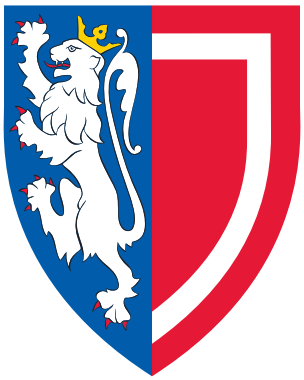
# Searches for $B \to D\mu^+\mu^-$ decays and test-beam studies of TORCH detector prototypes

Thomas H. Hancock

Balliol College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity 2020

This thesis is dedicated to
John Hancock, Sally-Ann Kitts, and Annabel Hancock
for their ceaseless love and support

# Acknowledgements

Firstly, my utmost thanks go to my supervisor, Neville Harnew. His guidance, wisdom, and eagerness to help no matter what problem I was facing were invaluble throughout the course of my studies.

Secondly, I'd like to thank Malcolm John; the $B \to D\mu^+\mu^-$ analysis benefited hugely from his help and guidance. Similar thanks go to Donal Hill. Thanks also go to my fellow Oxford LHCb students, both old and new, who provided me with many helpful discussions and suggestions, along with an office camaraderie I'm going to dearly miss.

My thanks also go to the TORCH collaboration. Alongside many enjoyable face-to-face meetings, the test-beam analyses would not have been possible without their hard work. In particular, I'd like to thank Thierry Gys and Emmy Gabriel, not only for their dedication, but for making the test-beams the very enjoyable experiences that they were.

I'd like to express my gratitude to the Science and Technologies Facilities Council (STFC) and the Oxford Department of Physics for funding my studies and my year abroad at CERN.

Outside of physics, I'd like thank Balliol College, and in particular Balliol MCR, for making my time at Oxford such a fantastic experience. My DPhil would have been very different without the support of its many wonderful members.

Finally, I'd like to thank Mum, Dad and Annabel. Without their constant love and support, I would not be where I am today.

# Abstract

Two complementary pieces of research, linked by the LHCb experiment, are presented. The first uses $8.7\,\text{fb}^{-1}$ of LHCb data to search for a set of rare decays in the $B \to D\mu^+\mu^-$ family: $B^0 \to \overline{D}^0\mu^+\mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0\mu^+\mu^-$, $B_s^0 \to \overline{D}^0 J/\psi$, $B^+ \to D_s^+\mu^+\mu^-$, $B^+ \to D_s^+ J/\psi$, and $B_c^+ \to D_s^+\mu^+\mu^-$. Although no signals are observed, the following world-best upper limits are set on the branching ratios at the 95% confidence level:

$$\begin{aligned}
&\mathcal{B}\left(B^0 \to \overline{D}^0\mu^+\mu^-\right) && < 6.1 \times 10^{-8},\\
&\mathcal{B}\left(B^0 \to \overline{D}^0 J/\psi\right) && < 1.4 \times 10^{-6},\\
&\mathcal{B}\left(B_s^0 \to \overline{D}^0\mu^+\mu^-\right) && < 1.8 \times 10^{-7},\\
&\mathcal{B}\left(B_s^0 \to \overline{D}^0 J/\psi\right) && < 1.4 \times 10^{-6},\\
&\mathcal{B}\left(B^+ \to D_s^+\mu^+\mu^-\right) && < 3.2 \times 10^{-8},\\
&\mathcal{B}\left(B^+ \to D_s^+ J/\psi\right) && < 3.5 \times 10^{-7},\\
\text{and}\quad &\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+\mu^+\mu^-\right) && < 9.6 \times 10^{-8},
\end{aligned}$$

where $f_c$ and $f_u$ are the probabilities that a $b$ quark will hadronize into a $B_c^+$ or $B^+$ mesons respectively. Additionally, the branching ratio of the previously-observed $B_c^+ \to D_s^+ J/\psi$ decay is measured to be

$$\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right) = (1.63 \pm 0.15 \pm 0.13) \times 10^{-5}\,, \tag{1}$$

where the first error is statistical and the second systematic.

The second piece of research details test-beam studies of novel time-of-flight detector for a future upgrade of LHCb, prototypes of the so-called TORCH detector. Three test-beam campaigns have been undertaken, two with a small-scale demonstrator and one with a half-sized LHCb module. Single-photon time resolutions better than 100 ps have been measured; a plan of future work required to reach the target 70 ps is discussed. The photon-counting efficiency of the half-sized demonstrator is also presented. A deficit of photons compared to the LHCb benchmark is seen, however with MCP-PMT photon detectors of improved quantum efficiency, photon yields close to requirement should be anticipated.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

**ACF** . . . . . Anisotropic Conductive Film

**AUC** . . . . . Area Under Curve

**BDT** . . . . . Boosted Decision Tree

**CERN** . . . . . Conseil Européen pour la Recherche Nucléaire

**DCB** . . . . . Double Crystal-Ball (shape)

**DAQ** . . . . . Data AQuisition

**DIRA** . . . . . DIRection Angle

**DIRC** . . . . . Detection of Internally Reflected Cherenkov light (detector)

**DNL** . . . . . Differential Non-Linearity

**DOCA** . . . . . Distance Of Closest Approach

**DUT** . . . . . Device Under Test

**EWP** . . . . . ElectroWeak Penguin

**FCNC** . . . . . Flavour-Changing Neutral Current

**FPGA** . . . . . Field Programmable Gate Array

**GEM** . . . . . Gas Electron Multiplier (chamber)

**HPTDC** . . . . High Performance Time to Digital Converter

**HQET** . . . . . Heavy Quark Effective Theory

**INL** . . . . . Integral Non-Linearity

**KDE** . . . . . Kernal Density Estimation

**LHC** . . . . . Large Hadron Collider

**LHCb** . . . . . Large Hadron Collider beauty (experiment)

**MCP** . . . . . MicroChannel Plate

**MCP-PMT** . . MicroChannel Plate PhotoMultiplier Tube

**MWPC** . . . . Multi-Wire Proportional Chamber

**PCB** . . . . . Printed Circuit Board

**PDF** . . . . . Probability Density Function

**PDG** . . . . . Particle Data Group

**PID** . . . . . Particle IDentification

**PMT** . . . . . PhotoMultiplier Tube

**PS** . . . . . Proton Synchrotron

**PSF** . . . . . Point Spread Function

**PV** . . . . . Primary Vertex

**RICH** . . . . . Ring Imaging CHerenkov (detector)

**ROC** . . . . . Receiver Operating Characteristic (curve)

**SM** . . . . . Standard Model

**TLU** . . . . . Trigger Logic Unit

**ToF** . . . . . Time of Flight

**TORCH** . . . . Time Of internally Reflected CHerenkov light

**TOT** . . . . . Time-Over-Threshold

**QE** . . . . . Quantum Efficiency

**VELO** . . . . . Vertex Locator

# 1 Introduction

## Contents



## 1.1 Overview

This chapter gives a brief introduction to the thesis, which details two independent pieces of research:

searches for $B \to D\mu^+\mu^-$ type decays, and test-beam studies of two TORCH detector prototypes.

These topics are linked through the LHCb experiment; the search is carried out using data collected by the LHCb collaboration, and the TORCH detector is a proposed upgrade for LHCb. As such, Chapter 2 is dedicated to detailing the LHCb detector in relevance to each project.

Chapters 3 and 4 motivate and detail the searches for rare $B \to D\mu^+\mu^-$ type decays. Decays forming either a $D^0$ or $D_s^+$ meson along with a pair of muons or a $J/\psi$ are considered for all possible species of $B$-mesons, resulting in eight possible modes:

$$\begin{aligned} B^0 &\to \overline{D}{}^0\mu^+\mu^- & \qquad B^0 &\to \overline{D}{}^0 J/\psi \\ B_s^0 &\to \overline{D}{}^0\mu^+\mu^- & \qquad B_s^0 &\to \overline{D}{}^0 J/\psi \\ B^+ &\to D_s^+\mu^+\mu^- & \qquad B^+ &\to D_s^+ J/\psi \\ B_c^+ &\to D_s^+\mu^+\mu^- & \qquad B_c^+ &\to D_s^+ J/\psi\,. \end{aligned}$$

Only the $B_c^+ \to D_s^+ J/\psi$ decay has been previously observed, and of the remaining seven, only one has an upper limit set. Hence this is a new and unique avenue of research. Furthermore, at time of writing the LHC is the only currently running collider in the world which can produce $B_c^+$ mesons, meaning the LHC experiments are the only ones where the latter two decays can be observed. For the most part these decays are expected to have branching ratios of $\mathcal{O}\left(10^{-7}\right)$ or less, meaning with the available LHCb dataset they will either show a small signal or enable a world-leading upper limit to be set.

The second topic, namely test-beam studies of two prototype TORCH detectors, is presented in Chapters 5 and 6. TORCH is a proposed upgrade for LHCb, and aims to provide particle identification (PID) in the sub-20 GeV/$c$ momentum region. Details of the design of TORCH are presented in Chapter 5. Two reduced-scale TORCH prototypes were tested in three test-beam periods, where the single-photon time resolution and photon counting efficiencies were measured. These studies are brought together in Chapter 6.

Finally, conclusions are summarised in Chapter 7, along with some closing remarks.

### 1.1.1 Contributions of the Author

The nature of the field of particle physics means no person can work in isolation, and as such the work detailed in this thesis has had input from many colleagues. Nevertheless, the vast majority of each piece of research described herein was carried out by the author, as detailed below.

**The search for $B \to D\mu^+\mu^-$ decays:** The choice of selection, construction of fits to data, calculation of the final results and the performance of systematic studies were all carried out by the author, making use of ROOT, a particle physics data analysis framework, and some pieces of specialised LHCb software. The data were collected by the LHCb collaboration, with the author being involved in data-taking shifts while based at CERN.

**The TORCH test-beam studies:** Unless specified otherwise, all analyses were performed by the author using custom-written software. The author was also heavily involved in the test-beam campaigns, aiding with the set-up, performing and directing the data taking, and was in charge of the "real-time" analysis during the test-beam periods to ensure quality data were being recorded. The TORCH prototypes being tested were built by collaborators.

# 2

# The LHC Machine and the LHCb Detector

## Contents



## 2.1 The CERN Accelerator Complex

Before describing the LHCb detector, the wider context of CERN (Conseil Européen pour la Recherche Nucléaire - "European Organization for Nuclear Research") is discussed. The main focus of CERN in recent years has been the Large Hadron Collider (LHC) due to its unique discovery potential [1], however CERN is host to a large complex of accelerators. Some of these are used to accelerate protons to

the required energy for injection into the LHC, following the chain:

$$\text{LINAC 2} \rightarrow \text{Booster} \rightarrow \text{PS} \rightarrow \text{SPS} \rightarrow \text{LHC},$$

however the same machines feed a variety of other experiments, as indicated in Fig. 2.1. Of particular interest to this thesis is the "East Area", fed by the Proton Synchrotron (PS), as this is where the TORCH test-beam runs were carried out.

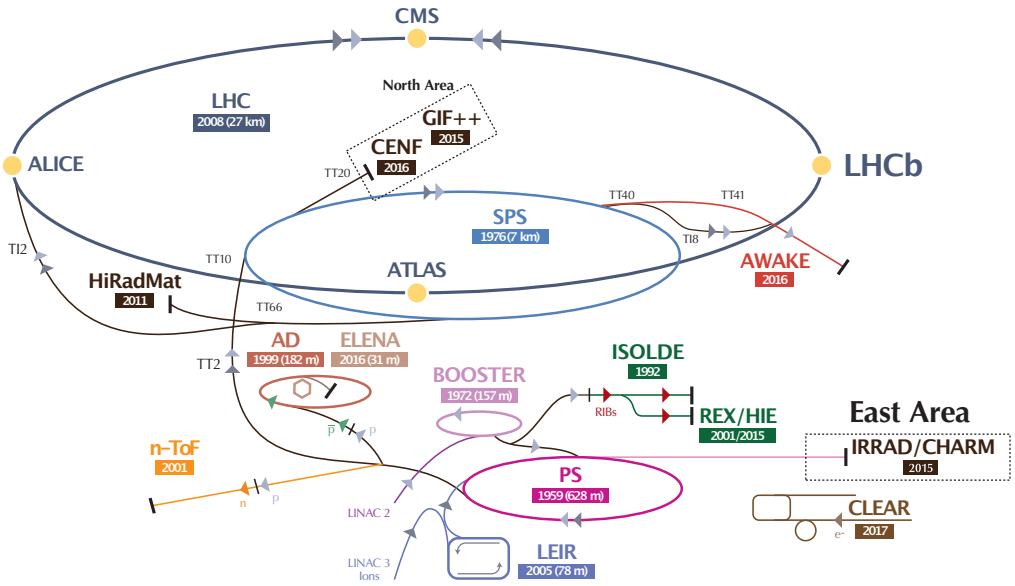



**Figure 2.1:** The CERN accelerator complex, adapted from Ref. [2]. Of particular note for this thesis is the LHCb experiment (upper right) and the East Area (lower right).

Returning to the LHC, this 26.7 km diameter accelerator is a proton-proton collider with a design centre-of-mass energy up to 14 TeV. A maximum of 2556 bunches per beam are accelerated in opposite directions and collide, with a peak crossing frequency of 40 MHz, at four points which host the large experiments: ATLAS, CMS, ALICE, and LHCb.

Two major data-taking periods have been carried out: Run 1, which occurred in 2011 and 2012, and Run 2, which occurred between 2015 and 2018, providing roughly $9\,\text{fb}^{-1}$ of data, broken down by year in Table 2.1. The LHC is currently the only active collider with sufficiently large collision energy to produce $B_c^+$ mesons, making it the only collider where $B_c^+ \rightarrow D_s^+ \mu^+ \mu^-$ and $B_c^+ \rightarrow D_s^+ J/\psi$ decays can be studied.

| Year | Collision Energy (TeV) | Integrated Luminosity ($fb^{-1}$) |
|---|---|---|
| **Run 1** | | |
| 2011 | 7 | 0.98 |
| 2012 | 8 | 1.99 |
| **Run 2** | | |
| 2015 | 13 | 0.28 |
| 2016 | 13 | 1.62 |
| 2017 | 13 | 1.69 |
| 2018 | 13 | 2.16 |

**Table 2.1:** The collision energy and integrated luminosity recorded in each year of LHCb data taking.

## 2.2 The LHCb Detector

### 2.2.1 Overview

Protons are collided at the LHCb interaction point and produce, amongst many other particles, $b$-quarks, which subseqently hadronise. The $b$ and $\bar{b}$-hadrons produced by LHC collisions are typically both boosted in the same hemisphere [3], being produced in either the forward or backward direction, as shown in Fig. 2.2. LHCb is designed to exploit this, with angular coverage in the forward direction of roughly 15 to 300 mrad in the bending plane of the LHCb dipole magnet (see Subsection 2.2.2) and 15 to 250 mrad in the non-bending plane [4].

LHCb is composed of several sub-detector systems, the layout of which is shown in Fig. 2.3 [3]. These systems fall into several broad categories; trackers, RICH (Ring Imaging CHerenkov) detectors, calorimeters, and muon detectors, with further details given in the following sections. Additionally, LHCb employs a dipole magnet to bend the path of charged particles, allowing their charge and momentum to be determined.

### 2.2.2 Magnet

In order to determine the charge and momentum of particles in LHCb, a conventional dipole magnet is employed to bend particle trajectories [3]. Formed of two identical symmetrical coils located between RICH1 and the outer tracking stations, the

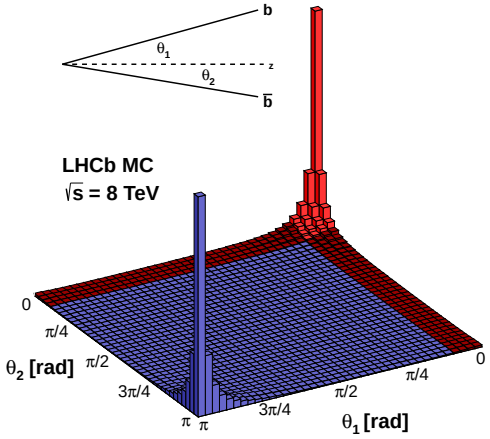



**Figure 2.2:** The $\bar{b}b$ production angles at an LHC collision energy of 8 TeV, with the LHCb angular acceptance shown in red. Taken from Ref. [5].

magnet has an integrated vertical field of 4 Tm. The direction of the field can be reversed between runs, leading to two kinds of datasets: "Mag Up" and "Mag Down". This is important for reducing systematics of CP-violation measurements at LHCb, however has the knock-on effect that the datasets used in the search for $B \to D\mu^+\mu^-$ decays are also split by magnet polarity.

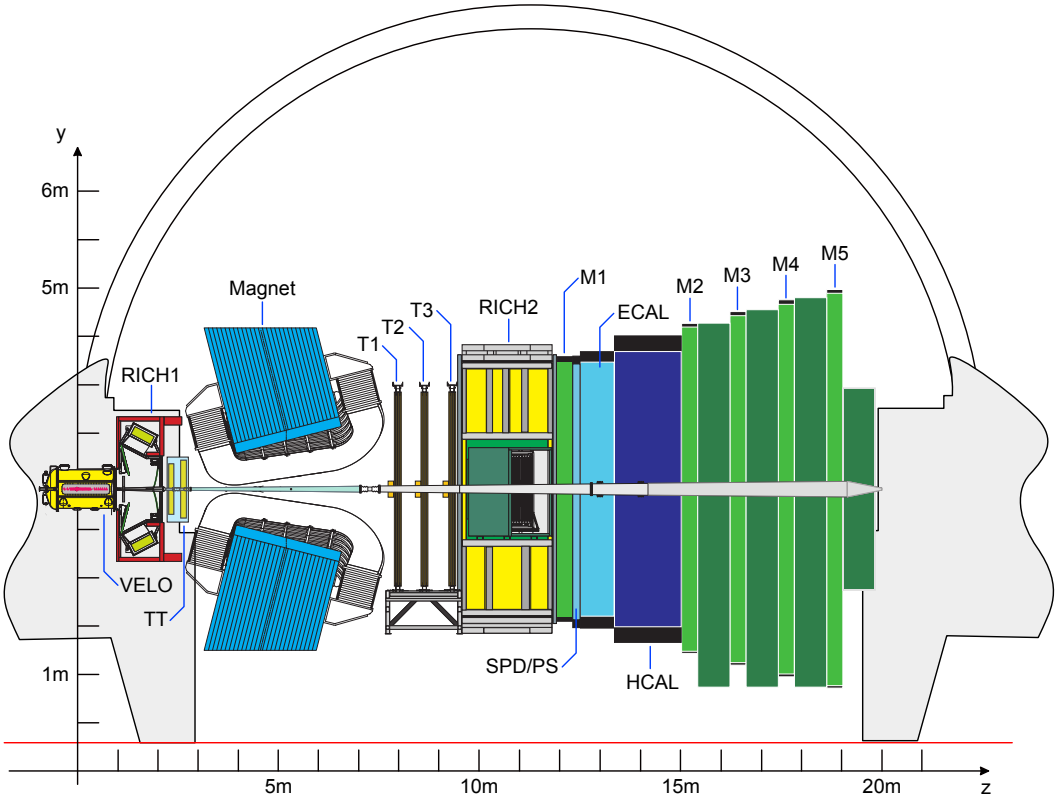



**Figure 2.3:** Side-view of the LHCb detector showing the arrangment of the sub-detector systems. Reproduced from Ref. [3].

### 2.2.3 Tracking

Four tracking systems are employed by LHCb. The VErtex LOcator (VELO) comprises two rows of 21 semicircular silicon micro-strip detectors with alternating micro-strip orientations, as shown in Fig. 2.4a. Sensors with $r$ micro-strip geometry (40 - 102 µm pitch) and $\phi$ geometry (36 - 97 µm pitch), shown in Fig. 2.4b, are employed to allow the reconstruction of 3-dimensional tracks close to the interaction region.

