data sample was considered. This data was gathered on the 26th of April this year and constitutes an integrated luminosity of $8\,\mathrm{pb}^{-1}$. The data is from a single fill (ID = 9566), and the polarity of the magnets was `MagDown` at the time. In this dataset 456371 events passed the `make_b2taunu_tau2pipipi_with_btracking` line, in accordance with the complete set of cuts defined in Section 6.1, including the VELO-hit requirement. The `make_b2dpipi_d2kpipi_with_btracking` line was implemented without VELO-hit requirement. 1474 events passed the this line line without a VELO-hit. 427 of these events pass an additional VELO-hit requirement. In Figure 18 the masses of the reconstructed B^+ and D^- candidates for $B^+ \to D^-\pi^+\pi^+$ as well as from the reconstructed τ^+ candidates for $B^+ \to \tau^+\nu_\tau$ are shown. The majority of the data collected can be assumed to originate from background, and not the signal decay channels. In particular for $B^+ \to \tau^+\nu_\tau$ the contribution of background is large.

An important quantity in the selection of signal candidates is the corrected mass (defined in Sec. 4.2). In her feasibility study, M.D. Galati [9] calculated the distribution of the corrected mass of the signal τ^+, as well as for several background modes. Figure 19 then gives an estimate of what the corrected mass in the real data is expected to look like for these signal and background modes. To prevent bias in the data analysis for the signal decays, the corrected mass should be blinded until the end of the analysis. As such, no plots of the corrected mass from real data are shown here.

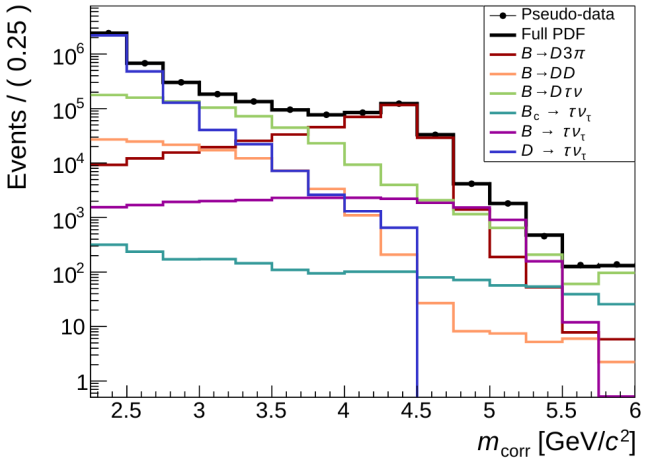
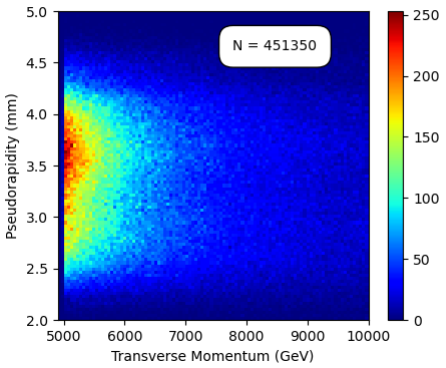


Figure 19: Corrected masses calculations of the signal τ^+ and of particles from several background modes. Taken from [9].

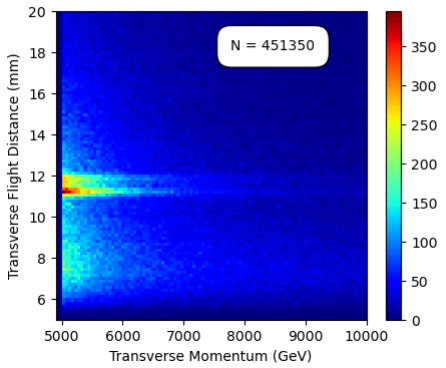
Assuming that the data sample of $8\,\mathrm{pb}^{-1}$ considered here will be representative of the full Run 3 + 4 dataset of $41\,\mathrm{fb}^{-1}$, based on the number of candidates passing the HLT2 lines here, the total number of candidates in Run 3 can be estimated. A naive estimate suggests that $41/(8 \times 10^{-3}) = 5125$ times as many events can be expected for the entirety of Run 3 + 4. Table 15 shows the result from this simple calculation. These estimated candidate yields greatly exceed the expected yields calculated in Table 13 (p. 35). This is again due to the large amount of background present in this data sample.

$\mathcal{L}$	$8\,\mathrm{pb}^{-1}$	$41\,\mathrm{fb}^{-1}$
$B^+ \to \tau^+\nu_\tau$	451350	2.31×10^9
$B^+ \to D^-\pi^+\pi^+$ no VELO-hit	1474	7.55×10^6
$B^+ \to D^-\pi^+\pi^+$ with VELO-hit	427	2.12×10^6

Table 15: Table showing the amount of candidate expected in Run 3 for based on the amount of candidates passing the normalisation mode candidate HLT2 lines.



(a) 2D histogram showing pseudorapidity and transverse momentum of the τ^+ reconstructed with the HLT2 line.



(b) 2D histogram showing transverse flight distance and transverse momentum of the τ^+ reconstructed with the HLT2 line.

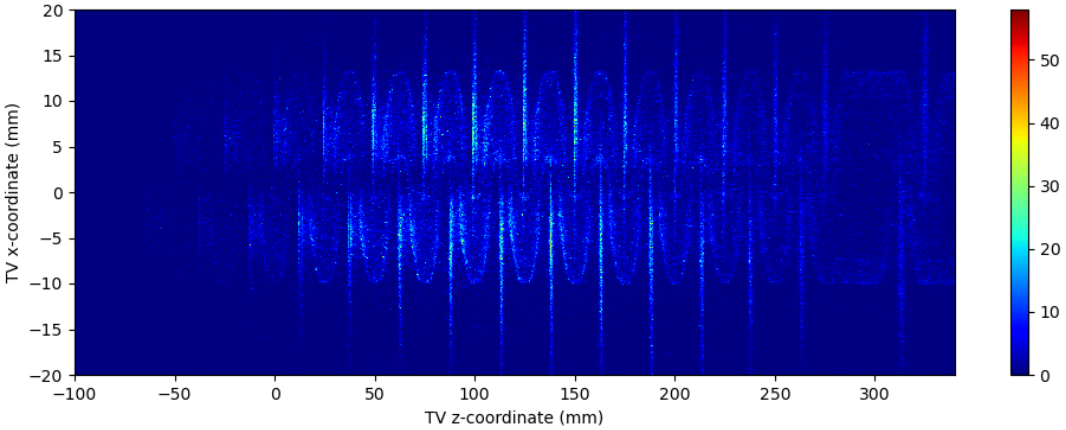
Figure 20: Real data reconstructed with the `make_b2taunu_tau2pipipi_with_btracking` line.

.

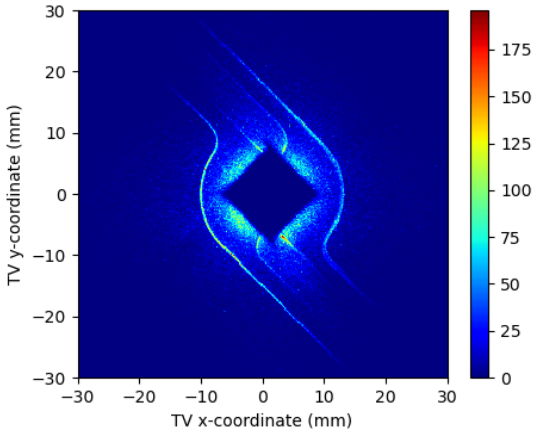
For further analysis of this data the same quantities as were plotted in the previous sections can again be plotted here. In Figure 20a the transverse momentum and the pseudorapidity of the reconstructed τ^+ candidates passing the `make_b2taunu_tau2pipipi_with_btracking` line are plotted. In Figure 20b the transverse momentum and transverse distance between the TV and the PV is shown. It can be observed that a lot of candidates have a transverse flight distance of $\sim$ 11-12 mm.

To investigate why this peak occurs, the coordinates of the TV were plotted. Figure 21 shows these hits projected on the $X-Z$ and $X-Y$ planes respectively. These figures clearly show elements of the VELO module and the RF foil. The peak occurs due to a circular section where the RF foil is thickest. These figures imply that the background is dominated by decays interacting with the detector material. As stated in Section 3.3, RapidSim does not take into account material interactions. Consequently, the accuracy of RapidSim simulations of background decays responsible for these material interactions may be limited.

For the `make_b2dpipi_d2kpipi_with_btracking` line plots were also made. However, due to the small amount of candidates passing the line for this decay, it is difficult to formulate conclusions based on this data. Still, figures are shown in Appendix A.5. Because here the decay can be fully reconstructed the kinematics of B^+ candidates can be investigated. In Figure 26a and 26b the transverse momentum and pseudorapidity for B^+ and D^- candidates are shown, and in Figure 26c the transverse momentum and the PV-TV transverse flight distance is shown.



(a) TV location of candidate taus projected on the $X-Z$ plane.



(b) TV location of candidate taus projected on the $X-Y$ plane.

Figure 21: TV location of τ^+ candidates in the real data sample passing the `make_b2taunu_tau2pipipi_with_btracking` line.

.

8 Conclusions

In recent years the measurement of $R(D^*)$ has put tension on the property of lepton flavour universality in the Standard Model. $B_c^+ \to \tau^+\nu$ has a similar decay structure to $B^0 \to D^-\tau^+\nu_\tau$, which was measured in the $R(D^{(*)})$ measurements. As such a measurement of the branching fraction of this decay can serve as a test for LFU in the SM. Potential violations in LFU are then evidence towards New Physics such as leptoquarks.

There are major difficulties in the measurement of this decay. The τ^+ is short-lived and hence not directly measurable. Therefore, the decay is reconstructed from the sub-decay $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$ final state pions. The difficulty then comes from the fact that the only signal final state particles are the three pions produced in the decay of the τ^+. However, just any reconstructed three-pion vertex does not uniquely identify the signal decay. Furthermore, two neutrinos that are produced in the full decay process. These cannot be measured in the LHCb detector, but hold vital information for the reconstruction of this decay. To be able to measure $B_c^+ \to \tau^+\nu$ then, a new method called B-tracking is used. This algorithm uses direct hits of the $B_{(c)}^+$ and/or τ^+ in the VELO to reconstruct the decay, this way part of the information carried by the neutrinos can be recovered.

In order to minimise uncertainty in the measurement value of the branching fraction of $B_c^+ \to \tau^+\nu$, it should be measured relative to a normalisation channel. Several normalisation mode candidates were considered. Based on measurement yields in Run 1 of the LHCb as well as simulations done in RapidSim, two potential candidates were selected to be investigated more in-depth: $B^+ \to \tau^+\nu$ and $B^+ \to D^-\pi^+\pi^+$. The chosen sub-decays for these normalisation mode candidates were $D^- \to K^+\pi^-\pi^-$ and $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$ respectively.