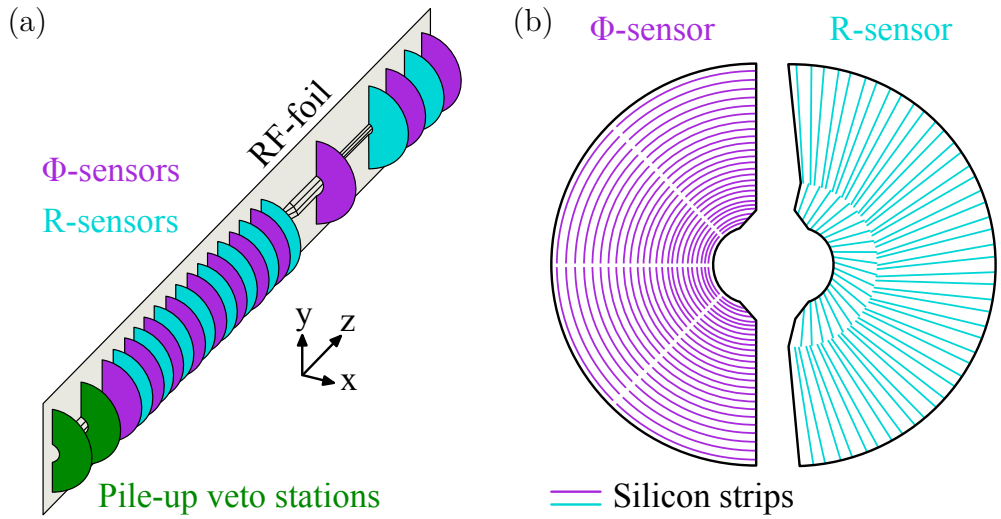



**Figure 2.4:** (a) The arrangement of the semicircular detectors comprising one half of the VELO. Also shown is a simplified RF-foil, which protects the VELO from the harsh LHC environment. Adapted from Ref. [6]. (b) The two geometries of VELO sensor, showing the layout of the silicon micro-strips. Adapted from Ref. [7].

The VELO design facilitates precise tracking around the interaction point, allowing the position of the primary vertex (PV) to be determined with a resolution of 13 μm in the transverse plane and 71 μm along the beam axis [8], and also aiding the identification of $b$-mesons which fly a short distance before decaying. The two opposite halves of the VELO begin in "open" position, where they sit 30 mm either side of the beam. Once stable beam conditions are reached, the modules move to "closed" position, where they encapsulate the beam with roughly a 8 mm radius.

Further tracking detectors are located downstream. The Tracker Turicensis (TT) and Inner Tracker (IT) employ silicon micro-strip sensors, with strip pitches of roughly 200 μm. The TT is positioned downstream of RICH1, before the magnet. The IT system is located downstream of the magnet, interspersed with the Outer Tracker. The hit resolution of both is roughly 50 μm, and the hit detection efficiencies are 99.7% and 99.8% for the TT and IT respectively [4].

The Outer Tracker (OT, composed of T1, T2, and T3 in Fig. 2.3) is a drift-time detector with a single hit resolution of 205 μm. It is positioned downstream of the magnet, and is essential for measuring the curvature of charged particle tracks after they traverse the magnetic field. As the OT has a maximum drift time of $\sim 35$ ns,

which is larger than the bunch spacing, multiple events are read out at once, limiting where tracking information can be used in the trigger (see Subsection 2.2.7). The momentum resolution of the tracking system is $\Delta p/p = 0.5\,\%$ for particles below $20\,\text{GeV}/c$, rising to $\Delta p/p \sim 0.8\,\%$ for particles around $100\,\text{GeV}/c$ [4].

Track reconstruction begins with either the VELO or the OT and IT (from here on treated as a single tracker and refered to as the OT). Hits are combined into tracks which are then propagated to the TT to form "upstream" (VELO and TT) and "downstream" (OT and TT) tracks. Additionally, "long" tracks are formed from hits in the VELO and OT, but may also use hits detected in the TT [9]. These track classifications are shown graphically in Fig. 2.5. Should a particle have multiple tracks of different types, only the track best suited to the analysis is kept. Long tracks are preferred over all others, and it is these tracks that are used in the search for $B \to D\mu^+\mu^-$ decays.

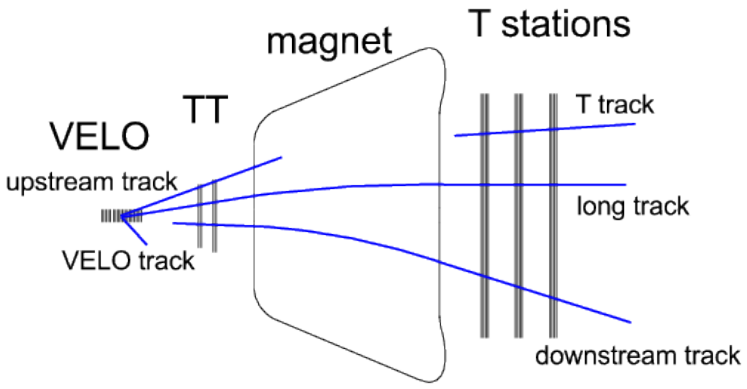



**Figure 2.5:** Classifications of tracks in LHCb depending on which detectors register the hits. Taken from Ref. [9].

### 2.2.4 The RICH detectors

LHCb makes use of two RICH detectors to perform Particle Identification (PID) [3]. The upstream RICH1 covers a momentum range of $\sim 2 - 60\,\text{GeV}/c$ and uses $C_4F_{10}$ (refractive index $n \sim 1.0014$ [10]) as a radiator. In Run 1 an aerogel radiator was also incorporated, however this was removed for Run 2 as the high occupancy limited its performance [11]. The downstream RICH2 uses $CF_4$ ($n \sim 1.0005$ [10]) as

a radiator to cover the momentum range $\sim 15 - 100\,\mathrm{GeV}/c$, and requires a smaller angular acceptance of $\pm 120\,\mathrm{mrad}$ (horizontal) and $\pm 100\,\mathrm{mrad}$ (vertical) to cover the high-momentum region. The fluorocarbon radiators were chosen as they have suitable momentum coverage together with low chromatic dispersion.

In both detectors, the radiated Cherenkov photons are focused using a combination of spherical and flat mirrors. The photons are reflected out of the RICH acceptance to planes of detectors, where they are imaged by HPDs (Hybrid Photon Detectors). In these detectors, photoelections produced by a traditional photocathode are accelerated onto a silicon pixel array, providing a spatial resolution of 0.72 mm, limited by the finite pixel size.

Clear separation of different particle species by Cherenkov angle can be seen in RICH1 over a range of momenta, shown in Fig. 2.6. Further details on the RICH system performance are given in Chapter 5 when discussing the motivation for TORCH.

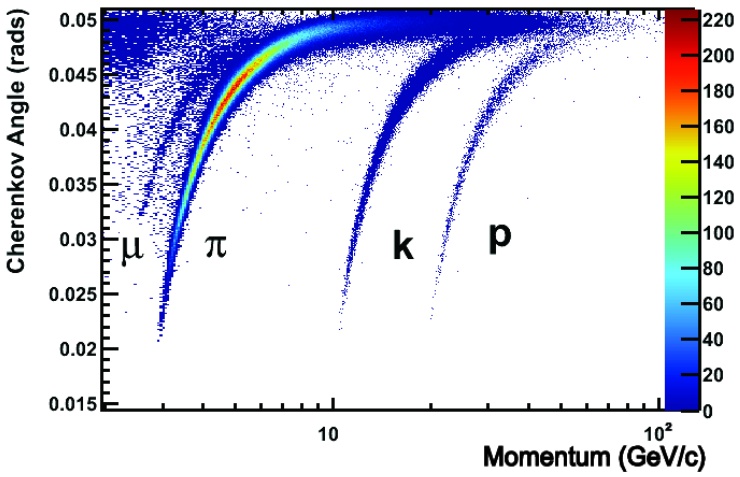



**Figure 2.6:** Reconstructed Cherenkov angle as a function of track momentum for isolated tracks in RICH1 with a $C_4F_{10}$ radiator. Four distinct bands can be seen, corresponding to the particles indicated. This demonstrates that the RICH1 is able to provide PID information over the specified range of momenta. Taken from Ref. [10].

### 2.2.5 The Calorimeters

Further PID is provided by the calorimeter system, which consists primarily of an Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL). Both

use light produced in organic scintillators to measure the energy of electrons and photons (ECAL), and hadrons (ECAL & HCAL) [4]. The energy resolution of the ECAL is parameterised as $\sigma_E/E = (8.5 - 9.5\%)/\sqrt{E} \oplus 0.8\%$ , and for the HCAL, $\sigma_E/E = (69 \pm 5)\,\%/\sqrt{E} \oplus (9 \pm 2)\,\%$ , with $E$ the particle energy in GeV [12].

Additionally, a pair of planes of scintillating pads sit in front of the ECAL (SPD/PS in Fig. 2.3), which provide a measure of the charged track multiplicity to the hardware trigger (see Subsection 2.2.7). Outside of the trigger requirements, the $B \to D\mu^+\mu^-$ search does not make specific use of the calorimeters, so additional focus is reserved for the other LHCb subsystems.

### 2.2.6 The Muon System

While hadron and electromagnetic PID are performed by the RICH and calorimeter systems respectively, muon PID relies primarily on a dedicated subsystem. The muon system is composed of five rectangular stations, named M1 through M5, as shown in Fig. 2.7a and Fig. 2.3. M2-5 are interleaved with iron absorbers which stop other particles and allow the system to maximally leverage the penetrating nature of muons. The calorimeters serve the same purpose between M1 and M2, with M1 placed upstream to provide an improved $p_{\mathrm{T}}$ measurement for the trigger [3].

The individual muon stations are projective, scaling in area perpendicular to the beam so that the entire detector acceptance is covered [3]. Figure 2.7b presents a front-view of a station, showing how each station is divided into four regions, chosen such that there is roughly constant particle flux in each. Each station consists of Multi-Wire Proportional Chambers (MWPCs) in which flows a mixture of gaseous $CO_2$, Ar, and $CF_4$ (in the ratio 40:55:5). The resolution of the system is region-dependent. The best resolution, in the R1 region of M1, is $4 \times 10\,\mathrm{mm}^2$ ($\sigma_x \times \sigma_y$), and the poorest is $150 \times 180\,\mathrm{mm}^2$, in the R4 region of M5.

A total of 19 different designs of MWPC are used in the muon system [14], four for each station, with the R1 region of M1 using Gas Electron Multiplier (GEM) Chambers instead of MWPCs. Excellent muon detection efficiency is essential for searching for $B \to D\mu^+\mu^-$ decays. Efficiencies of 99.3% or greater are measured for

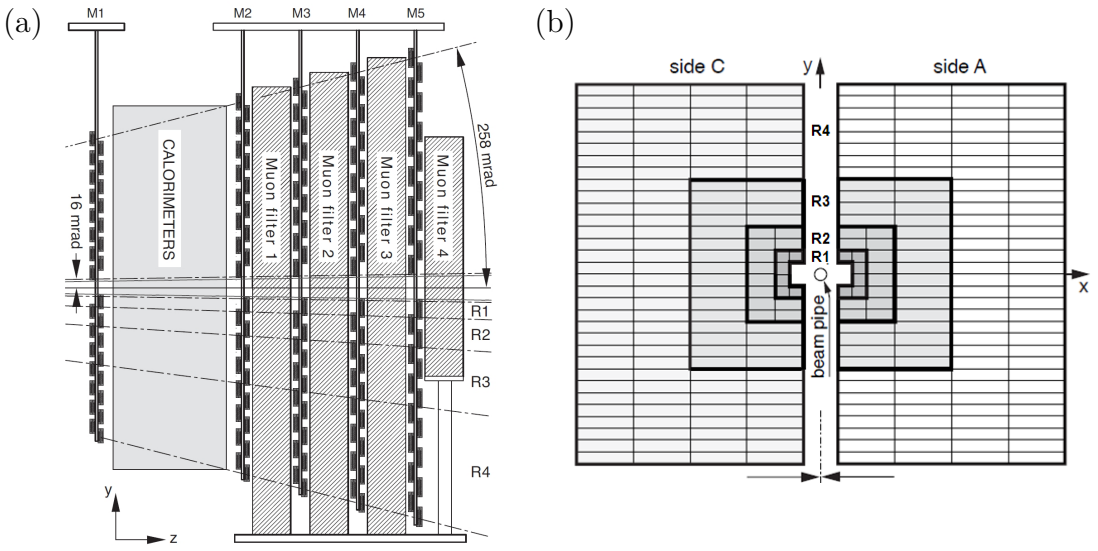



**Figure 2.7:** (a) Side-view of the LHCb muon system, which consists of five rectangular stations interleaved with iron absorber (denoted "muon filler"), taken from Ref. [3]. (b) Front-view of a single muon station, showing the segmentation into four areas of roughly equal particle flux, R1 through R4, taken from Ref. [13].

regions equipped with MWPCs, and 98.7% for the GEM chamber sector, achived by including redundancy in the chamber arrangement: M1 consists of two active layers, and M2-M5 four [13].

### 2.2.7 The LHCb Trigger

The LHCb trigger system is designed to reduce the output data rate from the detector to a rate which can be written to storage. The trigger is split into the hardware-based Level 0 (L0) and software-based High Level Trigger (HLT) systems, with the HLT further divided into two levels, HLT1 and HLT2 [15]. Each stage has access to more information than the previous stage, and reduces the rate subsequently, as shown in Table 2.2. The improvement in the HLT rates between Run 1 and Run 2 is due the addition of 800 extra computing nodes (almost doubling the Run 1 number), as well as a redesigned HLT software.

#### The L0 Trigger

The L0 trigger is implemented purely in hardware and, as it is required to make a decision within 5 μs, has limited information at its disposal.

| Stage in Processing | Data Rate | |
|---|---|---|
| | Run 1 | Run 2 |
| LHCb → L0 | 40 MHz | |
| L0 → HLT1 | 1 MHz | |
| HLT1 → HLT2 | 40/80 kHz | 110 kHz |
| HLT2 → storage | 3/5 kHz | 12.5 kHz |

**Table 2.2:** Data transfer rates between LHCb, the three LHCb trigger levels and permanent storage for each LHC run. For Run 1, the rates for 2011 and 2012 are separated by a slash [15, 16].

The $B \to D\mu^+\mu^-$ search presented in Chapter 4 relies most heavily on the muon trigger. Straight-line tracks are first found in the five muon stations [15]. The two tracks with the largest $p_{\mathrm{T}}$ in each quadrant of the muon system (see Fig. 2.7b) are then considered by the trigger. The event passes the trigger if:

- one of the muons has $p_{\mathrm{T}}$ greater than a set threshold, or
- the product of the $p_{\mathrm{T}}$ of two muons is greater than a set threshold.

The thresholds can be modified to keep the L0 output rate the same with varying beam conditions, and are typically around 3.6 GeV/$c$ (single muon) or 2.0 GeV$^2/c^2$ (dimuon).

A subset of $B \to D\mu^+\mu^-$ events are also triggered by the hadron trigger, using the calorimeters [15]. Deposits with $E_{\mathrm{T}}$ greater than a specific threshold, typically between 2.1 and 2.7 GeV, are required for the event to be retained. Hadrons are distinguished from electrons and photons by requiring an HCAL signal with an ECAL signal in front of it.

### The HLT1 Trigger

At HLT1, a partial event reconstruction is performed; charged particle tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$ which traverse the entire tracking system are found, and additionally a precise reconstruction of the PV is undertaken [15]. The majority of the PID algorithms are unable to run due to timing restrictions, however the clean muon signature allows muon tracks to be identified. For this reason, the requirement of muons which are displaced from the PV or dimuon combinations are

easy to reconstruct. Additional events are selected based purely on the properties of single- or two-track signatures in the event to select non-muonic decays. However, these are also relevant for $B \to D\mu^+\mu^-$ decays in the case the decay is triggered by the daughters of the $D$ meson rather than the muons.

**The HLT2 Trigger**

HLT2 runs a full reconstruction of the event. In Run 1, wherever possible, the same reconstruction was performed online as offline, however in some cases constraints on the run time prevented this. In Run 2 this was improved, with HLT2 running the offline reconstruction as part of the trigger [15].

Many "trigger lines" are run in HLT2, most to select signatures for specific heavy-flavour decay channels. However, of note for the search for $B \to D\mu^+\mu^-$ decays are the *topological trigger lines*, designed to inclusively select $b$-hadron decays by combining charged tracks [17]. Initially two tracks are combined, requiring a Distance Of Closest Approach (DOCA) $< 0.15\,\text{mm}$. An additional track is then added to the pair, requiring the same DOCA requirement between the new track and the 2-body object. This procedure is repeated for adding a fourth track, creating 2-, 3-, and 4-body combinations which are then subject to additional selection criteria to find partially or fully reconstructed decays. The process of adding each new track to the combination of the previous ones greatly improves the HLT2 efficiency on $B \to DX$ decays, making it especially relevant for the $B \to D\mu^+\mu^-$ searches.

# 3

# The Physics behind $B \to D\mu^+\mu^-$ Decays

## Contents



## 3.1 Introduction

The first half of this thesis focuses on the search for $B \to D\mu^+\mu^-$ type decays, with the specific modes of interest given in Table 3.1. In this chapter, the theoretical background is given for how these modes occur, as well as predictions of their branching ratios.

| | |
|---|---|
| $B^0 \to \overline{D}^0 \mu^+ \mu^-$ | $B^0 \to \overline{D}^0 J/\psi$ |
| $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | $B_s^0 \to \overline{D}^0 J/\psi$ |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | $B^+ \to D_s^+ J/\psi$ |
| $B_c^+ \to D_s^+ \mu^+ \mu^-$ | $B_c^+ \to D_s^+ J/\psi$ * |

**Table 3.1:** A summary of the decays being searched for and discussed in this thesis (charge-conjugate modes are implied). The asterisk (*) denotes a mode that has already been observed [18–20].

## 3.2 The Standard Model of Particle Physics

The Standard Model (SM) is the result of a century of discovery and development [21–23]. It consists of two families of spin-½ fermions, one family of spin-1 bosons, and the spin-0 Higgs boson, which was the final SM particle to be observed [24, 25].

Six leptons and six quarks constitute the two fermion families, and each can be sub-divided into three generations:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \qquad \text{(leptons)}\,,$$

and

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix} \qquad \text{(quarks)}\,.$$

The fermions interact through fundamental forces, mediated by the spin-1 bosons: the photon for electromagnetism, the gluon for the strong force, and the $W^\pm$ and $Z^0$ bosons for the weak force [26]. The weak force is of primary importance for $B \to D\mu^+\mu^-$ decays, and so discussions of its properties will form a large part of this chapter.

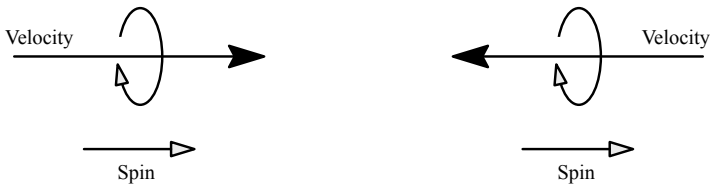
Despite a considerable number of experimental results agreeing with predictions made by the SM, there are still many features of nature the model is unable to explain. The fourth fundamental force, gravity, is not included in the SM, and the SM cannot answer questions such as why charge is quantized, what is dark matter, or why matter survived over anti-matter in the early universe [27]. As such, continuing to search for physics beyond the SM is a priority for the field of particle physics.