LHCb Monte Carlo data samples were then analysed for these two decays. First, the kinematic distributions of the B^+ meson in these decays in LHCb MC samples were compared to the RapidSim samples. It was observed that the initial distributions in transverse momentum and pseudorapidity are largely in agreement. However, when cuts are made on these quantities, their distributions begin to deviate. This confirms observations from Andrej Sarnatskiy [25] that these quantities are not correlated in RapidSim, while they are in truth correlated. In addition, it was found that the efficiency of a VELO-hit occurring was similar between RapidSim and the LHCb MC data samples for both decays. However, the probability of these hits occurring did not depend on the transverse distance of the tertiary decay vertex from the primary vertex in the same way for both simulation methods. For the LHCb MC sample it was found that decays exceeding a PV-TV transverse flight distance of 10 mm were almost guaranteed to leave a VELO-hit, while the probability was significantly lower for RapidSim.

The decays were also reconstructed by using their respective HLT2 lines. Using the resulting data samples as well as the true MC data, five sub-efficiencies were calculated for both normalisation mode candidates. The total reconstruction efficiency for $B^+ \to D^-\pi^+\pi^+$ was calculated to be $(8.54 \pm 0.66) \times 10^{-4}$, and for $B^+ \to \tau^+\nu$ this was calculated to be $(5.63 \pm 0.53) \times 10^{-4}$. From these values the expected Run 3+4 measurement yields at 41fb^{-1} could be calculated. The expected yield for $B^+ \to \tau^+\nu$ was ~ 16000, for $B^+ \to D^-\pi^+\pi^+$ the expected yield was calculated to be ~ 150000. Further, using the VELO-hit efficiency of $B_c^+ \to \tau^+\nu$ relative to $B^+ \to \tau^+\nu$ the full reconstruction efficiency of $B_c^+ \to \tau^+\nu$ was estimated to be $\sim 2.6 \times 10^{-5}$, resulting in a yield of ~ 500 for this decay.

A brief analysis of a small sample of real data collected in April 2024 revealed that a large number background passes the HLT2 lines for $B^+ \to D^-\pi^+\pi^+$, and $B^+ \to \tau^+\nu$ especially. More than 2 billion candidates are expected for the latter at the end of Run 4. It was found that this background is dominated by decays occurring in the detector material via material interactions. Consequently, these backgrounds can be removed by applying a detector geometry veto cut.

Due to a large uncertainty introduced through a theoretical calculation used to determine the ration of cross-sections between B^+ and B_c^+ mesons, as well as the low expected yield for $B^+ \to \tau^+\nu$, a large uncertainty in the final ratio of branching fractions was estimated. A naive calculation of the relative uncertainty in the ratio of branching fractions resulted in $\delta\mathcal{R}/\mathcal{R}(B^+ \to \tau^+\nu_\tau) = 24.4\text{-}27.8\%$ and $\delta\mathcal{R}/\mathcal{R}(B^+ \to D^-\pi^+\pi^+) = 24.3\text{-}27.2\%$. These calculations ignore the cancellations of sub-efficiencies, further cuts made to reduce background, as well as the calibration of the B-tracking algorithm.

Still, based on these values, the need for additional calibration and a much larger amount of background when $B^+ \to \tau^+\nu_\tau$ is chosen to be the normalisation channel rather than $B^+ \to D^-\pi^+\pi^+$, suggest that choosing he latter of these normalisation mode candidates will result in the least amount of uncertainty in the final measurement value. Branching fraction ratios for both normalisation modes can be determined as well. A part of potential NP contributions to the $B_c^+ \to \tau^+\nu_\tau$ branching fractions can be cancelled by normalising with respect to $B^+ \to \tau^+\nu_\tau$. As such, a comparison between these two ratios might prove useful in a test of LFU.

Although the uncertainty in $\mathcal{R}$ is large independent of chosen normalisation mode, the NP effects introduced by LFU through pseudoscalar interactions could result in an increase of branching fraction of $B_c^+ \to \tau^+\nu$ of $\sim$ 17-30 times the SM prediction [26]. Consequently, this theory can still be tested in spite of large uncertainties.

9 Outlook

Due to the low statistics of the MC samples after reconstruction, to better determine the reconstruction efficiency, and to see which sub-efficiencies cancel in the branching fraction ratio, larger data samples of both decays would be useful. In addition, an MC sample for $B_c^+ \to \tau^+\nu$ would be useful in order to better determine the efficiencies for this decay specifically. Similar comparisons between RapidSim and that LHC*b* data sample as were done in Section 6.2, can then be made for that decay as well. If larger simulation samples are considered, the uncertainty in the efficiencies will be reduced due to the possibilities of sub-efficiencies cancelling and due to the increased yields. This will help tighten the spread of values calculated for the relative uncertainty in the branching fraction ratio.

As discussed in Section 5.4, $B^+ \to J/\psi K^+$ is a good decay for the calibration of the B-tracking algorithm, where the effect of the VELO-hit of just the B^+ can be isolated. This decay was not deeply analysed in this thesis, but doing so may help to better understand the B-tracking algorithm. Daniel Martinez is currently investigating this decay.

A small sample of real data was considered in Section 7. However, the amount of background was too large and the amount of data too small to be able to formulate any definitive conclusions based on this sample. The next step in the analysis of the real data will be to analyse the entirety of the collected data.

M.D. Galati has performed an analysis of potential background for a measurement of $B_c^+ \to \tau^+\nu$ based on RapidSim data. However, in this thesis it was found that there are slight discrepancies in the simulation data samples produced by RapidSim and the LHC*b* MC samples. The real data shows the dominance of material interactions in the signal candidates. RapidSim does not take this into account. In addition, the pseudorapidity and transverse momentum of particles are not correlated in RapidSim. It was seen that these discrepancies can result in differences in the calculated VELO-hit efficiencies (of up to $\sim 25\%$). As a consequence, the estimates of these backgrounds could likely be improved by using LHC*b* MC data samples instead.

As more data is gathered in the coming months (until the shutdown starting 2026) it becomes very important to understand the backgrounds present in the collected data, and to find the right statistical methods to estimate their yield. It would be valuable to understand which decays are predominantly responsible for the material interactions discussed in Section 7. If everything goes successfully, finally a measurement value of the branching fraction ratio with the chosen normalisation mode can be reported.

10 Acknowledgements

According to the Mertonian norms communism, universalism, disinterestedness, and organised skepticism are the four pillars of science. The latter three of these, one can comprehend theoretically, however I have come to learn that the importance of the value of communism can only be experienced. The existence of an environment where scientific knowledge and the practical skills to achieve this knowledge are not just collected but also shared, is of great importance for the success of science.

Here I wish to acknowledge several people who intuitively understand this principle, and have helped and supported me personally during my master research project.

My most sincere gratitude goes to Mick Mulder for his supportive attitude, his unwavering enthusiasm for physics, and his personal commitment to the gathering and sharing of knowledge. You are the best supervisor a student could wish for, truly. Thank you.

I wish to show sincere gratitude to Maarten van Veghel for the development of software tools that were invaluable for this thesis, as well for the support he gave me to use those tools.

I want to thank Nick Sundstrom for helping me set up my software stack.

I want to thank Sander Bouma for first introducing me to the Groningen LHC*b* group, and for enabling my caffeine addiction.

I want to thank MD, Andrej, Jan, and the rest of the Groningen LHC*b* group for making me feel welcome and supported in the research group. I thank MD personally for letting me use her VELO-hit filtering script.

I want to thank the greater Nikhef Bfys group for putting the community in scientific *community*.

References

[1] LHCb Collaboration. *LHCb collaboration membership list*. 2024. URL: https://lbfence.cern.ch/membership/report/members-list.

[2] Heavy Flavour Averaging Group. *Preliminary average of R(D) and R(D*) for Summer 2023*. 2023. URL: https://hflav-eos.web.cern.ch/hflav-eos/semi/summer23/html/RDsDsstar/RDRDs.html.

[3] Jakub Kwasniak. *Bachelor Thesis: Beyond the Standard Model: Leptoquark Explanations for $R(D)$ and $R(D^*)$ Flavour Anomalies*. 2024.

[4] LHC*b* collaboration. *Standard Set of Performance Numbers*. 2024. URL: https://lhcb.web.cern.ch/speakersbureau/html/PerformanceNumbers.html.

[5] Mike Sikkens. *Bachelor Thesis: First estimate for noise inside the upgraded VELO detector*. 2023.

[6] LHCb collaboration et al. *The LHCb upgrade I*. 2023. arXiv: 2305.10515 [hep-ex].

[7] LHCb Starterkit. *LHCb data flow*. URL: https://lhcb.github.io/starterkit-lessons/first-analysis-steps/dataflow-run3.html.

[8] Lorenzo Paolucci and Dan Thompson. *Simulation and MC requests*. 2022. URL: https://indico.cern.ch/event/1206471/contributions/5157528/attachments/2557548/4407520/Simulation%20and%20MC%20Requests.pdf.

[9] K. De Bruyn et al. *Feasibility of the branching fraction measurement of the decays of $B^+ \to \tau^+\nu_\tau$ and $B_c^+ \to \tau^+\nu_\tau$ at LHCb (unpublished)*.

[10] Nicholas Sunstrom. *Master Thesis: Triggering Invisible Decays*. 2024.

[11] Jelte de Jong. *Master Thesis: Feasibility study of the branching fraction measurements of $B_c^+ \to \tau^+\nu_\tau$ and $B^+ \to \tau^+\nu_\tau$ at LHCb*. 2022.

[12] Roel Aaij et al. "Measurement of the B_c^- meson production fraction and asymmetry in 7 and 13 TeV pp collisions". In: *Phys. Rev. D* 100.11 (2019), p. 112006. DOI: 10.1103/PhysRevD.100.112006. arXiv: 1910.13404 [hep-ex].

[13] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[14] Roel Aaij et al. "Measurement of the ratio of branching fractions and difference in CP asymmetries of the decays $B^+ \to J/\psi\pi^+$ and $B^+ \to J/\psi K^+$". In: *JHEP* 03 (2017), p. 036. DOI: 10.1007/JHEP03(2017)036. arXiv: 1612.06116 [hep-ex].

[15] R. Aaij et al. "First observation and amplitude analysis of the $B^- \to D^+K^-\pi^-$ decay". In: *Phys. Rev. D* 91.9 (2015). [Erratum: Phys.Rev.D 93, 119901 (2016)], p. 092002. DOI: 10.1103/PhysRevD.91.092002. arXiv: 1503.02995 [hep-ex].

[16] Roel Aaij et al. "Measurement of the ratio of B_c^+ branching fractions to $J/\psi\pi^+$ and $J/\psi\mu^+\nu_\mu$". In: *Phys. Rev. D* 90.3 (2014), p. 032009. DOI: 10.1103/PhysRevD.90.032009. arXiv: 1407.2126 [hep-ex].

[17] R. Aaij et al. "Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$". In: *Phys. Rev. Lett.* 120.12 (2018), p. 121801. DOI: 10.1103/PhysRevLett.120.121801. arXiv: 1711.05623 [hep-ex].

[18] R Aaij et al. "Observation of $B_c^+ \to J/\psi D_s^+$ and $B_c^+ \to J/\psi D_s^{*+}$ decays". In: *Phys. Rev. D* 87.11 (2013). [Addendum: Phys.Rev.D 89, 019901 (2014)], p. 112012. DOI: 10.1103/PhysRevD.87.112012. arXiv: 1304.4530 [hep-ex].