## 3.3 The Weak Force

In the decay of hadrons containing a single $b$-quark such as $B \to D\mu^+\mu^-$ type decays, the $b$-quark must change flavour. Electromagnetism and the strong force both conserve flavour, and so these decays can only occur via the weak interaction. Furthermore, amongst the three SM forces, the weak force is particularly interesting as it is the only force mediated by massive gauge bosons; the $W^\pm$ and $Z^0$ have masses of $80.38\,\mathrm{GeV}/c^2$ and $91.19\,\mathrm{GeV}/c^2$ respectively, both far larger than that of the other force mediators [28]. Weak interactions can be categorised into *charged current* and *neutral current* interactions depending on whether they are mediated by a $W^\pm$ or $Z^0$ boson, respectively.

### 3.3.1 Helicity

Particles interact with the weak force via chiral states, which are linear combinations of helicity states [29]. The helicity of a particle is defined as the projection of the particle's spin onto its direction of motion, resulting in positive helicity when spin and direction are aligned. A particle with positive helicity is referred to as "right-handed", and analogously "left-handed" for negative helicity, as shown in Fig. 3.1. For massless particles, the helicity and chirality are equivalent.




(a) Right-handed (b) Left-handed

**Figure 3.1:** Definition of (a) positive (right-handed) and (b) negative (left-handed) helicity. Adapted from Ref. [26].

The $W^\pm$ only couples to left-handed chiral states. However, as helicity states are linear combinations of the chiral states, the $W^\pm$ can couple to both left- and right-handed helicity particles. This introduces a suppression factor most commonly demonstrated through the leptonic decay of the pion: due to lepton universality,

and after correcting for phase space effects, it would be expected that the decays $\pi^- \to \mu^- \nu_\mu$ and $\pi^- \to e^- \nu_e$ occur with equal rates, yet the branching ratio of the muonic decay is $\sim 0.99997$ compared to $\sim 0.00012$ for the electron decay [28].

The pion decay rates can be explained as follows. The pion has zero spin, therefore the decay products must be emitted with opposite spin angular momenta and linear momenta to conserve each quantity in total. With neutrinos being (almost) massless, antineutrinos are always right-handed, therefore the muon or electron must also be right-handed. Were the muon or electron massless, the decays would be forbidden due to the left-handed coupling of the $W^\pm$, however with mass, right-handed helicity contains a non-zero left-handed chiral component, and so the decays can proceed. The size of this left-handed chiral component depends on mass, and hence the electron mode is heavily suppressed relative to the muonic mode due to the electron having significantly less mass.

### 3.3.2 Electroweak Unification

It has been described that the weak interaction prefers particles with left-handed helicity and, that in the ultra-relativistic limit, the weak force only couples to left-handed helicity particles. Yet right-handed particles are seen to interact equally compared to their left-handed counterparts electromagnetically. Similarly, unlike the $W^\pm$, the $Z^0$-boson is experimentally observed to couple to both left- and right-handed particles. This apparent dilema is resolved through unification.

A new quantum number, weak isospin (analogous to isospin in the strong interaction) is introduced, denoted by $I$, with generator $SU(2)_L$ (where the $L$ indicates coupling to left-handed fermions only). The third component, $I_3$, is of particular interest as it is required to be conserved in charged current interactions. With three potential values, three bosons are generated, the $W^1$ with $I_3 = +1$, the $W^2$ with $I_3 = -1$, and a third, $W^3$, with $I_3 = 0$.

The second part of electroweak unification involves replacing the $U(1)_Q$ gauge group from QED with the $U(1)_Y$ group. Rather than coupling to charge, $U(1)_Y$

couples to weak hyper-charge, $Y$, which is related to charge and $I_3$ by

$$Q = I_3 + Y/2\,. \tag{3.1}$$

This replacement introduces another gauge boson: the $B^0$.

From here, spontaneous symmetry breaking leads to the physical particles observed in nature. The $W^1$ and $W^2$ bosons combine to form the $W^\pm$ bosons.

$$W^\pm = \frac{1}{\sqrt{2}}\left(W^1 \mp iW^2\right), \tag{3.2}$$

whilst the neutral electroweak bosons are a result of mixing between the $B^0$ and $W^3$ bosons,

$$\begin{pmatrix}\gamma\\ Z^0\end{pmatrix} = \begin{pmatrix}\cos(\theta_W) & \sin(\theta_W)\\ -\sin(\theta_W) & \cos(\theta_W)\end{pmatrix}\begin{pmatrix}B\\ W^3\end{pmatrix}. \tag{3.3}$$

Here the electroweak mixing angle, $\theta_W$, is a free parameter of the SM [30].

The left-handed preference of the $W^\pm$ can now be explained in terms of weak isospin. As the $W^\pm$ is solely a combination of bosons arising from the $SU(2)_L$ group which generates isospin, it only interacts with particles with non-zero isospin, that is to say left-handed chirality fermions which have $I = \pm\frac{1}{2}$. Fermions with right-handed chirality have $I = 0$, and so cannot interact with the $W^\pm$. The $Z^0$, on the other hand, is a combination of the $W^3$ and $B^0$ bosons, and is hence not restricted in the same way.

### 3.3.3 Quark Mixing

Mixing does not only occur in the force-carrying bosonic sector. Quark mixing is key to explaining phenomena such as the suppressed decays of particles containing strange quarks. Again using $\pi^- \to \mu^-\nu_\mu$ as an example, the rate of the kaon equivalent decay $K^- \to \mu^-\nu_\mu$ is roughly a quarter of $\pi^- \to \mu^-\nu_\mu$ [28], unexpected by phase space considerations but explained by quark mixing.

Cabibbo first introduced the idea of quark mixing, postulating that the mass eigenstates observed in nature are linear combinations of weak eigenstates with which the weak force couples [31]. This was expanded upon by S.L. Glashow,

J. Iliopoulos, and L. Maiani, who introduced the GIM mechanism in relation to the then-known three-quark system [32]. Denoting weak eigenstates with primes, this mixing can be expressed as

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos(\theta_c) & \sin(\theta_c) \\ -\sin(\theta_c) & \cos(\theta_c) \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} , \tag{3.4}$$

where $s$ and $d$ are the quark flavours and the mixing parameter $\theta_c$ is called the Cabibbo angle. The GIM mechanism predicted the existence of a fourth quark [32, 33], the charm quark, which was later discovered at SLAC and BNL [34, 35].

Expanding the $2 \times 2$ matrix to $3 \times 3$ allows the third generation of quarks to be taken into account[1], resulting in the Cabibbo-Koboyashi-Maskawa (CKM) matrix [36]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} , \tag{3.5}$$

where the probability for an up-type quark, $x$, to transition into a down-type quark of a different flavour, $y$, is given by $|V_{xy}|^2$, and the magnitudes of the CKM elements are measured experimentally to be [28]:

$$\begin{array}{lll} |V_{ud}| = 0.97420 \pm 0.00021 & |V_{us}| = 0.2243 \pm 0.0005 & |V_{ub}| = (3.94 \pm 0.36) \times 10^{-3} \\ |V_{cd}| = 0.218 \pm 0.004 & |V_{cs}| = 0.997 \pm 0.017 & |V_{cb}| = (42.2 \pm 0.8) \times 10^{-3} \\ |V_{td}| = (8.1 \pm 0.5) \times 10^{-3} & |V_{ts}| = (39.4 \pm 2.3) \times 10^{-3} & |V_{tb}| = 1.019 \pm 0.025 \ . \end{array}$$

The down-type quarks form an orthonomal basis, therefore the CKM matrix generates a change of basis and so must be unitary. This means that while charged current interactions can facilitate flavour changes, neutral current interactions must conserve flavour. Hence, there are no tree-level flavour-changing neutral currents (FCNCs) allowed by the SM. Conversely, without such cross-generational interactions, mesons such as $K$ and $B$ would be stable.

[1] Although only three quarks had been discovered when the matrix was proposed by Koboyashi and Maskawa.

## 3.4 The Strong Force

Although $B \to D\mu^+\mu^-$ decays are mediated by the weak force as they involve the decay of a hadron, attention must also be paid to the force which governs the interaction of quarks. The strong force is described by the theory of Quantum ChromoDynamics (QCD), so named because it is associated with colour charge [37], analogous to electric charge for electromagnetism and weak hyper-charge for the weak force.

The stength of the QCD interaction is determined by the QCD coupling constant, $\alpha_s$. This constant is a "running" coupling, meaning its strength depends on the energy scale of the interaction. $\alpha_s$ is unique compared to the couplings for electromagnetism and the weak force in that it increases with decreasing energy, or in other words increases with distance. Due to this phenomenon, it is energetically impossible to separate quark pairs [38]; should a pair of quarks be produced together and subsequently move apart, it becomes energetically favourable to produce additional pairs, resulting in the production of colourless hadrons. Similarly, it is impossible to isolate individual quarks from the hadrons they form, and so quarks cannot be observed in isolation. This separation-distance dependent coupling strength is referred to as *asymptotic freedom*.

At high energies, the low coupling strength allows perturbation theory to be applied to perform QCD calculations. Interactions over short-distance scales are treated as perturbations, and their effects calculated via a series of higher-order interactions [39]. Due to the short-range nature of the interactions, the series can be truncated while still providing a good approximation of the system. This approach is referred to as perturbative QCD (pQCD).

Another approach to QCD calculations is Lattice QCD. Lattice QCD involves evaluating the field equations on a quantized grid, with the ability to reduce the grid spacing to approximate the continuum limit. Unfortunately this procedure is not calculated analytically and is very computationally expensive, which imposes constraints on its use [40].

When performing QCD calculations for systems with $B$-mesons, due to its large rest mass, the $b$-quark is essentially static in the rest frame of the hadron [41, 42]. Expanding the QCD correlation functions in terms of $1/m_b$ allows the contributions of the $b$-quark to be removed analytically, significantly simplifying the calculation by isolating the system to the ligher quarks. This creates an effective field theory, hence this method is refered to as Heavy Quark Effective Theory (HQET) [43].

## 3.5 $B \to D\mu^+\mu^-$ Decay Topologies

Of interest in this thesis are $B$-meson decays to final states with a $D$-meson and two muons. Using the rules imposed by electroweak theory and the CKM matrix, topologies for several $B \to D\mu^+\mu^-$ type decays can be determined.

### 3.5.1 Exchange

In decays such as $B^0 \to \overline{D}^0\mu^+\mu^-$ (where the charge-conjugate state is also implied), both quark currents change flavour from the initial state, necessitating a $W^\pm$ exchange as shown in Fig. 3.2. The muons are pair produced from an emitted $\gamma$ or $Z^0$ which can originate from any of the quarks. Due to the CKM elements involved, $B^0 \to \overline{D}^0\mu^+\mu^-$ is favoured over $B^0 \to D^0\mu^+\mu^-$ by a factor of

$$\left|\frac{V_{cb}V_{ud}}{V_{ub}V_{cd}}\right|^2 \approx 2500\,. \tag{3.6}$$

$B_s^0 \to \overline{D}^0\mu^+\mu^-$ decays occur via the same process through substituting the down quark with a strange quark.

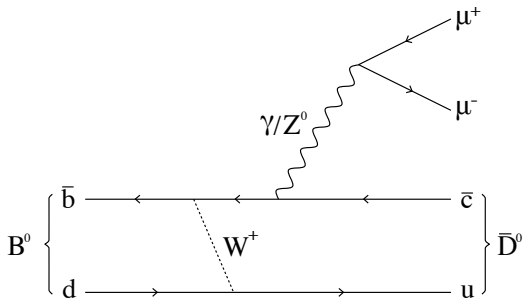



**Figure 3.2:** Exchange topology for $B^0 \to \overline{D}^0\mu^+\mu^-$ decays. By substituting the $d$ for an $s$ quark, $B_s^0 \to \overline{D}^0\mu^+\mu^-$ can also occur via the same mechanism.

### 3.5.2 Annihilation

In $B^\pm$ decays, the two quarks can annihilate to produce an intermediate $W^\pm$ which subsequently pair produces two further quarks. Similar to the exchange topology, a $\gamma$ or $Z^0$ can be emitted from any quark line, leading to the $B \to D\mu^+\mu^-$ final state, as shown in Fig. 3.3. Although $B^+ \to D_s^+\mu^+\mu^-$ is shown, by replacing the up quark with a charm quark, $B_c^+ \to D_s^+\mu^+\mu^-$ can also proceed via this process.

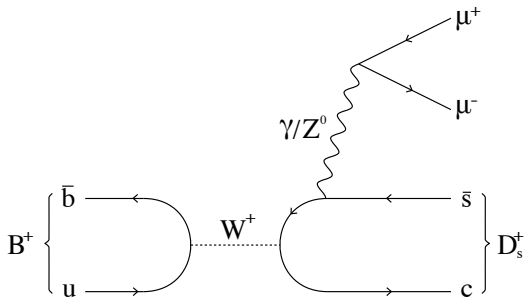



**Figure 3.3:** Annihilation topology for $B^+ \to D_s^+\mu^+\mu^-$. By substituting the $u$ for a $c$, $B_c^+ \to D_s^+\mu^+\mu^-$ can also occur via the same mechanism.

### 3.5.3 Electroweak-Penguin

Possibly the most interesting topology is the electroweak-penguin (EWP) decay, of which $B_c^+ \to D_s^+\mu^+\mu^-$ is an example. In this decay, the second quark in the $B$-meson (the spectator quark) remains unchanged, necessitating a flavour change of only the $b$-quark. This creates a flavour-changing neutral current, which as previously

mentioned is forbidden at tree level in the SM, and therefore must proceed via a loop [44]. To conserve momentum and energy, an additional particle must be emitted in the transition, causing the emission of a $\gamma$ or $Z^0$ which subsequently pair-produces two muons. This topology for $B_c^+ \to D_s^+\mu^+\mu^-$ is shown in Fig. 3.4.

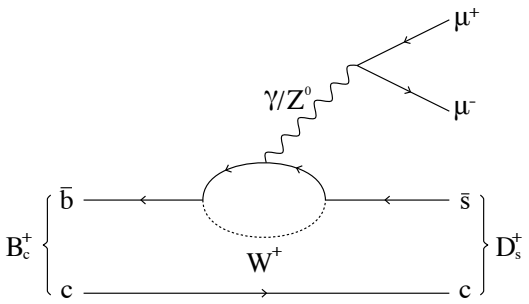


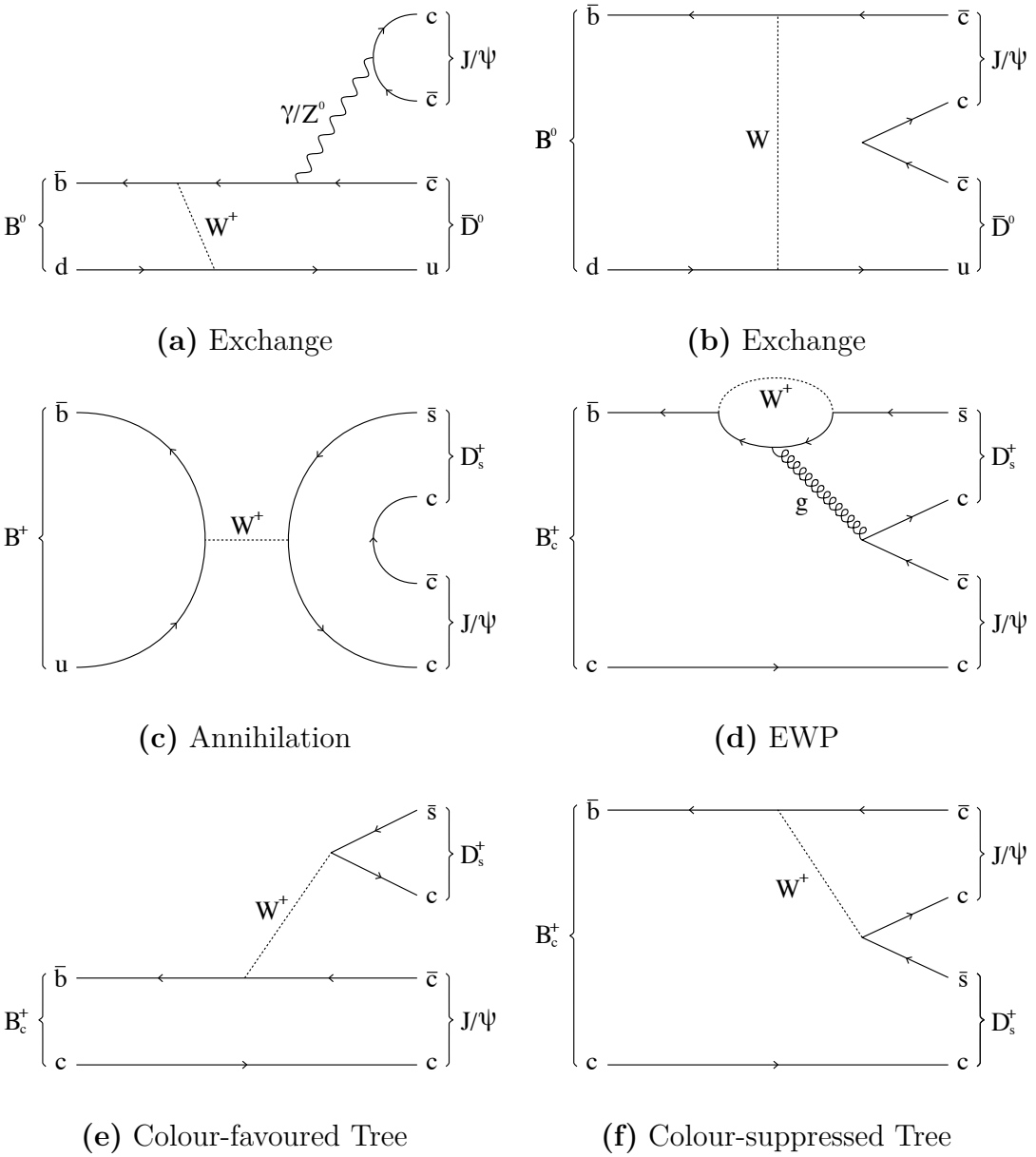

**Figure 3.4:** EWP decay topology for $B_c^+ \to D_s^+\mu^+\mu^-$.

EWP decay topologies provide an excellent platform to search for new physics [44]. As the particles in the loop are produced off-mass shell, these processes are very sensitive to unobserved heavy particles in the loop such as a $Z'$, a hypothetical massive boson with neutral electric charge [44, 45].

### 3.5.4 Intermediate $J/\psi$ modes

Decays of the type $B \to D\mu^+\mu^-$ can also occur through an intermediate $J/\psi$, with the $J/\psi$ subsequently decaying to a pair of muons via annihilation to a neutral boson. Several $B \to DJ/\psi$ topologies are possible, demonstrated in Fig. 3.5. The exchange diagrams show how $B^0 \to \overline{D}^0 J/\psi$ proceeds, and by replacing the down quark with a strange quark, $B_s^0 \to \overline{D}^0 J/\psi$ can occur via the same mechanism. Similar to their non-resonant counterparts, both $B^+ \to D_s^+ J/\psi$ and $B_c^+ \to D_s^+ J/\psi$ can proceed via annihilation (Fig. 3.5c), however due to the requirement of a charmed spectator quark, only $B_c^+ \to D_s^+ J/\psi$ can proceed via the tree diagrams and EWP. It should be noted that possible higher-order diagrams exist for all these modes.

The measured branching ratio for $J/\psi \to \mu^+\mu^-$ is $5.961 \pm 0.033$ % [28]. This mode can compete with hadronic decays of the $J/\psi$ since the latter are suppressed by the OZI rule [46, 47].




**(a)** Exchange **(b)** Exchange

**(c)** Annihilation **(d)** EWP

**(e)** Colour-favoured Tree **(f)** Colour-suppressed Tree

**Figure 3.5:** Possible decay topologies for (a, b) $B^0 \to \overline{D}^0 J/\psi$ and $B_s^0 \to \overline{D}^0 J/\psi$ (with $d$ substituted for $s$ for $B_s^0 \to \overline{D}^0 J/\psi$ decays), (c) $B^+ \to D_s^+ J/\psi$ decays, and (d, e, f) $B_c^+ \to D_s^+ J/\psi$ decays (which are also possible with $u$ substituted for $c$ in (c)).

In a colloqual manner, the OZI rule states that any diagram which can be bisected to split the initial and final state particles, while severing only gluon lines, is suppressed. Such diagrams must contain three gluons (a single gluon would violate conservation of colour, and decays of mesons such as the $J/\psi$, which have a C-parity of $-1$, cannot proceed via two gluon states as they have a C-parity of $+1$) , meaning

there is a factor of $(\alpha_s)^6$ in the decay amplitude and resulting branching ratio. In the case of the $J/\psi$, the result is that the branching fractions of electromagnetic decays are enhanced to a measurable rate relative to hadronic decays.

## 3.6 Theoretical Predictions and Current Limits

Not all the modes of interest have branching ratio predictions available, however those which do are listed in Table 3.2. Lepton universality means predictions with $e^+e^-$ final states should also be valid for modes with $\mu^+\mu^-$.

Whilst the $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B^+ \to D_s^+\mu^+\mu^-$ predictions are of $\mathcal{O}(10^{-10})$, of particular note is the $B^0 \to \overline{D}^0\mu^+\mu^-$ mode, which has two conflicting predictions which differ by almost four orders of magnitude. One prediction was made using HQET and results derived from Lattice QCD [48], whilst the other with perturbative QCD [49]. It has been suggested that the prediction made by Ref. [49] is missing factor of $\alpha^2$ ($\sim 5 \times 10^{-5}$) in the squared matrix element, and multiplying by this factor brings the predictions much closer together.

| Mode | Predicted Branching Ratio | |
|---|---|---|
| $B^0 \to \overline{D}^0\mu^+\mu^-$ | $9.7 \times 10^{-6}$ * | [49] |
| $B^0 \to \overline{D}^0 e^+e^-$ | $2.6 \times 10^{-9}$ | [48] |
| $B_s^0 \to \overline{D}^0 e^+e^-$ | $1.3 \times 10^{-10}$ | [48] |
| $B^+ \to D_s^+ e^+e^-$ | $2.7 \times 10^{-10}$ | [50] |
| $B_c^+ \to D_s^+\mu^+\mu^-$ | $\sim \mathcal{O}(10^{-7})$ † | [28] |

**Table 3.2:** Available theoretical predictions for search modes. All modes consider the range $q^2(\ell^+\ell^-) > 1$ GeV$^2/c^4$ except that marked with *, which considers the range $q^2(\mu^+\mu^-) \in [1, 5]$ GeV$^2/c^4$. Due to lepton universality, modes with an $e^+e^-$ pair should have the same branching ratio as their dimuon counterpart. † from analogy to $B^+ \to K^+\mu^+\mu^-$; no literature reference is available.

The single decay which has been previously observed, $B_c^+ \to D_s^+ J/\psi$, has a ratio of branching fractions relative to the $B_c^+ \to J/\psi\pi^+$ decay of $3.1 \pm 0.5$ [28]. Of the unobserved decays, only $B^0 \to \overline{D}^0 J/\psi$ has a current upper limit: $\mathcal{B}\left(B^0 \to \overline{D}^0 J/\psi\right) < 1.3 \times 10^{-5}$ [51]. No limits have been placed on the others, and this thesis aims to put upper limits or measure the branching fractions of these decays.

# 4

# The Search for $B \to D\mu^+\mu^-$ Decays

## Contents

## 4.1 Analysis Strategy

### 4.1.1 General

The searches for the family of $B \to D\mu^+\mu^-$ decays are made using the full Run 1 and Run 2 LHCb datasets, as detailed in Section 2.1, alongside several simulation samples which were generated by the collaboration using the full LHCb Monte Carlo.

Samples of $D$-mesons are selected, to which pairs of muons are subsequently added. A Boosted Decision Tree (BDT), trained to reduce combinatorial background based on the properties of the candidate decays, is then applied. Fits to the $B$-meson invariant mass distributions are used to measure the number of events and determine the significance of the observed signals.

The analysis was performed fully blind except for the normalisation channels: only once the analysis was fully developed and the systematic uncertainties evaluated was the blinding removed from the invariant mass distributions and the final results extracted.

### 4.1.2 Decays with a $B^0$ or $B_s^0$ meson

A similar selection is used to search for $B^0 \to \overline{D}^0\mu^+\mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0\mu^+\mu^-$, and $B_s^0 \to \overline{D}^0 J/\psi$ decays. Yields for the two non-resonant modes are measured simultaneously by a fit to the $B$-meson invariant mass spectrum, with the same procedure

adopted for the two $J/\psi$ modes. The $\overline{D}^0 \to K^+\pi^-$ channel is used to reconstruct $\overline{D}^0$ candidates, with the branching ratio $\mathcal{B}\left(\overline{D}^0 \to K^+\pi^-\right) = 3.950 \pm 0.031\%$ [28].

Both $B^0 \to \overline{D}^0\mu^+\mu^-$ and the corresponding charge-conjugate decay $\overline{B}^0 \to D^0\mu^+\mu^-$ are searched for. As $\overline{D}^0 \to K^-\pi^+$ is doubly cabbibo suppressed compared to $\overline{D}^0 \to K^+\pi^-$, $D^0$ and $\overline{D}^0$ mesons can be differentiated by the charges of their decay products, given that LHCb has kaon/pion identification. However, with no way to differentiate between $B^0$ or $\overline{B}^0$ mesons, $B^0 \to D^0\mu^+\mu^-$ and $\overline{B}^0 \to \overline{D}^0\mu^+\mu^-$ will both pass the selections, but in the charge-conjugate samples. However, in the SM, $B^0 \to D^0\mu^+\mu^-$ decays are expected to be surpressed by a factor $|V_{ub}V_{cd}/V_{cb}V_{ud}|^2 = 0.00039$ with respect to $B^0 \to \overline{D}^0\mu^+\mu^-$ decays [28], so given the anticipated size of the signal, the contribution from $B^0 \to D^0\mu^+\mu^-$ decays will be negligible[1].

The $B^0 \to (J/\psi \to \mu^+\mu^-)\,(K^{*0} \to K^+\pi^-)$ decay is used to normalise the $B^0 \to \overline{D}^0\mu^+\mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0\mu^+\mu^-$, and $B_s^0 \to \overline{D}^0 J/\psi$ measurements due to having an identical final state and large branching ratio of $(1.27 \pm 0.05) \times 10^{-3}$ [28].

At LHCb, $B_s^0$ and $B^0$ mesons are not produced at the same rate and, as a decay with a primary $B^0$ meson is used to normalise the $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$ measurements, the production fractions must be accounted for. The ratio of production rates is given by the fragmentaton fraction $f_s/f_d$, where $f_x$ is the probability that a $b$ quark will hadronize into a $B_x$ meson. This ratio is applied when determining the upper limits on $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$ decays, and has a measured value $f_s/f_d$ of $0.256 \pm 0.020$ [52].

### 4.1.3 Decays with a $B^+$ or $B_c^+$ meson

Due to the significant difference in mass between $B^+$ and $B_c^+$ mesons, an independent selection is used to select $B_c^+ \to D_s^+\mu^+\mu^-$ and $B_c^+ \to D_s^+ J/\psi$ decays compared to $B^+ \to D_s^+\mu^+\mu^-$ and $B^+ \to D_s^+ J/\psi$ decays. For all four decays, however, $D_s^+$ candidates are reconstructed in the $D_s^+ \to K^+K^-\pi^+$ channel, with $\mathcal{B}\left(D_s^+ \to K^+K^-\pi^+\right) = 5.45 \pm 0.17\%$ [28].

[1] It is worth noting that enhancements from new physics could lead to $B^0 \to D^0\mu^+\mu^-$ decays being non-negligible, however, for the reason previously given, this analysis would not be able to differentiate them from $B^0 \to \overline{D}^0\mu^+\mu^-$ decays.

The $B^+ \to J/\psi K^+$ decay is used to normalise the $B^+ \to D_s^+ \mu^+ \mu^-$, $B^+ \to D_s^+ J/\psi$, and $B_c^+ \to D_s^+ J/\psi$ measurements, with the difference in the number of final state tracks accounted for with efficiency corrections. $B^+ \to J/\psi K^+$ has a branching ratio of $(1.01 \pm 0.03) \times 10^{-3}$ [28].

The non-resonant $B_c^+ \to D_s^+ \mu^+ \mu^-$ decay measurement is normalised against $B_c^+ \to D_s^+ J/\psi$, which has a branching ratio relative to the $B_c^+ \to J/\psi \pi^+$ decay of $3.1 \pm 0.5$ [28]. To convert this relative to the $B^+ \to J/\psi K^+$ decay, which is used to normalise $B_c^+ \to D_s^+ J/\psi$, the ratio [53]

$$\frac{f_c}{f_u} \cdot \frac{\mathcal{B}(B_c^+ \to J/\psi \pi^+)}{\mathcal{B}(B^+ \to J/\psi K^+)} = (0.638 \pm 0.018 \pm 0.009)\,\%$$

is used.

## 4.2 Selection

### 4.2.1 Pre-selection

A pre-selection was applied to the data for all modes, primarily to reduce the size of the dataset being worked with. For individual tracks, cuts on the track-fit $\chi^2$, track momenta, and distance to the primary vertex (PV) were applied. Loose PID cuts were made on the final-state particles, however tighter cuts applied in the later selection supersede these. Additionally, basic requirements on the properties of the $B$-meson were made. In particular, the $B$-meson lifetime was required to be $> 0.05\,\text{ps}$, rather than a traditionally longer-lifetime cut, to ensure a good efficiency for selecting $B_c^+$-meson decays, which have a shorter lifetime than other $B$-mesons.

A complete list of pre-selection cuts is given in Table 4.1 for the $B \to D\mu^+\mu^-$ modes. A simpler set of pre-selection cuts was used to select $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ candidates, given in Table 4.2.

| Variable | Cut |
|---|---|
| **Event-level cuts** | |
| Number of "long tracks" (see Subsection 2.2.3) | $< 500$ |
| Number of PVs | $> 0$ |
| **Cuts applied to all tracks** | |
| Track $\chi^2$/d.o.f | $< 4$ |
| Track momentum transverse to beam | $> 100\,\text{MeV}/c$ |
| Track momentum | $> 1000\,\text{MeV}/c$ |
| Minimum $\chi^2$ distance to the PV | $> 4$ |
| Ghost track probability | $< 0.4$ |
| ***D*-meson related cuts** | |
| Sum of daughter momenta transverse to beam | $> 1800\,\text{MeV}/c$ |
| Sum of daughter momenta | $> 5000\,\text{MeV}/c$ |
| DOCA of kaon(s) to pion | $< 0.5\,\text{mm}$ |
| $D$-meson end vertex $\chi^2$/d.o.f | $< 10$ |
| $\chi^2$-distance from the PV | $> 36$ |
| $D$-meson DIRA | $> 0$ |
| Kaon PID Kaon | $> -10$ |
| Pion PID Kaon | $< 20$ |
| *For $D^0$ mesons only:* | |
| Invariant mass of daughters | $\in 100\,\text{MeV}/c^2$ of the $D^0$ mass |
| *For $D^+/D_s^+$ mesons only:* | |
| Invariant mass of daughters | $> D^+$ mass - $100\,\text{MeV}/c^2$ |
| Invariant mass of daughters | $< D_s^+$ mass + $100\,\text{MeV}/c^2$ |
| **Cuts applied to all muons** | |
| Track momentum | $> 250\,\text{MeV}/c$ |
| Minimum $\chi^2$ distance to the PV | $> 4$ |
| ***B*-meson related cuts** | |
| Invariant Mass | $\in [4900, 7000]\ \text{MeV}/c^2$ |
| Sum of daughter momenta | $> 5000\,\text{MeV}/c$ |
| $B$-vertex $\chi^2$/d.o.f | $< 10$ |
| $B$-lifetime | $> 0.05\,\text{ps}$ |
| $B$-meson Impact Parameter to PV $\chi^2$ | $< 25$ |
| $B$-meson DIRA | $> 0.999$ |
| *At least one daughter track has:* | |
| Momentum | $> 10000\,\text{MeV}/c$ |
| Momentum transverse to beam | $> 1700\,\text{MeV}/c$ |
| $\chi^2$/d.o.f | $< 4$ |
| Minimum $\chi^2$ distance to the PV | $> 16$ |
| Minimum IP to the PV | $> 0.1\,\text{mm}$ |
| *More than one daughter track has:* | |
| Momentum | $> 5000\,\text{MeV}/c$ |
| Momentum transverse to beam | $> 500\,\text{MeV}/c$ |
| $\chi^2$/d.o.f | $< 4$ |

**Table 4.1:** Pre-selection cuts used to select $B \to D\mu^+\mu^-$ candidates. DOCA stands for Distance Of Closest Approach. DIRA stands for DIRection Angle, and is the angle between the reconstructed momentum of the particle and a line drawn between the decay vertex and the PV.

| Variable | Cut |
|---|---|
| ***K*-meson related cuts for $B^+ \to J/\psi K^+$** | |
| $K$-meson $P_T$ | $> 500\,\text{MeV}$ |
| ***K**-meson related cuts for $B^0 \to J/\psi K^{*0}$** | |
| Invariant mass | $\in 70\,\text{MeV}/c^2$ of the $K^*$ mass |
| Momentum transverse to beam | $> 1300\,\text{MeV}/c$ |
| Vertex $\chi^2$ | $< 25$ |
| $K$ $\chi^2$/d.o.f | $< 5$ |
| $\pi$ $\chi^2$/d.o.f | $< 5$ |
| Kaon PID kaon | $> 0$ |
| Pion PID kaon | $< 0$ |
| **Cuts applied to all muons** | |
| Momentum transverse to beam | $> 500\,\text{MeV}$ |
| Muon PID muon | $> 0$ |
| ***J/ψ*-related cuts** | |
| Invariant mass | $\in 150\,\text{MeV}$ of $J/\psi$ mass |
| $\mu^+\mu^-$ $\chi^2$-DOCA | $< 20$ |
| Vertex $\chi^2$ | $< 16$ |
| ***B*-meson related cuts** | |
| Invariant Mass | $\in [5050, 5550]\ \text{MeV}/c^2$ |
| **For $B^0 \to J/\psi K^{*0}$ only** | |
| $B$-vertex $\chi^2$/d.o.f | $< 20$ |
| Invariant mass of $\mu^+\mu^-$ | within $80\,\text{MeV}$ of $J/\psi$ mass |
| $c\tau$ | $> 0.0598$ |
| **For $B^+ \to J/\psi K^+$ only** | |
| $B$-vertex $\chi^2$/d.o.f | $< 10$ |

**Table 4.2:** Pre-selection cuts used to select $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ candidates. DOCA stands for Distance Of Closest Approach.

As part of the pre-selection, a kinematic fit was run on the remaining candidates [54]. This is a least-squares fit which parametrises the decay chain in terms of the positions of vertices, particle momenta, and decay times. The fit constrains the $B$-meson direction to point to the PV, and the mass of the $D$-meson is constrained to its nominal value. For the $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ normalisation modes, the reconstructed $J/\psi$ mass is constrained to its nominal value. The result is an improved $B$-meson mass resolution.

### 4.2.2 Trigger

A specific choice of trigger decisions was required as this improves the accuracy of the selection efficiency estimation and reduces background. A breakdown of the three trigger levels and requirements made at each stage were given in Subsection 2.2.7. At L0, one of the muon or hadron signatures were required. At HLT1, the event had to contain a high-$p_{\mathrm{T}}$ track or a two-track vertex, and in each case the track or vertex had to be displaced from the PV. At HLT2, the event was required to have passed one or more of the topological trigger lines.

Within the trigger, events are broken down into two categories [55]. In "TOS" (Trigger On Signal) events, the trigger requirement is fulfilled by a particle within the reconstructed signal decay (for $B \to D\mu^+\mu^-$ decays, one or more of the muons or $D$-meson daughters). An event is classified as "TIS" (Trigger Independent of Signal) if the event would still pass the trigger with the signal removed (i.e. if the trigger requirement is met by another particle(s) in the event). It is also possible for candidates to fall into both categories.

The candidates remaining after applying the pre-selection have already passed the trigger requirements, in both data and simulation. This means the efficiency of the trigger requirements cannot be determined by comparing the number of events which pass the trigger to the total number initially produced. Instead, the TIS and TOS categories can be used to measure the trigger efficiency [55].

The efficiency of the trigger selection, $\varepsilon_{Trig}$, is given by

$$\varepsilon_{Trig} = \frac{N_{Trig}}{N_{Total}}, \tag{4.1}$$

where $N_{Trig}$ is the number of events passing the trigger requirement, and $N_{Total}$ is the total initial number of events. Note $N_{Total}$ is unknown because, as mentioned previously, only events passing the trigger requirement are retained and written to disk.

Equation (4.1) can be rewritten

$$\varepsilon_{Trig} = \frac{N_{Trig}}{N_{TIS}} . \frac{N_{TIS}}{N_{Total}} = \frac{N_{Trig}}{N_{TIS}} . \varepsilon_{TIS}, \tag{4.2}$$

where $N_{TIS}$ is the number of triggered events matching the TIS criteria, and $\varepsilon_{TIS}$ is the efficiency of selecting events classified as TIS. As the TIS criteria depends on the underlying event rather than the signal, it can be assumed that the efficiency of a TIS trigger signal is the same on any subset of events (essentially assuming the background composition is the same). This means $\varepsilon_{TIS}$ can be written as

$$\varepsilon_{Trig} = \frac{N_{Trig}}{N_{TIS}} . \frac{N_{TISTOS}}{N_{TOS}}, \tag{4.3}$$

where $N_{TOS}$ is the number of events matching the TOS criteria, and $N_{TISTOS}$ is the number of events meeting both the TIS and TOS criteria. All the quantities in Eq. (4.3) can be measured, allowing $\varepsilon_{Trig}$ to be determined despite $N_{Total}$ being unknown[2].

For each mode, the efficiencies of the trigger requirements were measured using Eq. (4.3), and are given in Table 4.3. The measurements were made after the pre-selections and the selections described later in this chapter were applied, except for the BDT output cut (see Subsection 4.2.6). For the $B_c^+ \to D_s^+ \mu^+ \mu^-$ and $B_c^+ \to D_s^+ J/\psi$ modes specifically, the PID cuts were also relaxed to provide an increase in statistics. This slightly-modified selection was applied to ensure the trigger efficiency measurement was being made on events which are signal-like while also leaving enough statistics to make an accurate measurement. Note the measurements described here refer to trigger efficiencies only; other efficiency measurements accounting for the detector acceptance and the selections themselves are given later, in Section 4.3.