[19] B. Kronenbitter et al. "Measurement of the branching fraction of $B^+ \to \tau^+\nu_\tau$ decays with the semileptonic tagging method". In: *Phys. Rev. D* 92.5 (2015), p. 051102. DOI: 10.1103/PhysRevD.92.051102. arXiv: 1503.05613 [hep-ex].

[20] J. P. Lees et al. "Evidence of $B^+ \to \tau^+\nu$ decays with hadronic B tags". In: *Phys. Rev. D* 88.3 (2013), p. 031102. DOI: 10.1103/PhysRevD.88.031102. arXiv: 1207.0698 [hep-ex].

[21] Bernard Aubert et al. "Dalitz Plot Analysis of B- —> D+ pi- pi-". In: *Phys. Rev. D* 79 (2009), p. 112004. DOI: 10.1103/PhysRevD.79.112004. arXiv: 0901.1291 [hep-ex].

[22] Kazuo Abe et al. "Study of B- —> D**0 pi- (D**0 —> D(*)+ pi-) decays". In: *Phys. Rev. D* 69 (2004), p. 112002. DOI: 10.1103/PhysRevD.69.112002. arXiv: hep-ex/0307021.

[23] Yasmine Amhis et al. "Prospects for $B_c^+ \to \tau^+\nu_\tau$ at FCC-ee". In: *Journal of High Energy Physics* 2021.12 (Dec. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep12(2021)133. URL: http://dx.doi.org/10.1007/JHEP12(2021)133.

[24] R.L. Workman et al. *PDG $D^{+/-}$listing*. 2023.

[25] Andrej Sarnatskiy. *Master Thesis: Production of charmed B mesons at high transverse momentum in proton-proton collisions*. 2023.

[26] A.F. Esselink. *Bachelor Thesis: Potential Constraints from a $B_c \to \tau\nu_\tau$ Measurement on the New Physics Parameter Space*. 2023.

[27] Bernard Aubert et al. "Dalitz Plot Analysis of B- —> D+ pi- pi-". In: *Phys. Rev. D* 79 (2009), p. 112004. DOI: 10.1103/PhysRevD.79.112004. arXiv: 0901.1291 [hep-ex].

[28] S. Choudhury et al. "Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays". In: *JHEP* 03 (2021), p. 105. DOI: 10.1007/JHEP03(2021)105. arXiv: 1908.01848 [hep-ex].

[29] J. Z. Bai et al. "A Measurement of J / psi decay widths". In: *Phys. Lett. B* 355 (1995). [Erratum: Phys.Lett.B 363, 267 (1995)], pp. 374–380. DOI: 10.1016/0370-2693(95)00712-T.

[30] V. V. Kiselev. "Exclusive Decays and Lifetime of B_c Meson in QCD sum rules". In: (2003). arXiv: hep-ph/0211021 [hep-ph].

[31] LHCb collaboration et al. *A search for rare $B \to D\mu^+\mu^-$ decays*. 2023. arXiv: 2308.06162 [hep-ex].

[32] Medina Ablikim et al. "Amplitude analysis and branching fraction measurement of $D_s^+ \to K^+K^-\pi^+$". In: *Phys. Rev. D* 104.1 (2021), p. 012016. DOI: 10.1103/PhysRevD.104.012016. arXiv: 2011.08041 [hep-ex].

A Appendices

A.1 Candidate Normalisation Channel Branching Fractions

The table below shows the branching fraction for each of the candidate normalisation channels (Second column), as well as the relevant additional sub-decays. The last column shows the combined Branching fraction of the full decay including its sub-decays. Uncertainties are not considered for these values, because their purpose in this thesis is to calculate crude estimations only.

Candidate Decay	$\mathcal{BF}$	Sub-Decay	Sub-Decay $\mathcal{BF}$	Combined $\mathcal{BF}$
$B^+ \to D^- \pi^+ \pi^p i$	$(1.08 \pm 0.06) \times 10^{-3}$ [27]	$D^- \to K^+ \pi^- \pi^-$	0.0938 ± 0.0016 [13]	1.013×10^{-4}
$B^+ \to J/\psi K^+$	$(1.03 \pm 0.03) \times 10^{-3}$ [28]	$J/\psi \to \mu^+ \mu^-$	0.05961 ± 0.00033 [29]	6.14×10^{-5}
$B_c^+ \to J/\psi \mu^+ \nu_\tau$	1.9% [30]	$J/\psi \to \mu^+ \mu^-$	0.05961 ± 0.00033 [29]	1.13×10^{-3}
$B_c^+ \to J/\psi \pi^+$	1.3×10^{-3} [30]	$J/\psi \to \mu^+ \mu^-$	0.05961 ± 0.00033 [29]	7.75×10^{-5}
$B_c^+ \to J/\psi D_s^+$	2.17×10^{-3} [31]	$J/\psi \to \mu^+ \mu^-$	0.05961 ± 0.00033	6.95×10^{-6}
-	-	$D_s^+ \to K^+ K^- \pi^+$	$5.37 \pm 0.01\%$ [32]	-
$B_c^+ \to \tau^+ \nu_\tau$	1.95×10^{-2} [23]	$\tau^+ \to \pi^+ \pi^+ \pi^- \bar{\nu}_\tau$	9.31% [13]	1.82×10^{-3}
$B^+ \to \tau^+ \nu_\tau$	1.09×10^{-4} [13]	$\tau^+ \to \pi^+ \pi^+ \pi^- \bar{\nu}_\tau$	9.31% [13]	1.015×10^{-5}