[2] Note as the numerator in Eq. (4.3) is not a subset of the denominator, $\varepsilon_{Trig}$ values greater than 100% are possible. While this does not impact the analysis presented in this thesis in a significant way, a Bayesian analysis could be performed to ensure the estimates and uncertainties do not exceed the physical bound.

| Decay | Data Efficiency (%) | | Sim. Efficiency (%) | |
|---|---|---|---|---|
| | Run 2 | Run 1 | Run 2 | Run 1 |
| $B^0 \to \overline{D}^0\mu^+\mu^-$ | $83.5 \pm 3.1$ | $83.1 \pm 8.2$ | $91.5 \pm 1.2$ | $92.1 \pm 2.4$ |
| $B^0 \to \overline{D}^0 J/\psi$ | $92.9 \pm 5.8$ | $82.6 \pm 19.0$ | $95.1 \pm 0.8$ | $98.2 \pm 1.5$ |
| $B_s^0 \to \overline{D}^0\mu^+\mu^-$ | *same as* $B^0 \to \overline{D}^0\mu^+\mu^-$ | | $90.5 \pm 2.5$ | – |
| $B_s^0 \to \overline{D}^0 J/\psi$ | *same as* $B^0 \to \overline{D}^0 J/\psi$ | | $93.1 \pm 1.8$ | – |
| $B^0 \to J/\psi K^{*0}$ | $94.2 \pm 0.3$ | $94.3 \pm 0.6$ | $93.4 \pm 0.3$ | $95.4 \pm 0.5$ |
| $B^+ \to D_s^+\mu^+\mu^-$ | $83.3 \pm 4.5$ | $89.4 \pm 7.4$ | $90.9 \pm 1.5$ | $93.3 \pm 3.5$ |
| $B^+ \to D_s^+ J/\psi$ | $87.2 \pm 5.6$ | $97.2 \pm 6.2$ | $95.7 \pm 1.8$ | – |
| $B^+ \to J/\psi K^+$ | $94.2 \pm 0.2$ | $95.4 \pm 0.3$ | $93.3 \pm 0.2$ | $94.1 \pm 0.6$ |
| $B_c^+ \to D_s^+\mu^+\mu^-$ | $79.1 \pm 3.9$ | $84.1 \pm 6.6$ | $90.0 \pm 2.6$ | $86.6 \pm 7.2$ |
| $B_c^+ \to D_s^+ J/\psi$ | $101.4 \pm 2.5$ | $103.0 \pm 5.6$ | $93.1 \pm 1.9$ | $96.6 \pm 1.9$ |

**Table 4.3:** Efficiencies measured with the TISTOS method [55] in data and simulation for Runs 1 and 2. The $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$ yields in data are measured using the same selection and from the same sample of candidates as $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B^0 \to \overline{D}^0 J/\psi$, hence the efficiencies measured on data are also the same. A dash (–) indicates the relevant simulation sample was not available. Two of the measured efficiencies are larger than 100%. This is possible as the numerator in the efficiency calculation is not a subset of the denominator.

Differences in the trigger efficiencies between simulation and data are possible. The efficiency of the trigger is not individually measured in the analysis, and so to account for these differences, correction factors are derived which can be applied to the ratio of normalisation efficiency to signal efficiency in the final branching ratio calculation. This correction makes the trigger efficiency representative of data without requiring it to be measured as a separate value from the overall efficiency.

The above correction factors for each signal-mode/normalisation-mode pair are derived using

$$C_{trig} = \frac{\varepsilon_{\text{norm}}^{\text{trig,data}}}{\varepsilon_{\text{norm}}^{\text{trig,sim}}} \cdot \frac{\varepsilon_{\text{sig}}^{\text{trig,sim}}}{\varepsilon_{\text{sig}}^{\text{trig,data}}}\,, \tag{4.4}$$

where $C_{trig}$ is the correction factor and $\varepsilon_{\text{mode}}^{\text{trig,data/sim}}$ is the measured Run 2 trigger efficiency for the specified signal or normalisation mode on data/simulation. The derived correction factors are given in Table 4.4, and are applied when calculating the final results in Section 4.6. Using the Run 1 efficiencies provides consistent factors.

| Signal Mode | Normalisation Mode | Correction Factor |
|---|---|---|
| $B^0 \to \overline{D}^0 \mu^+\mu^-$ | $B^0 \to J/\psi\, K^{*0}$ | $1.10 \pm 0.04$ |
| $B^0 \to \overline{D}^0 J/\psi$ | $B^0 \to J/\psi\, K^{*0}$ | $1.03 \pm 0.07$ |
| $B_s^0 \to \overline{D}^0 \mu^+\mu^-$ | $B^0 \to J/\psi\, K^{*0}$ | $1.09 \pm 0.05$ |
| $B_s^0 \to \overline{D}^0 J/\psi$ | $B^0 \to J/\psi\, K^{*0}$ | $1.01 \pm 0.07$ |
| $B^+ \to D_s^+ \mu^+\mu^-$ | $B^+ \to J/\psi\, K^+$ | $1.10 \pm 0.06$ |
| $B^+ \to D_s^+ J/\psi$ | $B^+ \to J/\psi\, K^+$ | $1.11 \pm 0.07$ |
| $B_c^+ \to D_s^+ \mu^+\mu^-$ | $B_c^+ \to D_s^+ J/\psi$ | $1.24 \pm 0.08$ |
| $B_c^+ \to D_s^+ J/\psi$ | $B^+ \to J/\psi\, K^+$ | $0.93 \pm 0.03$ |

**Table 4.4:** The correction factors derived to account for differences between simulation and data trigger efficiencies for each signal and normalisation mode pair.

### 4.2.3 Separating Non-Resonant and $J/\psi \to \mu^+\mu^-$ Decays

To differentiate between decays proceeding non-resonantly and via $J/\psi \to \mu^+\mu^-$, appropriate cuts in $M(\mu^+\mu^-)$ or $q^2(\mu^+\mu^-)$ must be chosen, where $q^2(\mu^+\mu^-) = [M(\mu^+\mu^-)]^2$. Fig. 4.1 shows the $q^2(\mu^+\mu^-)$ distribution seen for $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B^0 \to \overline{D}^0 J/\psi$ candidates after the pre-selection. Based on this, a cut of $q^2(\mu^+\mu^-) \in [0.044, 8.0]$ GeV$^2/c^4$ was chosen to select non-resonant decays.

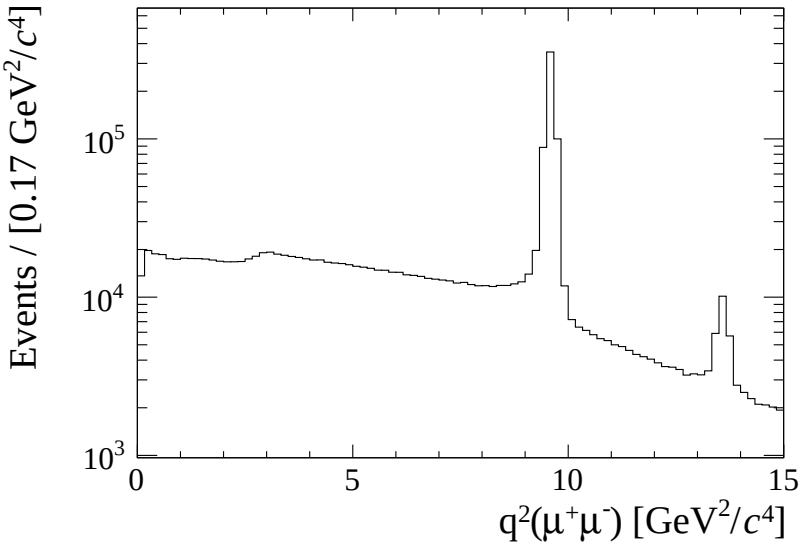



**Figure 4.1:** $q^2(\mu^+\mu^-)$ for $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B^0 \to \overline{D}^0 J/\psi$ candidates passing the pre-selection requirements. The significant peak seen at $q^2(\mu^+\mu^-) \sim 9.6\,\text{GeV}^2/c^4$ corresponds to the $J/\psi$ resonance, and the smaller peak at $q^2(\mu^+\mu^-) \sim 13.5\,\text{GeV}^2/c^4$ the $\psi(2S)$ resonance. The excess at $\sim 3\,\text{GeV}^2/c^4$ is attributed to $D^0 \to \pi^+\pi^-$ and $D^0 \to K^\pm\pi^\mp$ decays, with the hadrons misidentified as muons.

To select the samples of $J/\psi$ decays, a cut was made on the dimuon invariant

mass distribution. A $3\sigma$ cut around the $J/\psi$ mass was chosen, which was determined using $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B^0 \to \overline{D}^0 J/\psi$ candidates in data passing the pre-selection, shown in Fig. 4.2. This gives $M(\mu^+\mu^-) \in [3061, 3133]$ MeV/$c^2$ (taking $\sigma$ as the width of the inner, narrower Gaussian in the fit).

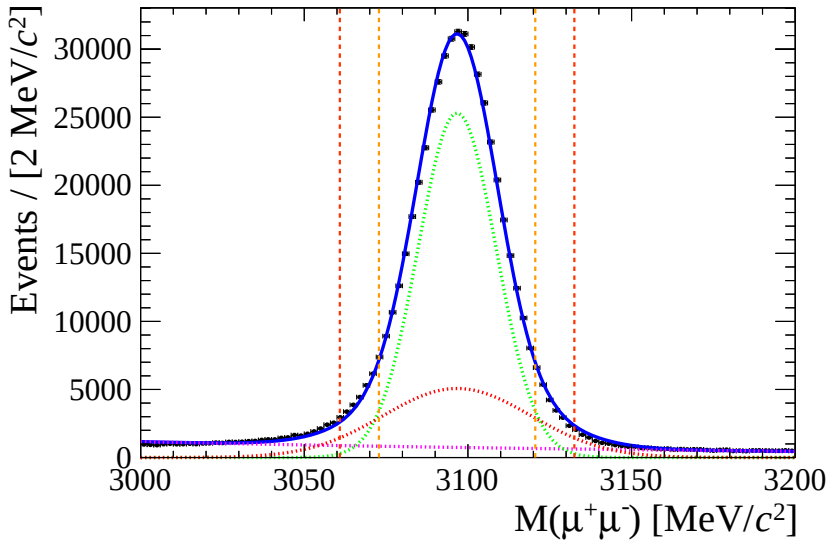



**Figure 4.2:** The dimuon invariant mass distribution around the $J/\psi$ mass for $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B^0 \to \overline{D}^0 J/\psi$ candidates in data passing the pre-selection. The distribution is fit with a sum of two Gaussian shapes to model the signal, shown individually by the green and red dotted lines. The background is modelled with an exponential shape, shown by the pink dotted line. The solid blue line shows the combined fit. The core (green) Gaussian has a mean of $3096.75 \pm 0.03$ MeV/$c^2$ and a width of $11.69 \pm 0.09$ MeV/$c^2$; using these values, the vertical orange dotted lines show the 2- and 3-sigma bounds of the core Gaussian. The 3-sigma bound is used to determine a cut of $M(\mu^+\mu^-) \in [3061, 3133]$ MeV/$c^2$ for selecting $J/\psi$ decays.

### 4.2.4 *D*-meson Selection

$D$-meson candidates were selected by making cuts on the invariant mass of the daughters, the vertex $\chi^2$/d.o.f, and the $D$-meson flight significance. The flight significance is defined as

$$\Sigma_D = \frac{z_B - z_D}{\sqrt{\sigma_{z_B}^2 + \sigma_{z_D}^2}}, \tag{4.5}$$

where $z_i$ is the z-coordinate of the decay vertex of particle $i$, and $\sigma_{Z_i}$ is the associated error. Low values of $\Sigma_D$ indicate the reconstructed $D$ meson did not travel a

significant distance, and thus is likely to have been background rather than a real $D$ meson.

Additionally, PID cuts were applied to the $D$-meson daughters, using likelihood-based PID variables. The $B^+ \to D_s^+ \mu^+ \mu^-$ and $B^+ \to D_s^+ J/\psi$ selection is a minor exception, where a PID cut is only applied to a single kaon rather than all final state hadrons, to match the normalisation mode ($B^+ \to J/\psi K^+$). When refering to a PID cut in this thesis, the first particle given is the one the cut is applied to, and the species following "PID" is what the likelihood refers to. The typical kaon identification efficiency in LHCb is $\sim$ 95%, with a $\sim$ 5% pion to kaon misidentification probability.

A summary of the $D$ meson selection criteria is given in Table 4.5, and the resulting samples of $D^0$ and $D_s^+$ candidates are shown in Figs 4.3 and 4.4 (with the $D$-mass cuts shown graphically before they have been applied).

| Variable | $D^0$ Cut | $D_s^+$ Cut |
|---|---|---|
| $D$ Mass | $\pm 25\,\mathrm{MeV}/c^2$ | $\pm 20\,\mathrm{MeV}/c^2$ |
| $D$ Flight Significance | $> 2$ | $> 2$ |
| $D$ Vertex $\chi^2$/d.o.f | $< 2$ | $< 2$ |
| $K$ PID kaon | $> 1$ | $> 1$ * |
| $\pi$ PID kaon | $< -1$ | $< -1$ † |

**Table 4.5:** The cuts used to select $D^0$ and $D_s^+$ mesons in the $\overline{D}^0 \to K^+\pi^-$ and $D_s^+ \to K^+K^-\pi^+$ channels. The asterisk (*) indicates this cut is only applied to a single kaon in the selection used for $B^+ \to D_s^+ \mu^+ \mu^-$ and $B^+ \to D_s^+ J/\psi$ decays. The dagger (†) indicates this cut is only applied in the $B_c^+ \to D_s^+ \mu^+ \mu^-$ and $B_c^+ \to D_s^+ J/\psi$ selection.

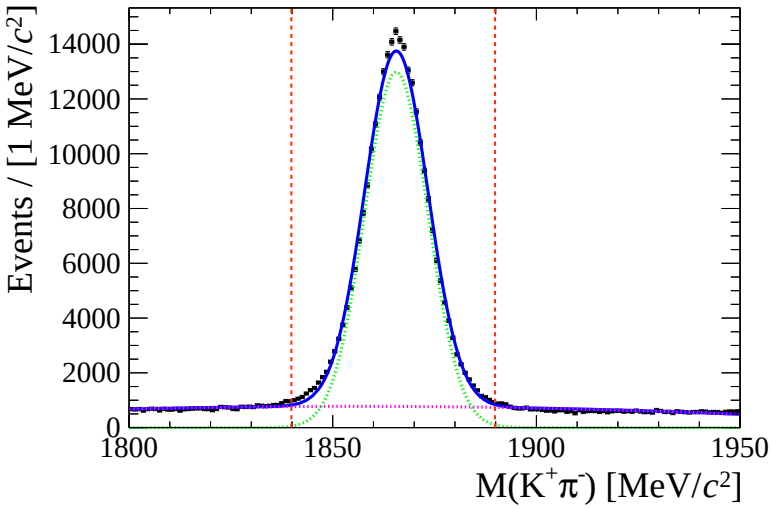



**Figure 4.3:** The $K\pi$ invariant mass for $D^0$ candidates used in the $B^0 \to \overline{D}^0\mu^+\mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0\mu^+\mu^-$, and $B_s^0 \to \overline{D}^0 J/\psi$ selections after the cuts in Table 4.5 have been applied, with the exception of the $D$ invariant mass cut which is indicated by the vertical dotted lines. A mass window has also been applied to the reconstructed $B$ mass, selecting only events with a $B$ mass within the range $[5000, 5800]\,\text{MeV}/c^2$. A Gaussian shape is used to model the $D$-meson peak (green dotted line) and a second-order Chebychev polynomial for the background (pink dotted line). The mean and sigma of the Gaussian are $1865.61 \pm 0.02$ MeV/$c^2$ and $7.76 \pm 0.02$ MeV/$c^2$, respectively.

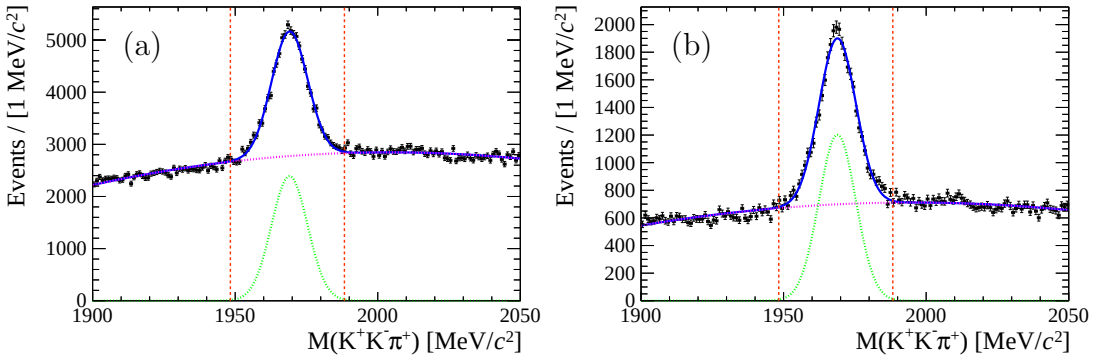



**Figure 4.4:** The $KK\pi$ invariant mass for $D_s^+$ candidates used in (a) the $B^+ \to D_s^+\mu^+\mu^-$ and $B^+ \to D_s^+ J/\psi$ and (b) the $B_c^+ \to D_s^+\mu^+\mu^-$ and $B_c^+ \to D_s^+ J/\psi$ selections after the cuts in Table 4.5 have been applied, with the exception of the $D$ invariant mass cut which is indicated by the vertical dotted lines. A mass window has also been applied to the reconstructed $B$ mass, selecting only events with a $B$ mass within the range (a) $[5100, 5800]\,\text{MeV}/c^2$ and (b) $[5600, 6700]\,\text{MeV}/c^2$. The larger background level observed in (a) is due to the lack of PID cuts on the pion and second kaon compared to (b). Both plots are fit with a Gaussian shape to model the $D_s^+$-meson peak (green dotted line) and a second order Chebychev polynomial for the background (pink dotted line). The means and sigmas of the Gaussians are (a) $1969.06 \pm 0.07$ MeV/$c^2$ and $6.34 \pm 0.08$ MeV/$c^2$ and (b) $1968.88 \pm 0.08$ MeV/$c^2$ and $6.46 \pm 0.09$ MeV/$c^2$ respectively.

### 4.2.5 *B*-meson Selection

With the $D$-meson samples identified, $B$-meson candidates are constructed by adding two muons. In addition to the $M(\mu^+\mu^-)$ and $q^2(\mu^+\mu^-)$ cuts detailed previously, a PID cut is applied to both muons, using similar likelihood-based PID variables as for the hadrons. Additionally, for all modes except $B_c^+ \to D_s^+\mu^+\mu^-$ and $B_c^+ \to D_s^+ J/\psi$, the $B$ lifetime is required to be $> 0.2\,\text{ps}$, superceeding the cut applied in the pre-selection.

### 4.2.6 Multivariate Analysis to Reduce Background

After the previously-detailed requirements have been made, a Boosted Decision Tree (BDT) is trained and applied for each family of modes based on the primary $B$-meson species. Three BDTs are used:

- a "$B^0$ BDT", for $B^0 \to \overline{D}^0\mu^+\mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0\mu^+\mu^-$, and $B_s^0 \to \overline{D}^0 J/\psi$ decays,
- a "$B^+$ BDT", for $B^+ \to D_s^+\mu^+\mu^-$ and $B^+ \to D_s^+ J/\psi$ decays,
- a "$B_c^+$ BDT", for $B_c^+ \to D_s^+\mu^+\mu^-$ and $B_c^+ \to D_s^+ J/\psi$ decays.

These BDTs are designed to classify events as signal or background based on the properties of the decay in a more optimal way than using simple "square" cuts.

To create the BDTs, a list of input variables is provided, as well as two data samples: a "signal" sample, which contains signal-like events, and a "background" sample, which contains background or background-like events. The samples are merged, but retain their signal or background label. The BDT uses the merged sample to optimise a set of decisions on the input variables to provide the best separation between events labelled signal and background, using the following procedure:

- find a cut value for each input variable which maximises the separation between signal and background;

- select the cut which provides the best separation;
- split the sample into two subsamples, those which pass the cut and those which do not;
- repeat the procedure independently on the two subsamples, iterating until a stopping criteria is met (usually a maximum number of decisions).

This procedure is referred to as "training". Once the decision tree has been created, the output for a given event is the purity of the final subsample in which it is contained,

$$purity = \frac{s}{s+b}, \tag{4.6}$$

where $s/b$ are the number of events labelled signal/background in the subsample.

The above algorithm describes the "Decision Tree" part of the BDTs. The "Boosted" in BDT refers to a procedure where multiple decision trees are created from the same sample by applying arbitrary weights. After successive reweightings, the ensemble of decision trees is combined into a single output by taking the weighted average of the outputs of the individual trees. This procedure boosts performance and avoids the case with a single decision tree where the output can be dominated by specific features of small subsets of events.

The BDTs used in this thesis had a maximum depth of 3 (i.e. the procedure outlined above was repeated 3 times), and the end result is a weighted average of 850 individual trees.

**Signal and Background Samples**

For all modes, the signal samples used in training were simulated samples of the respective non-resonant signal mode, with truth-level information used to ensure only signal decays were selected. For the $B^0$ BDT specifically, only simulated $B^0 \to \overline{D}^0\mu^+\mu^-$ decays were used in the training.

All years of data were used in forming the background samples. For the $B^+$ and $B_c^+$ BDTs, unphysical reconstructed decays of the form $B_c^+ \to D_s^-\mu^+\mu^+$ were

used. For the $B^0$ BDT, events in the upper-mass side band in data, defined with a reconstructed invariant $B^0$ mass greater than $5800\,\mathrm{MeV}/c^2$, were used.

For both the signal and background samples, the $D$-meson selection and additional $B$-meson cuts were applied to the samples, except with the PID requirements relaxed to those given in Table 4.6 to boost the statistics of the samples.

| Variable | Cut |
|---|---|
| $K$ PID kaon | $> 0.1$ |
| $\pi$ PID kaon | $< -0.1$ |
| $\mu$ PID muon | $> 0.1$ |

**Table 4.6:** The PID cuts used in the selection of the BDT signal and background samples.

### Variable Lists

The same variables were used for all three BDTs, chosen primarily to be properties of the $B$ meson or properties of the $D$ meson related to its production from the $B$ meson. An additional pair of $D$-specific variables were included to further clean the $D$-meson sample, and for the $B^+$ and $B_c^+$ BDTs, the transverse momenta of the kaons from the $D_s^+$ decay were also included. The complete list of variables is given in Table 4.7.

| Variable |
|---|
| **$B$-meson properties** |
| • Angle between the $B$-meson vertices and the $B$ reconstructed momentum |
| • $B$-meson transverse momentum |
| • $B$-meson pseudorapidity |
| • $\chi^2$ of the $B$-meson end vertex |
| • Maximum DOCA between the $B$ meson and other particles |
| • $B$-meson DOCA to the PV |
| • Distance in the $y$-$z$-plane from the $B$-meson end vertex to the PV |
| • $B$-meson lifetime |
| **$D$-meson properties related to the $B$-meson** |
| • $D$-meson transverse momentum |
| • $D$-meson pseudorapidity |
| • Distance from the $D$-meson end vertex to the PV |
| • $D$-meson DOCA to the PV |
| • Distance in the $y$-$z$-plane from the $D$-meson end vertex to the PV |
| **Additional $D$-meson-related variables** |
| • Angle between the $D$-meson vertices and the $D$ reconstructed momentum |
| • Maximum DOCA between the $D$ meson and other particles |
| • * Transverse momenta of the $D_s^+$-meson daughter kaons |

**Table 4.7:** Variables used in training the BDTs. The asterisk (*) indicates the variable is only used in the $B^+$ and $B_c^+$ BDTs, where the decay includes a $D_s^+$ meson. DOCA stands for Distance Of Closest Approach.

### BDT Performance

The behaviour of the BDTs was validated through the use of testing samples, produced in the same way as the training samples but with independent events. By testing on independent samples, it can be checked that the BDT has learned to distinguish events based on the properties of the signal decay and not any specific properties of the training samples.

The *Area Under the ROC*[3] *Curve* (AUC) score was used to quantitatively benchmark each BDT. The AUC score is a normalised quantity, hence a value of 1.0 corresponds to perfect separation. Table 4.8 gives the AUC score achieved by each BDT. Measured on the independent testing samples, it can be seen that each BDT has excellent separating power and so will provide strong background rejection

[3]ROC stands for Receiver Operating Characteristic.

when used in the selection. The ROC curves for each BDT, as well as comparisons of the BDT response on the training and testing samples, are shown in Figs 4.5, 4.6, and 4.7. The similarity in the distributions of BDT response between the testing and training samples indicates the BDTs are able to correctly distinguish signal and have not learned any specific properties of the training samples.

| Mode | AUC |
|---|---|
| $B^0$ BDT | 0.988 |
| $B^+$ BDT | 0.987 |
| $B_c^+$ BDT | 0.978 |

**Table 4.8:** The Area Under Curve (AUC) score for each BDT.

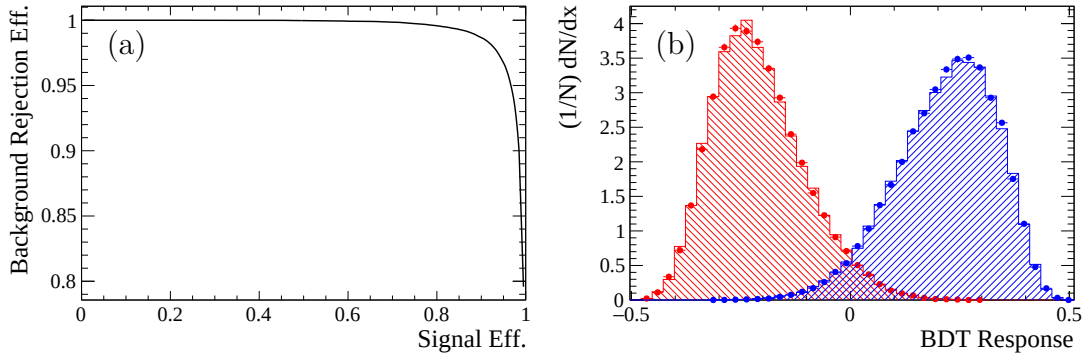



**Figure 4.5:** (a) The ROC curve for the $B^0$ BDT. (b) The BDT response for the signal (blue) and background (red) training (points) and testing (bars) datasets used in optimising the $B^0$ BDT.

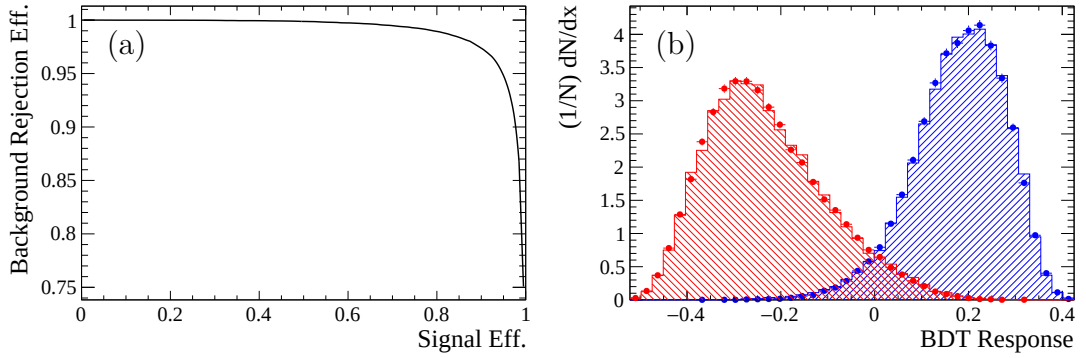



**Figure 4.6:** (a) The ROC curve for the $B^+$ BDT. (b) The BDT response for the signal (blue) and background (red) training (points) and testing (bars) datasets used in optimising the $B^+$ BDT.

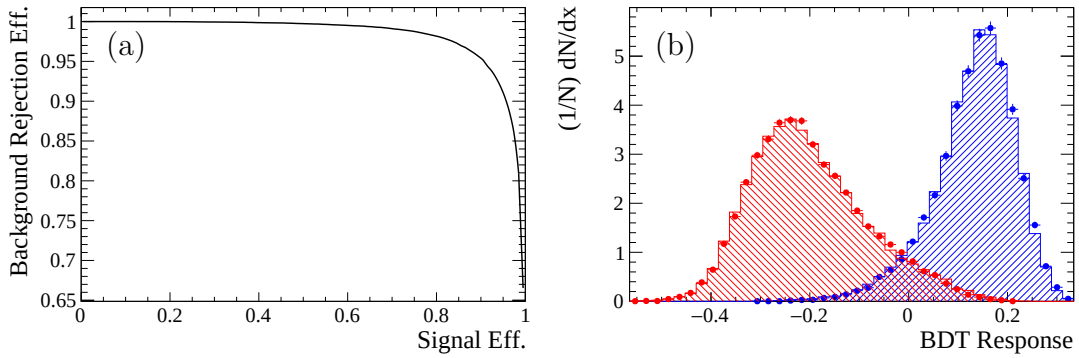



**Figure 4.7:** (a) The ROC curve for the $B_c^+$ BDT. (b) The BDT response for the signal (blue) and background (red) training (points) and testing (bars) datasets used in optimising the $B_c^+$ BDT.

### BDT Output Cut Optimisation

To obtain the optimal trade-off between reducing background and selecting signal decays, a figure of merit was maximised with respect to the cut on the BDT output. The Punzi figure of merit [56] was used, defined as

$$\mathcal{F}_p = \frac{\varepsilon_S}{\sqrt{B} + \frac{a}{2}}, \tag{4.7}$$

where $\varepsilon_S$ is the signal efficiency, $B$ is the number of background events ($\sqrt{B}$ its uncertainty) and $a$ is the statistical significance in standard deviations of the expected signal (introduced in the derivation by considering a one-sided Gaussian test at a particular significance, and then maximising the search sensitivity at the threshold of signal visibility). Somewhat colloquially, the Punzi figure of merit aims to optimise for searches where a signal may not be observed, and hence the signal efficiency is of interest rather than the signal yield. A constant term is desired in the denominator to prevent the figure of merit breaking down at very small values of $B$. The value $a = 3$ was chosen for the optimisation procedure as this was seen to provide a stronger background rejection than the traditional $a = 5$, and thus provide a better optimisation for setting upper limits.

An example of the variation in $\varepsilon_S$, $B$, and $\mathcal{F}_p$ with BDT output cut is shown in Fig. 4.8 for the $B^+ \to D_s^+ J/\psi$ selection. $\varepsilon_S$ was measured using simulation, and $B$ measured in data with a fit used to extract the background component. The

optimal cut for each mode is taken as the value which maximises $\mathcal{F}_p$, resulting in the cut values presented in Table 4.9.

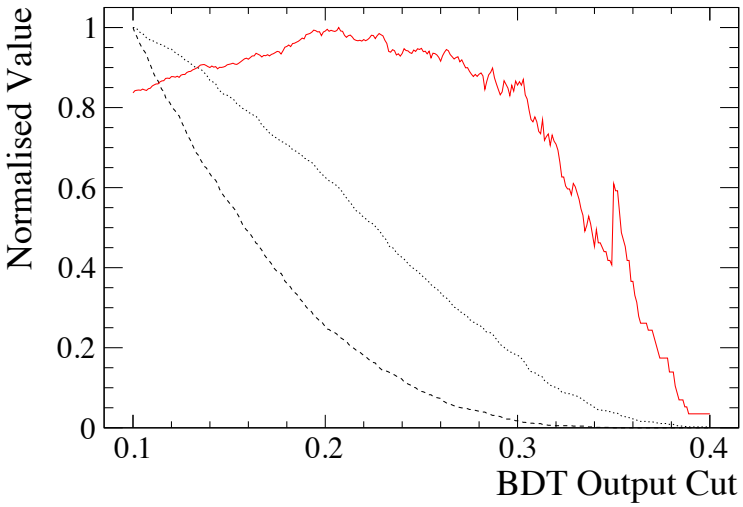



**Figure 4.8:** The variation in the signal efficiency (dotted), background level beneath the signal peak (dashed), and Punzi Figure of Merit (red, solid) as a function of BDT output cut, used to optimise the $B^+$ BDT output cut to select $B^+ \to D_s^+ J/\psi$ decays. All $y$-axis values have been scaled so the maximum point of each curve occurs at $y = 1.0$.

| Mode | BDT | Cut Value |
|---|---|---|
| $B^0 \to \overline{D}^0 \mu^+ \mu^-$ and $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | $B^0$ | 0.272 |
| $B^0 \to \overline{D}^0 J/\psi$ and $B_s^0 \to \overline{D}^0 J/\psi$ | $B^0$ | 0.301 |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | $B^+$ | 0.223 |
| $B^+ \to D_s^+ J/\psi$ | $B^+$ | 0.207 |
| $B_c^+ \to D_s^+ \mu^+ \mu^-$ | $B_c^+$ | 0.156 |
| $B_c^+ \to D_s^+ J/\psi$ | $B_c^+$ | 0.123 |

**Table 4.9:** The BDT output cut values chosen for each decay mode through maximising the Punzi figure of merit. As $B^0 \to \overline{D}^0 \mu^+ \mu^-$ and $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ candidates undergo the same selection, the same optimised cut is used for both. The same is true for $B^0 \to \overline{D}^0 J/\psi$ and $B_s^0 \to \overline{D}^0 J/\psi$ candidates.

### BDT Validation

To ensure the BDTs are able to rely on the simulation training samples and correctly distinguish signal, the $B^0$ BDT was validated using both $B^0 \to J/\psi K^{*0}$ data and simulation by substituting the $D^0$ variables with the equivalent variables for the $K^{*0}$. After applying the simple, cuts-based selection presented in Subsection 4.2.10

to select a $B^0 \to J/\psi K^{*0}$ data sample, the data were $_s$Weighted [57] to give a representative sample of $B^0 \to J/\psi K^{*0}$ signal decays in data which could be compared to simulation. Qualitatively, $_s$Weighting involves applying a weight to each event derived from a fit with defined signal and background shapes. Events where the signal component is significant receive a large weight, while events in regions dominated by the background component receive a small weight. This allows data, where the two components cannot be separated, to be weighted such that the signal is better represented.

The variation in BDT output between $B^0 \to J/\psi K^{*0}$ simulation and $_s$Weighted data is shown in Fig. 4.9. Good agreement is seen, especially in the upper region where the majority of the optimal BDT output cuts are. To further investigate the agreement between simulation and data, the efficiency was investigated as a function of BDT output cut. The number of retained events compared to the total available is shown in Fig. 4.10 as a function of BDT cut for $B^0 \to J/\psi K^{*0}$ simulation and $_s$Weighted $B^0 \to J/\psi K^{*0}$ data. A systematic for the difference in efficiency between simulation and data was derived from this by considering the ratio of the two efficiencies at a specific BDT cut (see Subsection 4.5.4).

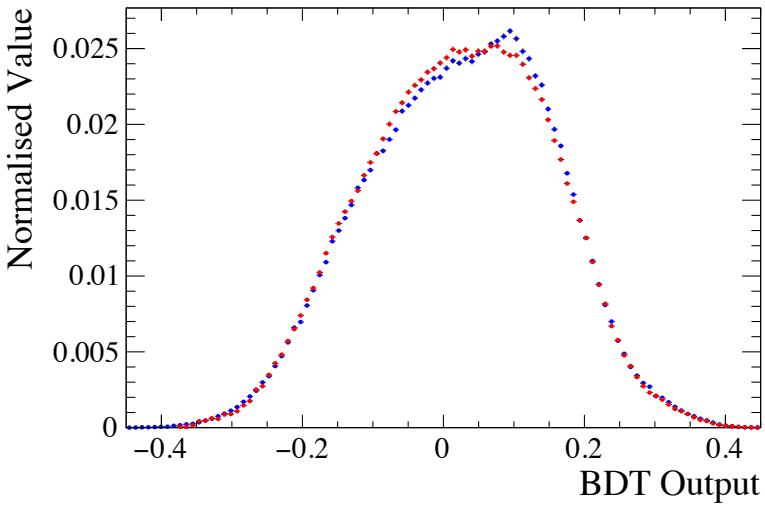



**Figure 4.9:** Comparison of the BDT output distribution when the $B^0$ BDT is applied to $B^0 \to J/\psi K^{*0}$ simulation (blue) and $_s$Weighted data (red).

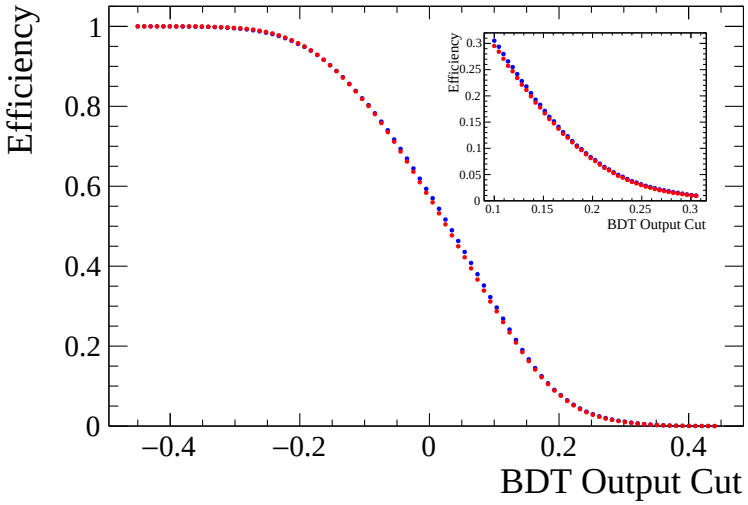



**Figure 4.10:** Comparison of the efficiency in $B^0 \to J/\psi K^{*0}$ simulation (blue) and ${}_s$Weighted data (red) as a function of BDT output cut, where the efficiency is defined as the number of events passing the cut divided by the total events available in the initial sample.

### 4.2.7 $B^0 \to D^{*-}\mu^+\nu_\mu$ Semileptonic Background

The $B^0 \to \left(D^{*-} \to \overline{D}{}^0\pi^-\right)\mu^+\nu_\mu$ decay, where the $\nu$ is missed and the $\pi^-$ is misidentified as a $\mu^-$, appears as a background at low $q^2\,(\mu^+\mu^-)$ in the $B^0 \to \overline{D}{}^0\mu^+\mu^-$ search. By taking a $\pi^-$ mass hypothesis for the $\mu^-$, the $D^{*+}$ mass can be found, and a mass difference, $\Delta M = M\,(D^{*+}) - M\,(D^0)$, calculated. $\Delta M$ is expected to peak just above the $\pi^-$ mass.

Figure 4.11 shows the $\Delta M$ distribution for $B^0 \to \overline{D}{}^0\mu^+\mu^-$ candidates in data, with the full selection previously described applied except for the $\Delta M$ cut, with both charge combinations considered. A cut of width $10\,\mathrm{MeV}/c^2$ centred around $146\,\mathrm{MeV}/c^2$ has been chosen to remove this background, and is shown applied to the data in Fig. 4.12.

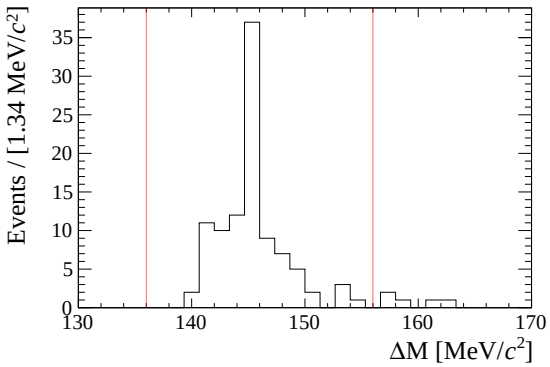



**Figure 4.11:** The $\Delta M = M\left(D^{*+}\right) - M\left(D^0\right)$ distribution in data, after applying the selection given in Table 4.10 except for the $\Delta M$ cut, with both charge combinations considered. The red lines show the chosen cut, of width $10\,\text{MeV}/c^2$ centred around $146\,\text{MeV}/c^2$.