A.2 Calculation of Efficiency Uncertainties

The values for the uncertainties reported in this thesis are of course not the *true* efficiencies. Instead they are the statistical estimators for the efficiencies $\hat{\epsilon} = \frac{k}{N} \approx \epsilon_{true}$, where ϵ_{true} is the true efficiency for a certain cut, N is the number of events before the cut, and k is the number of events after the cut. In this case we can assume a Poisson distribution with the large sample approximation for k and N, such that $\delta k = \sqrt{k}$ and $\delta N = \sqrt{N}$. Standard error propagation for quotients gives the following equation for the error on the uncertainty:

$$(\frac{\delta\hat{\epsilon}}{\hat{\epsilon}})^2 = (\frac{\delta k}{k})^2 + (\frac{\delta N}{N})^2 \tag{16}$$

Using Poisson statistics is only valid for uncorrelated events. However, k is a subset of N and therefore correlated. As such, N should be split into two independent datasets k and r such that $N = k + r$, where again $\delta r = \sqrt{r}$. The uncertainty in N can then be expressed as:

$$\delta N = \sqrt{\delta k^2 + \delta r^2} \tag{17}$$

Calculating the uncertainty then yields the result:

$$\begin{aligned}
\delta\hat{\epsilon} &= \hat{\epsilon}\sqrt{(\frac{\delta k}{k})^2 + (\frac{\delta N}{N})^2} \\
&= \hat{\epsilon}\sqrt{(\frac{\delta k}{k})^2 + \frac{\delta k^2 + \delta r^2}{(k+r)^2}} \\
&= \hat{\epsilon}\sqrt{\frac{k}{k^2} + \frac{k+r}{(k+r)^2}} \\
&= \frac{k}{N}\sqrt{\frac{1}{k} + \frac{1}{N}} \\
&= \sqrt{\frac{k(N+k)}{N^3}}
\end{aligned} \tag{18}$$

This equation can be used to calculate the uncertainty of the different sub-efficiencies. The uncertainty of a total uncertainty ϵ_{tot} with n sub-efficiencies ϵ_i can then be calculated as follows:

$$\delta\epsilon_{tot} = (\prod_{i=1}^{n}\epsilon_i)\sqrt{\sum_{i=1}^{n}(\frac{\delta\epsilon_i}{\epsilon_i})^2} \tag{19}$$

A.3 2D Histograms of B^+ Pseudorapidity and Transverse Momentum

$B^+ \to D^- \pi^+ \pi^+$

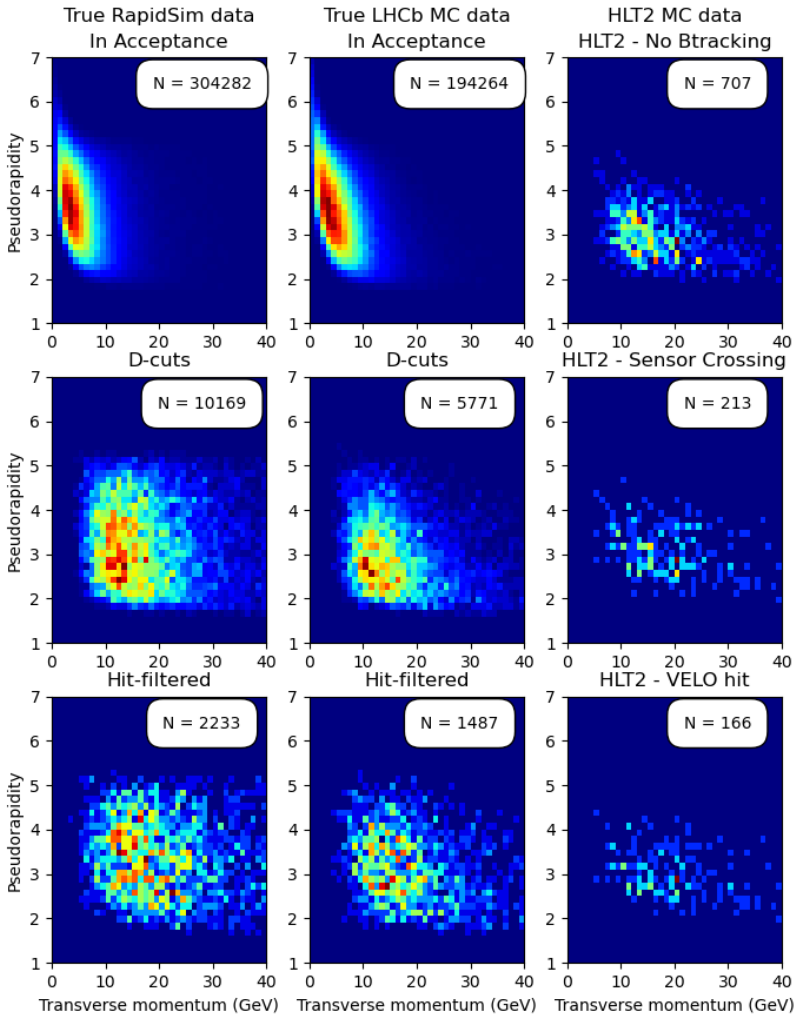


Figure 22: 2D density plots of pseudorapidity (x axes) and transverse momentum (y axes) of the B^+ meson in the decay $B^+ \to D^- \pi^+ \pi^+$ for the true MC data samples (left) and reconstructed samples on the right.