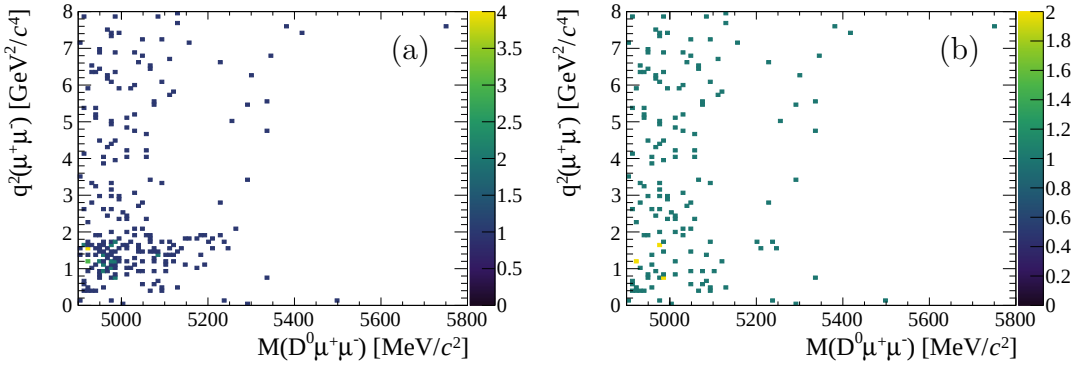



**Figure 4.12:** $q^2\left(\mu^+\mu^-\right)$ as a function of the $B^0$ invariant mass in the $B^0 \to \overline{D}^0\mu^+\mu^-$ data sample after applying the final selection criteria described in this section: (a) except for the cut related to the $B^0 \to D^{*-}\mu^+\nu_\mu$ background and (b) applying all the cuts. The excess at low $q^2\left(\mu^+\mu^-\right)$ and $B^0$ mass is seen to be removed by the cut.

### 4.2.8 Final Search Mode Selections

In summary, the final selections for each search mode are given in Tables 4.10, 4.11 and 4.12.

| Variable | Requirement |
|---|---|
| **$D^0$ Selection** | |
| $D^0$ Mass | $\lvert M(D^0) - M_{PDG}(D^0)\rvert < 25\,\text{MeV}/c^2$ |
| $D^0$ Flight Significance | $> 2$ * |
| $D^0$ Vertex $\chi^2$/d.o.f | $< 2$ |
| $K$ PID kaon | $> 1$ |
| $\pi$ PID kaon | $< -1$ |
| **$B^0$ Selection** | |
| $B^0$ Mass | $\in [5000, 5800]\,\text{MeV}/c^2$ |
| $\mu$ PID muon | $> 4$ |
| $B^0$ Lifetime | $> 0.0002\,\text{ns}$ |
| $\Delta M = M\,(D^{*+}) - M\,(D^0)$ | $\ni [136, 156]\,\text{MeV}/c^2$ |
| **$B^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0\mu^+\mu^-$ Specific** | |
| $q^2\,(\mu^+\mu^-)$ | $\in [0.25, 8.0]\ \text{GeV}^2/c^4$ |
| BDT Output | $> 0.272$ |
| **$B^0 \to \overline{D}^0 J/\psi$ and $B_s^0 \to \overline{D}^0 J/\psi$ Specific** | |
| $M\,(\mu^+\mu^-)$ | $\in [3061, 3133]\ \text{MeV}/c^2$ |
| BDT Output | $> 0.301$ |
| $D^0$ Flight Significance | $> 3$ |

**Table 4.10:** Final requirements for the $B^0 \to \overline{D}^0\mu^+\mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0\mu^+\mu^-$, and $B^0 \to \overline{D}^0 J/\psi$ selections. $M_{PDG}$ indicates the mass as given in Ref. [28]. The asterisk (*) is to highlight that the flight significance cut of $> 2$ is superseded by a tighter cut of $> 3$ for the $B^0 \to \overline{D}^0 J/\psi$ and $B_s^0 \to \overline{D}^0 J/\psi$ decay modes.

| Variable | Requirement |
|---|---|
| **$D_s^+$ Selection** | |
| $D_s^+$ Mass | $\lvert M(D_s^+) - M_{PDG}(D_s^+)\rvert < 20\,\text{MeV}/c^2$ |
| $D_s^+$ Flight Significance | $> 2$ * |
| $D_s^+$ Vertex $\chi^2$/d.o.f | $< 2$ |
| $K$ PID kaon | $> 1$ (only applied to one kaon) |
| **$B^+$ Selection** | |
| $B^+$ Mass | $\in [5100, 5800]\,\text{MeV}/c^2$ |
| $\mu$ PID muon | $> 4$ |
| $B^+$ Lifetime | $> 0.0002\,\text{ns}$ |
| **$B^+ \to D_s^+\mu^+\mu^-$ Specific** | |
| $q^2\,(\mu^+\mu^-)$ | $\in [0.25, 8.0]\ \text{GeV}^2/c^4$ |
| BDT Output | $> 0.223$ |
| **$B^+ \to D_s^+ J/\psi$ Specific** | |
| $M\,(\mu^+\mu^-)$ | $\in [3061, 3133]\ \text{MeV}/c^2$ |
| BDT Output | $> 0.250$ |
| $D_s^+$ Flight Significance | $> 3$ |

**Table 4.11:** Final requirements for the $B^+ \to D_s^+\mu^+\mu^-$ and $B^+ \to D_s^+ J/\psi$ selections. $M_{PDG}$ indicates the mass as given in Ref. [28]. The asterisk (*) is to highlight that the flight significance cut of $> 2$ is superseded by a tighter cut of $> 3$ for the $B^+ \to D_s^+ J/\psi$ decay mode.

| Variable | Requirement |
|---|---|
| **$D_s^+$ Selection** | |
| $D_s^+$ Mass | $\lvert M(D_s^+) - M_{PDG}(D_s^+)\rvert < 20\,\mathrm{MeV}/c^2$ |
| $D_s^+$ Flight Significance | $> 2$ |
| $D_s^+$ Vertex $\chi^2$/d.o.f | $< 2$ |
| $K$ PID kaon | $> 1$ |
| $\pi$ PID kaon | $< -1$ |
| **$B_c^+$ Selection** | |
| $B_c^+$ Mass | $\in [5600, 6700]\,\mathrm{MeV}/c^2$ |
| $\mu$ PID muon | $> 4$ |
| **$B_c^+ \to D_s^+\mu^+\mu^-$ Specific** | |
| $q^2\,(\mu^+\mu^-)$ | $\in [0.25, 8.0]\ \mathrm{GeV}^2/c^4$ |
| BDT Output | $> 0.156$ |
| **$B_c^+ \to D_s^+ J/\psi$ Specific** | |
| $M\,(\mu^+\mu^-)$ | $\in [3061, 3133]\ \mathrm{MeV}/c^2$ |
| BDT Output | $> 0.123$ |

**Table 4.12:** Final requirements for the $B_c^+ \to D_s^+\mu^+\mu^-$ and $B_c^+ \to D_s^+ J/\psi$ selections. $M_{PDG}$ indicates the mass as given in Ref. [28].

### 4.2.9 Backgrounds Without a $D$ Meson

After the selection is applied, it is possible for decays with the same final state, but without the hadrons being produced by a $D$ meson, to appear. This can be investigated by searching for structure in the $B$-meson invariant mass distribution in the lower and upper sidebands of the $D$-meson invariant mass distribution. After applying the selections given in the previous section, residual peaking structure is only seen for $B^0 \to \overline{D}^0 J/\psi$ decays, as shown in Fig. 4.13, attributed to $B^0 \to K^+\pi^- J/\psi$ decays. This contribution will be accounted for in the $B^0$ invariant mass fit, described in Subsection 4.4.1.

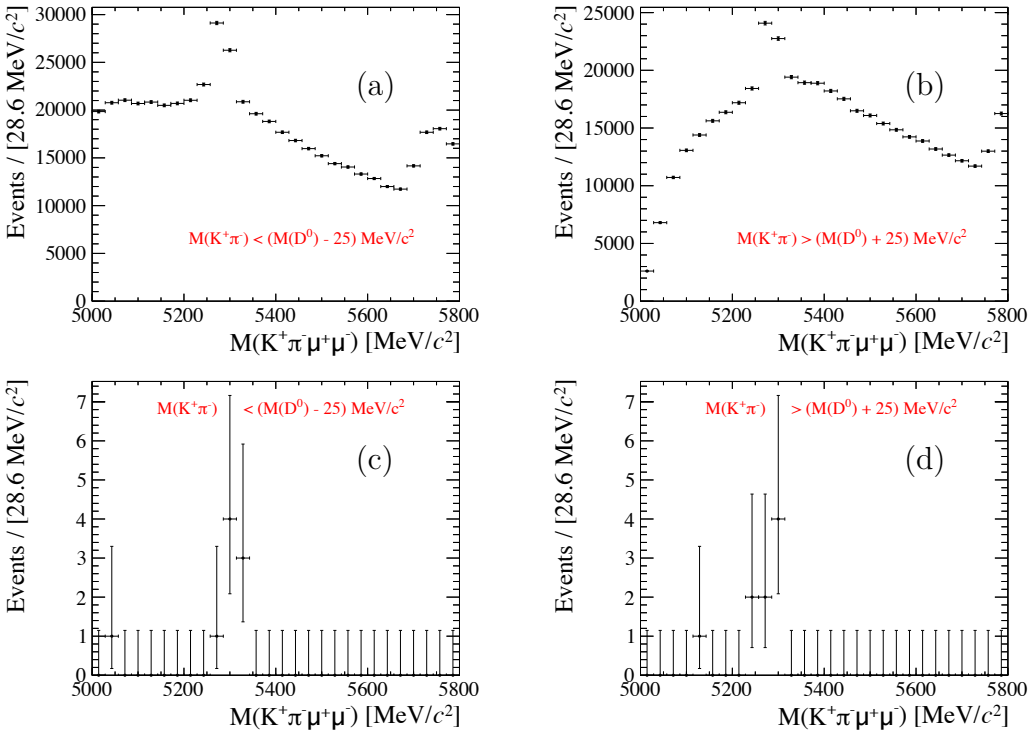



**Figure 4.13:** The unconstrained (i.e. no kinematic fit constraints) $B^0$-invariant-mass spectra for $B^0 \to \overline{D}^0 J/\psi$ candidates in the (a, c) lower and (b, d) upper sidebands of the $D^0$ invariant mass distribution, defined as being outside the mass cut given in Table 4.5 and shown by the vertical lines in Fig. 4.3a. The spectra are shown in (a, b) with only the pre-selection, trigger, and $M(\mu^+\mu^-)$ cuts applied and in (c, d) with the selection given in Table 4.5 (except for the $D^0$ mass cut) applied. Some residual structure remains after the selections have been applied, attributed to $B^0 \to K^+\pi^- J/\psi$ decays.

### 4.2.10 Normalisation-Mode Selections

Simple cut-based selections, given in Tables 4.13 and 4.14, were used to create samples of $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ decays from the pre-selected data. For the $B^0 \to (J/\psi \to \mu^+\mu^-)(K^{*0} \to K^+\pi^-)$ selection, an additional PID requirement was imposed on the $\pi$ to remove a possible background resulting from $\Lambda_b^0 \to J/\psi pK^-$ decay, with the proton misidentified as a pion. This cut-based approach was seen to provide a sufficiently clean sample for the purposes of normalising the signal decays, and so a multivariate technique was not employed in this case.

| Variable | Requirement |
|---|---|
| $B^0$ Mass | $\in [5175, 5350]\,\text{MeV}/c^2$ |
| $K^*$ Mass | $\lvert M(K\pi) - M_{PDG}(K^*)\rvert < 50\,\text{MeV}/c^2$ |
| $K$ PID kaon | $> 1$ |
| $\pi$ PID kaon | $< -1$ |
| $\pi$ PID proton | $< -4$ |
| $\mu$ PID muon | $> 4$ |
| $M(\mu^+\mu^-)$ | $\in [3061, 3133]$ MeV/$c^2$ |

**Table 4.13:** Requirements for the $B^0 \to J/\psi K^{*0}$ selection. $M_{PDG}(K^*)$ is the mass of the $K^*$ meson taken from Ref. [28].

| Variable | Requirement |
|---|---|
| $B^+$ Mass | $\in [5175, 5425]\,\text{MeV}/c^2$ |
| $K$ PID kaon | $> 1$ |
| $\mu$ PID muon | $> 4$ |
| $M(\mu^+\mu^-)$ | $\in [3061, 3133]$ MeV/$c^2$ |

**Table 4.14:** Requirements for the $B^+ \to J/\psi K^+$ selection.

## 4.3 Efficiencies

The efficiencies for each mode are measured using simulation by comparing the number of events retained at each different stage of the analysis. The efficiencies for Run 1 and Run 2 are measured separately and combined together with a weighted average.

The efficiency in Run 2 was determined individually for all modes. For Run 1, an independent efficiency measurement is made with Run 1 conditions wherever the simulation sample is available.

The efficiency measurements are broken down into three parts:

- the **generator efficiency** is the efficiency of decays to be within the LHCb acceptance;
- the **online efficiency** is the efficiency of the reconstruction and pre-selection;
- the **offline efficiency** is the efficiency of the selection presented in Section 4.2.

These three efficiencies are given for each mode, individually and combined, in Table 4.15 for Run 2 and Table 4.16 for Run 1.

The trigger efficiency is separated between online and offline; the efficiency of the overall trigger to record an event is included in the online efficiency, while the efficiency of the specific subset of trigger requirements made in this analysis, as detailed in Subsection 4.2.2, are included in the offline efficiency. For this reason, trigger efficiency corrections for the differences between data and simulation derived in Subsection 4.2.2 are applied later to the ratio of efficiencies in the branching ratio calculation, and so are not included in the values given in Tables 4.15 and 4.16.

| Decay Mode | Efficiency (%) | | | |
|---|---|---|---|---|
| | Generator | Online | Offline | Total |
| $B^0 \to \overline{D}^0\mu^+\mu^-$ | $16.39 \pm 0.05$ | $11.87 \pm 0.02$ | $6.72 \pm 0.05$ | $0.131 \pm 0.001$ |
| $B^0 \to \overline{D}^0 J/\psi$ | $16.57 \pm 0.05$ | $13.83 \pm 0.02$ | $5.43 \pm 0.05$ | $0.124 \pm 0.001$ |
| $B_s^0 \to \overline{D}^0\mu^+\mu^-$ | $16.44 \pm 0.05$ | $12.47 \pm 0.04$ | $6.78 \pm 0.10$ | $0.139 \pm 0.002$ |
| $B_s^0 \to \overline{D}^0 J/\psi$ | $16.74 \pm 0.12$ | $11.35 \pm 0.04$ | $5.59 \pm 0.10$ | $0.106 \pm 0.002$ |
| $B^+ \to D_s^+\mu^+\mu^-$ | $16.35 \pm 0.05$ | $8.30 \pm 0.02$ | $6.52 \pm 0.06$ | $0.089 \pm 0.001$ |
| $B^+ \to D_s^+ J/\psi$ | $16.56 \pm 0.12$ | $8.61 \pm 0.04$ | $10.84 \pm 0.16$ | $0.135 \pm 0.003$ |
| $B_c^+ \to D_s^+\mu^+\mu^-$ | $14.8 \pm 0.1$ | $5.02 \pm 0.02$ | $3.72 \pm 0.06$ | $0.0277 \pm 0.0005$ |
| $B_c^+ \to D_s^+ J/\psi$ | $14.48 \pm 0.05$ | $5.66 \pm 0.02$ | $9.97 \pm 0.13$ | $0.082 \pm 0.001$ |
| $B^0 \to J/\psi K^{*0}$ | $16.77 \pm 0.04$ | $15.30 \pm 0.01$ | $27.49 \pm 0.04$ | $0.705 \pm 0.002$ |
| $B^+ \to J/\psi K^+$ | $17.40 \pm 0.06$ | $31.84 \pm 0.01$ | $37.18 \pm 0.03$ | $2.060 \pm 0.008$ |

**Table 4.15:** Summary of the efficiencies for each mode measured with simulation corresponding to Run 2 conditions.

| Decay Mode | Efficiency (%) | | | |
|---|---|---|---|---|
| | Generator | Online | Offline | Total |
| $B^0 \to \overline{D}^0 \mu^+ \mu^-$ | $15.34 \pm 0.12$ | $9.93 \pm 0.03$ | $5.55 \pm 0.08$ | $0.085 \pm 0.001$ |
| $B^0 \to \overline{D}^0 J/\psi$ | $15.60 \pm 0.12$ | $11.66 \pm 0.04$ | $4.31 \pm 0.08$ | $0.078 \pm 0.002$ |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | $15.45 \pm 0.12$ | $7.10 \pm 0.03$ | $5.49 \pm 0.12$ | $0.060 \pm 0.001$ |
| $B_c^+ \to D_s^+ \mu^+ \mu^-$ | $13.07 \pm 0.07$ | $4.73 \pm 0.03$ | $2.37 \pm 0.12$ | $0.015 \pm 0.001$ |
| $B_c^+ \to D_s^+ J/\psi$ | $12.73 \pm 0.07$ | $6.28 \pm 0.02$ | $5.49 \pm 0.10$ | $0.044 \pm 0.001$ |
| $B^0 \to J/\psi K^{*0}$ | $15.75 \pm 0.03$ | $17.36 \pm 0.02$ | $15.76 \pm 0.03$ | $0.431 \pm 0.001$ |
| $B^+ \to J/\psi K^+$ | $16.39 \pm 0.07$ | $29.33 \pm 0.04$ | $34.68 \pm 0.09$ | $1.667 \pm 0.009$ |

**Table 4.16:** Summary of the efficiencies for each mode measured with simulation corresponding to Run 1 conditions.

For $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$, $B_s^0 \to \overline{D}^0 J/\psi$ and $B^+ \to D_s^+ J/\psi$, no simulation samples with Run 1 conditions are available. To account for this, the efficiency of the most similar mode for which Run 1 simulation is available is used to estimate the Run 1 efficiency through comparison with the corresponding Run 2 efficiencies,

$$\varepsilon^{\text{ch.X}}_{\text{Missing,Run 1}} = \varepsilon^{\text{ch.X}}_{\text{Available,Run 2}} \cdot \frac{\varepsilon^{\text{ch.Y}}_{\text{Available,Run 1}}}{\varepsilon^{\text{ch.Y}}_{\text{Available,Run 2}}}, \tag{4.8}$$

where $\varepsilon^{\text{ch.}\beta}_{\alpha,\text{Run i}}$ is the efficiency for decay $\beta$ in Run i, and $\alpha$ refers to whether a Run 1 simulation sample is available for the decay. For $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$, $B^0 \to \overline{D}^0 \mu^+ \mu^-$ and $B^0 \to \overline{D}^0 J/\psi$ decays are used respectively as the similar decay (decay Y in Eq. (4.8)), while for $B^+ \to D_s^+ J/\psi$, $B^+ \to D_s^+ \mu^+ \mu^-$ is used. The resulting efficiency estimates are given in Table 4.17.

| Decay Mode | Efficiency (%) |
|---|---|
| $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | $0.090 \pm 0.002$ |
| $B_s^0 \to \overline{D}^0 J/\psi$ | $0.067 \pm 0.002$ |
| $B^+ \to D_s^+ J/\psi$ | $0.092 \pm 0.003$ |

**Table 4.17:** The Run 1 selection efficiencies calculated using Eq. (4.8) for modes without available Run 1 simulation samples.