$B^+ \to \tau^+ \nu_\tau$

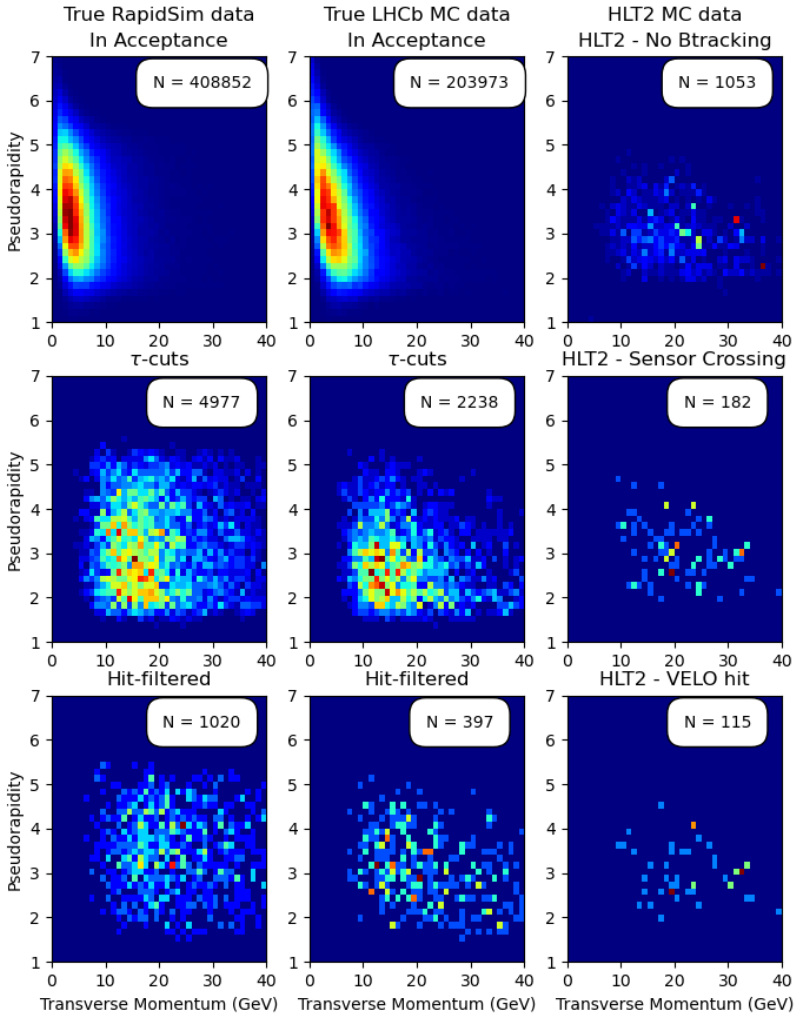


Figure 23: 2D density plots of pseudorapidity (x axis) and transverse momentum (y axis) of the B^+ meson in the decay $B^+ \to \tau^+ \nu_\tau$ for the true MC data samples (left) and reconstructed samples on the right.

A.4 2D Histograms of τ/D FD_T and P_T

$B^+ \to D^- \pi^+ \pi^+$

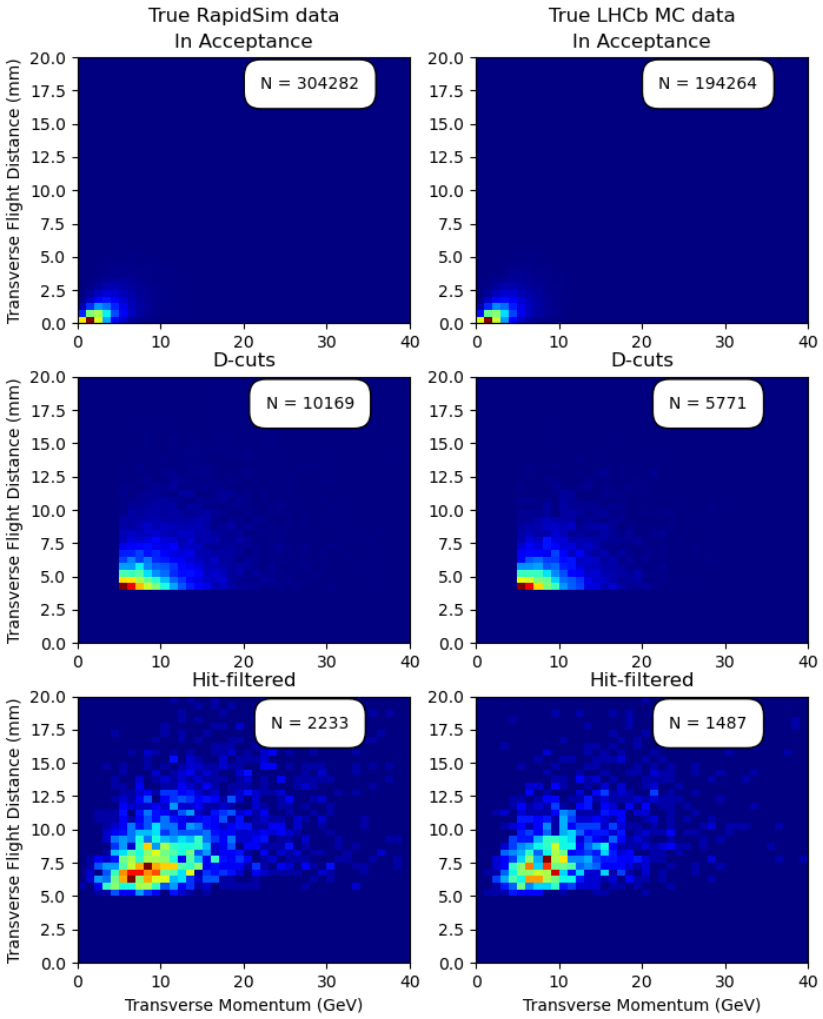


Figure 24: 2D density plots of the transverse momentum (x axes) and transverse flight distance (y axes) of the D^- meson in the decay $B^+ \to D^- \pi^+ \pi^+$ for the true MC data samples (left) and reconstructed samples on the right.

$B^+ \to \tau^+ \nu_\tau$

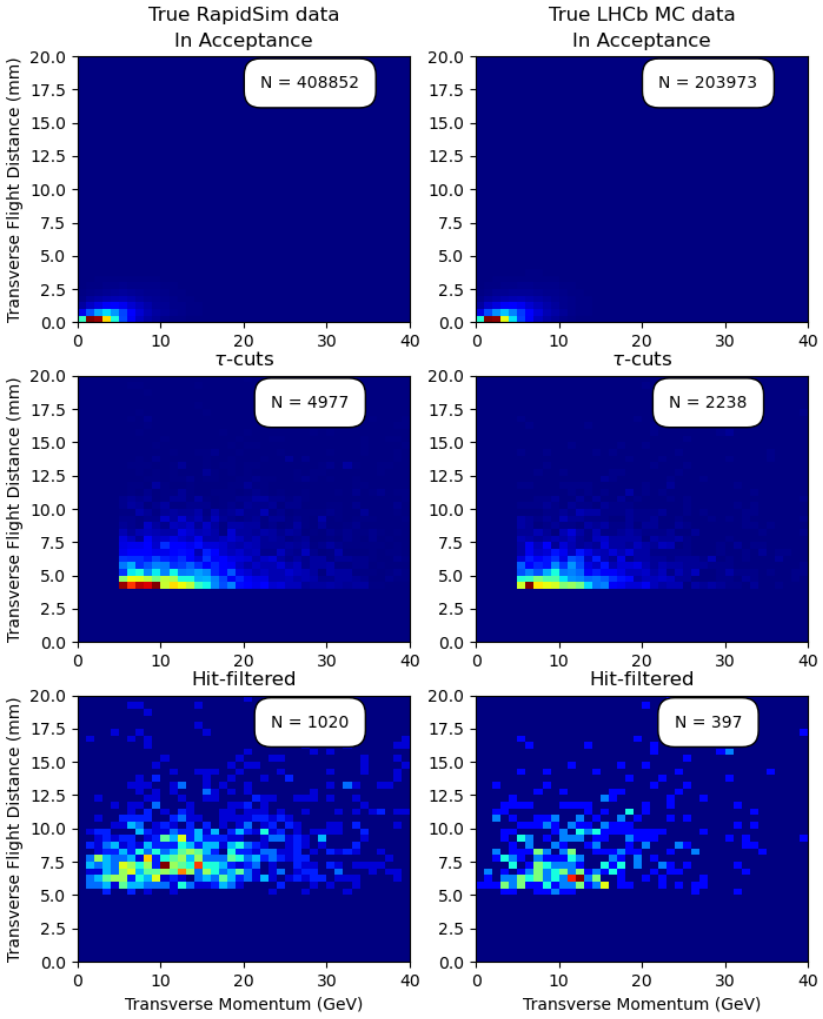
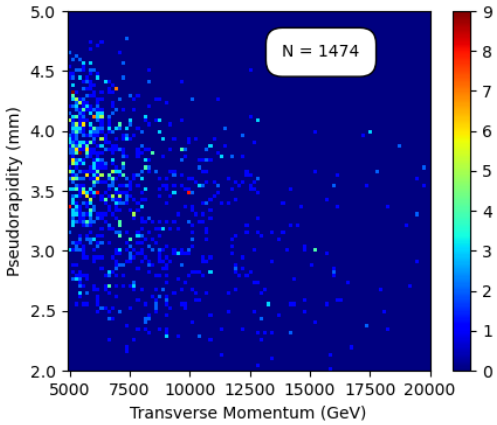
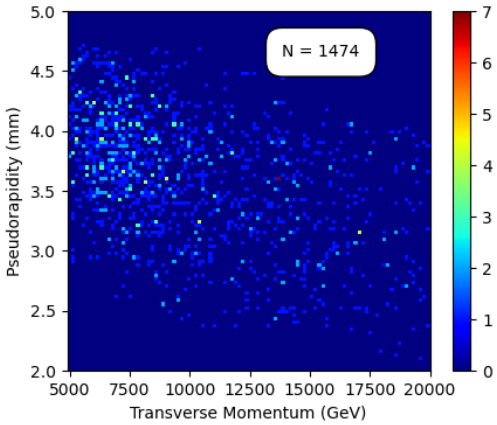


Figure 25: 2D density plots of transverse momentum (x axis) and transverse flight distance (y axis) of the τ^+ in the decay $B^+ \to \tau^+ \nu_\tau$ for the true MC data samples (left) and reconstructed samples on the right.

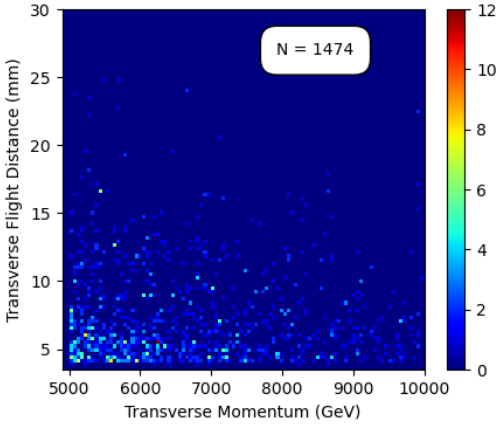
A.5 2D Histograms for the `make_b2dpipi_d2kpipi_with_btracking` line



(a) 2D histogram showing pseudorapidity and transverse momentum of reconstructed B^+ candidates.



(b) 2D histogram showing pseudorapidity and transverse momentum of reconstructed D^- candidates.



(c) 2D histogram showing transverse flight distance and transverse momentum of reconstructed D^- candidates.

Figure 26: Plots showing the first real data reconstructed with the new `make_b2dpipi_d2kpipi_with_btracking` line. Data gathered in April and May 2024 with MagDown polarity.

.