### 4.3.1 Tracking Efficiency Correction

Differences in the tracking efficiency between simulation and data are known to exist,

$$i.e.\ r_{track} = \frac{\varepsilon_{data}}{\varepsilon_{sim}} \neq 1.0\,, \tag{4.9}$$

where $\varepsilon_{data/sim}$ is the tracking efficiency in data/simulation. The ratio depends on the track momentum and pseudorapidity, and has been measured as a function of each [9]. In data, the efficiency is measured using a tag-and-probe method with $J/\psi \to \mu^+\mu^-$ decays, where one muon track (the "tag") is fully reconstructed, and the second muon track (the "probe") is only partially reconstructed. If the probe muon has sufficient momentum, the dimuon invariant mass can be used to discriminate against background without the full track reconstruction, allowing an efficiency measurement to be made.

As branching fractions are, in general, measured as a ratio between two modes, the differences in tracking efficiencies cancel out to first order when the number and types of final-state tracks are the same. However, the final states of the $B^+ \to D_s^+\mu^+\mu^-$, $B^+ \to D_s^+ J/\psi$, and $B_c^+ \to D_s^+ J/\psi$ modes have two additional tracks compared to the normalisation mode $B^+ \to J/\psi\, K^+$ ($K^+K^-\pi^+\mu^+\mu^-$ compared to $K^+\mu^+\mu^-$). This means $r_{track}$ must be calculated and included in the branching ratio calculation.

Measurements of $r_{track}$ as a function of track momentum ($p$) and pseudorapidity ($\eta$) are provided by the LHCb collaboration[4], and are shown in Fig. 4.14. Two values of $r_{track}$ are determined for each relevant decay and year of data taking, one for the missing kaon track ($r_{K_2}$) and one for the missing pion track ($r_\pi$). Each is calculated from the weighted average of the $r_{track}$ values given in Fig. 4.14, the weights determined from the occupancy of the extra tracks in the specific $\eta$ and $p$ bins in simulation. The calculated $r$ values are given in Tables 4.18 and 4.19, with the final correction applied being the product of the two. The systematic uncertainty on the ratios is given as 0.8% for Run 2 and 0.4% for Run 1, with material interactions imposing a further 1.4% uncertainty for pions and 1.1% for kaons. These are included as systematic uncertainties in the final measurements.

[4]https://twiki.cern.ch/twiki/bin/viewauth/LHCbInternal/LHCbTrackingEfficiencies, accessed 15th August 2019

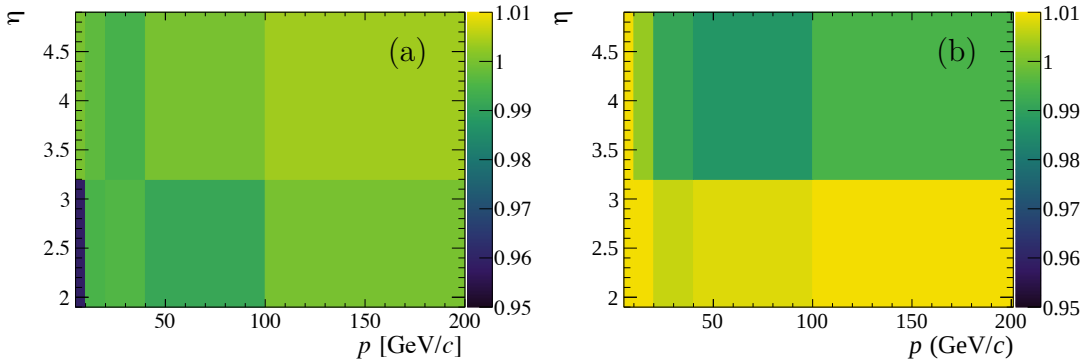



**Figure 4.14:** The ratio $r_{track} = \varepsilon_{data}/\varepsilon_{sim}$ between data and simulation as a function of track momentum ($p$) and pseudorapidity ($\eta$) for (a) Run 2 and (b) Run 1 samples. These plots are taken from the LHCb twiki on 15th August 2019.

| Decay | $r_{K_2}$ | |
|---|---|---|
| | Run 2 | Run 1 |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | $0.9897 \pm 0.0036$ | $1.0097 \pm 0.0097$ |
| $B^+ \to D_s^+ J/\psi$ | $0.9892 \pm 0.0067$ | – |
| $B_c^+ \to D_s^+ J/\psi$ | $0.9897 \pm 0.0056$ | $1.0123 \pm 0.0072$ |

**Table 4.18:** The values of $r_{track}$ determined as a function of decay and data-taking period for the extra kaon track. The errors quoted are purely statistical. Additional systematic uncertainties apply, as described in the text. A dash (–) indicates the relevant simulation sample was not available.

| Decay | $r_{\pi}$ | |
|---|---|---|
| | Run 2 | Run 1 |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | $0.9888 \pm 0.0036$ | $1.0515 \pm 0.0100$ |
| $B^+ \to D_s^+ J/\psi$ | $0.9873 \pm 0.0072$ | – |
| $B_c^+ \to D_s^+ J/\psi$ | $0.9883 \pm 0.0058$ | $1.0153 \pm 0.0073$ |

**Table 4.19:** The values of $r_{track}$ determined as a function of decay and data-taking period for the extra pion track. The errors quoted are purely statistical. Additional systematic uncertainties apply, as described in the text. A dash (–) indicates the relevant simulation sample was not available.

### 4.3.2 Final Efficiency Values

As Run 1 and Run 2 data are fitted together, the final efficiency used in measuring the branching fraction for each mode is taken as a weighted combination of the efficiencies measured for each Run, calculated according to,

$$\varepsilon_{Total} = \frac{\mathcal{L}_1 \varepsilon_1 + R_b \mathcal{L}_2 \varepsilon_2}{\mathcal{L}_1 + R_b \mathcal{L}_2}, \tag{4.10}$$

where $\varepsilon_{1/2}$ is the efficiency measured with Run 1/2 simulation, $\mathcal{L}_{1/2}$ is the integrated luminosity of Run 1/2, and $R_b$ is the ratio of $b$-quark production cross-sections between 13 and 7 TeV, given in [58, 59].

For signal modes where the normalisation channel has two fewer tracks, a modified form of Eq. (4.10) must be used which incorporates the tracking efficiency correction,

$$\varepsilon_{Total} = \frac{\mathcal{L}_1 . r_{K_2}^{Run1} . r_{\pi}^{Run1} . \varepsilon_1 + R_b \mathcal{L}_2 . r_{K_2}^{Run2} . r_{\pi}^{Run2} . \varepsilon_2}{\mathcal{L}_1 + R_b \mathcal{L}_2} \,. \tag{4.11}$$

Note for $B^+ \to D_s^+ J/\psi$ there is no Run 1 simulation available, and so the correction values measured for $B^+ \to D_s^+ \mu^+ \mu^-$ are used. The corrections determined from simulation with Run 2 conditions are seen to be consistent between the two decays, indicating this is a valid substitution.

The final efficiency values used in the calculation of branching ratios are given in Table 4.20. The $B_c^+ \to D_s^+ J/\psi$ decay appears twice as no tracking correction is required when it is used to normalise the $B_c^+ \to D_s^+ \mu^+ \mu^-$ measurement, but the correction is needed when measuring the $B_c^+ \to D_s^+ J/\psi$ branching fraction itself (see Subsection 4.6.1).

| Decay Mode | Efficiency (%) |
|---|---|
| $B^0 \to \overline{D}^0 \mu^+ \mu^-$ | $0.121 \pm 0.003$ |
| $B^0 \to \overline{D}^0 J/\psi$ | $0.115 \pm 0.003$ |
| $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | $0.129 \pm 0.004$ |
| $B_s^0 \to \overline{D}^0 J/\psi$ | $0.098 \pm 0.004$ |
| $B^+ \to D_s^+ \mu^+ \mu^-$ * | $0.084 \pm 0.002$ |
| $B^+ \to D_s^+ J/\psi$ * | $0.129 \pm 0.006$ |
| $B_c^+ \to D_s^+ \mu^+ \mu^-$ | $0.025 \pm 0.001$ |
| $B_c^+ \to D_s^+ J/\psi$ | $0.074 \pm 0.002$ |
| $B_c^+ \to D_s^+ J/\psi$ * | $0.076 \pm 0.002$ |
| $B^0 \to J/\psi K^{*0}$ | $0.649 \pm 0.007$ |
| $B^+ \to J/\psi K^+$ | $1.979 \pm 0.024$ |

**Table 4.20:** The final efficiencies for each mode, calculated using Eq. (4.10), or Eq. (4.11) for decays with an asterisk (*). The errors quoted here are derived purely from the statistics of the simulated samples and the errors on $R$ in Eq. (4.10) and Eq. (4.11). The error introduced by the tracking correction is applied later as a separate systematic uncertainty.

## 4.4 Signal Extraction through Fits to Data

Fits to the invariant mass distributions of $B$-meson candidates are used to extract signal yields for each decay. All fits to data use the maximum likelihood method, were performed using the RooFit package [60], and the fit characteristics are summarised in Table 4.21. Explanations of the different function shapes used to fit the data are given in Appendix A.

| Decay Mode(s) | Type | Figure(s) | Components |
|---|---|---|---|
| $B^0 \to \overline{D}^0 \mu^+ \mu^-$ & $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | Search | 4.22a | $B^0 \to \overline{D}^0 \mu^+ \mu^-$ DCB<br>$B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ DCB<br>$B^0$ semileptonic KDE shape<br>$B_s^0$ semileptonic KDE shape<br>background exponential |
| $B^0 \to \overline{D}^0 J/\psi$ & $B_s^0 \to \overline{D}^0 J/\psi$ | Search | 4.22b | $B^0 \to \overline{D}^0 J/\psi$ DCB<br>$B_s^0 \to \overline{D}^0 J/\psi$ DCB<br>$B^0 \to K^+\pi^- J/\psi$ Gaussian<br>background exponential |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | Search | 4.23a | $B^+ \to D_s^+ \mu^+ \mu^-$ DCB<br>background exponential |
| $B^+ \to D_s^+ J/\psi$ | Search | 4.23b | $B^+ \to D_s^+ J/\psi$ DCB<br>$B^+ \to K^+\pi^-\pi^+ J/\psi$ KDE shape<br>background exponential |
| $B_c^+ \to D_s^+ \mu^+ \mu^-$ | Search | 4.24a | $B_c^+ \to D_s^+ \mu^+ \mu^-$ DCB<br>background exponential |
| $B_c^+ \to D_s^+ J/\psi$ | Search & Norm. | 4.24b | $B_c^+ \to D_s^+ J/\psi$ DCB<br>$\mathcal{A}^{\pm\pm}$ Little Horns & Hill Analytic PDFs<br>$\mathcal{A}^{00}$ Horns & Hill Analytic PDFs<br>background exponential |
| $B^0 \to J/\psi K^{*0}$ | Norm. | 4.25a | $B^0 \to J/\psi K^{*0}$ Hypatia<br>$B_s^0 \to J/\psi K^{*0}$ Hypatia<br>background exponential |
| $B^+ \to J/\psi K^+$ | Norm. | 4.25b | $B^+ \to J/\psi K^+$ Hypatia<br>background exponential |

**Table 4.21:** Summary of the fits performed to extract signal yields for the search and normalisation modes. DCB stands for Double Crystal-Ball (see Section A.3), and KDE for Kernal Density Estimation (see Section A.8). $\mathcal{A}^{\pm\pm}$ and $\mathcal{A}^{00}$ are helicity amplitudes of partially reconstructed $B_c^+ \to D_s^{*+} J/\psi$ decays, explained in more detail in Subsection 4.4.1.

### 4.4.1 Validation of Fit Components

Each component except for the background exponentials were fit to simulated data to ensure they properly model the distributions and also to determine appropriate parameters to use in the final fit to data.

#### $B \to D\mu^+\mu^-$ Signal Shape

The invariant-mass peaks for all $B \to D\mu^+\mu^-$ modes are modelled with a Double Crystal Ball (DCB) shape with shared tail parameters, as detailed in Appendix A.3. Two DCB shapes are included in the fit to $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0\mu^+\mu^-$ candidates, one for each decay, and the same is true for the $B^0 \to \overline{D}^0 J/\psi$ and $B_s^0 \to \overline{D}^0 J/\psi$ invariant mass fit.

In the fit to simulation, the exponent in the power-law tail ($n$) is fixed to unity, and the tail is constrained to be on the low-mass side of the shape ($\alpha > 0$). An example fit to the $B^0 \to \overline{D}^0\mu^+\mu^-$ simulated sample is shown in Fig. 4.15.

In the fit to data, all parameters are fixed to the values from the simulation fit, with the exception of the mean $B$ mass, which is fixed to the mean measured by the fit to the respective normalisation mode. This is done to account for any differences between the data and the values from simulation and the literature. For the $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$ modes specifically, due to the low $B_s^0 \to J/\psi K^{*0}$ yield, the mean $B_s^0$ mass is fixed to the PDG $B_s^0$ mass [28], except offset by the same amount the $B^0$ mass is offset from the PDG value in the $B^0 \to J/\psi K^{*0}$ fit ($\sim 0.63\,\mathrm{MeV}/c^2$).

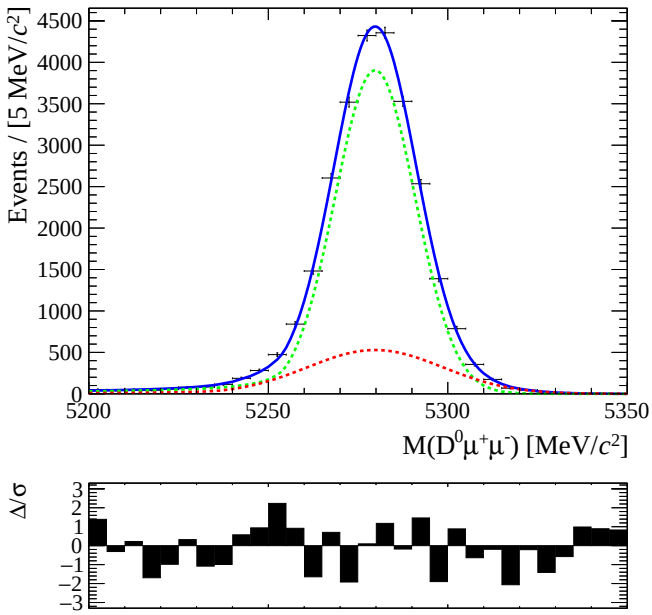



**Figure 4.15:** Double crystal-ball fit to simulated $B^0 \to \overline{D}^0 \mu^+ \mu^-$ decays, with the two crystal-ball shapes shown by the dotted lines, and the sum by the blue solid line.

### $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ Signal Shapes

The invariant mass distribution for the $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ modes are modelled by a two-tailed Hypatia distribution [61], detailed in Section A.4. In the fits to simulation and data, $\zeta$ (a parameter of the hyperbolic core) and $\beta$ (which introduces a variance-dependence on the mean) are fixed to 0. All other parameters are floated in the fits to simulation, with $\lambda$ (another parameter of the hyperbolic core) constrained to be less than $-1$. Fits to simulated data for both normalisation modes are shown in Fig. 4.16.

In the fits to data, for both the $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ samples, the value of $\lambda$ and the exponent of the high-mass tail ($n_r$) are fixed to the values determined by the fit to simulation. All other parameters are left floating. In the $B^0 \to J/\psi K^{*0}$ mass fit, the $B_s^0 \to J/\psi K^{*0}$ decay component is modelled by a second Hypatia shape which shares its floating parameters with the $B^0 \to J/\psi K^{*0}$ shape except for the mean mass.

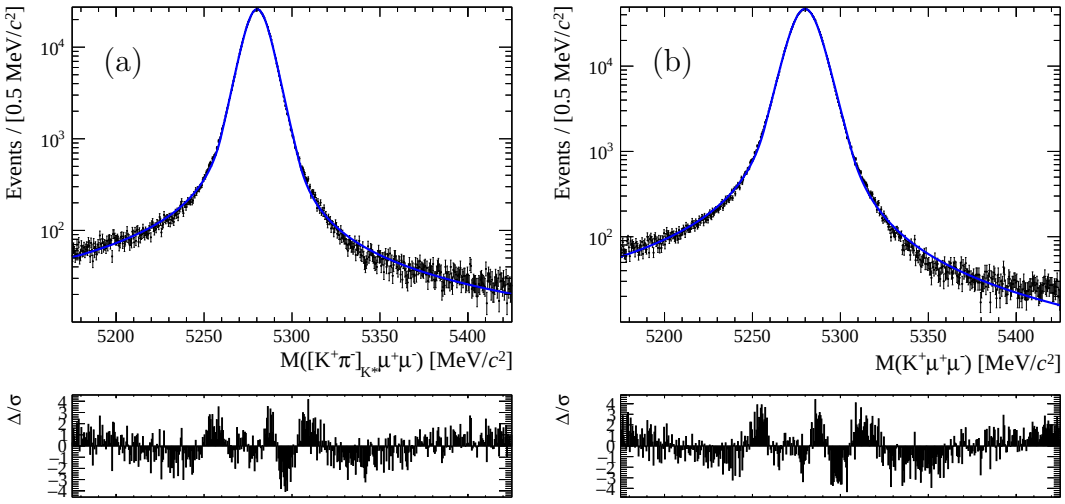



**Figure 4.16:** Fits with a two-tailed Hypatia shape to (a) $B^0 \to J/\psi\, K^{*0}$ simulated decays, and (b) $B^+ \to J/\psi\, K^+$ simulated decays.

### Semileptonic $B^0$ and $B_s^0$ Decay Background

While $B^0 \to \left(D^{*-} \to \overline{D}^0\pi^-\right)\mu^+\nu_\mu$ decays are removed with a cut in the selection (see Subsection 4.2.7), several other semileptonic backgrounds remain which must be accounted for in the fit:

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

where the $\pi^-$ is misidentified as a $\mu^-$, and

- $B_s^0 \to \left(D_{s2}^*(2573)^- \to D^0K^-\right)\mu^+\nu_\mu$,
- $B_s^0 \to \left(D_{s1}(2536)^- \to D^{*0}K^-\right)\mu^+\nu_\mu$,

where $D^{*0} \to D^0\,(\pi^0/\gamma)$, and the $K^-$ decays within LHCb via $K^- \to \mu^-\nu_\mu$. All these decays give the same final state as $B^0 \to \overline{D}^0\mu^+\mu^-$ when the neutral particles are not reconstructed.

The mass distributions of the above decays, over the $B^0 \to \overline{D}^0\mu^+\mu^-$ invariant mass plot range (*i.e.* $M(B^0) \in [5000, 5800]\,\text{MeV}/c^2$), are investigated using a simplified simulation [62], and kernal density estimation [63] (see Section A.8)

is used to create a PDF for each. The resultant shapes are shown applied to independent simulation samples in Fig. 4.17.

The shapes shown in Fig. 4.17 are combined into two fit components which are then included in the fit to the $B^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0\mu^+\mu^-$ invariant masses. The first combines the $B^0 \to \overline{D}^0\pi^-\mu^+\nu_\mu$ and $B^0 \to \overline{D}^{*0}\pi^-\mu^+\nu_\mu$ shapes, with the ratios between them derived from the branching fractions of the respective decays along with the efficiency for each to be within the $M(B^0) \in [5000, 5800]\,\mathrm{MeV}/c^2$ invariant mass window. The overall yield of $B^0 \to \overline{D}^0\pi^-\mu^+\nu_\mu$ and $B^0 \to \overline{D}^{*0}\pi^-\mu^+\nu_\mu$ decays is then fixed relative to the thirty-five $B^0 \to D^{*-}\mu^+\nu_\mu$ decays removed by the $\Delta M$ cut described in Subsection 4.2.7. Again accounting for the branching ratios and efficiencies to be within the plot range, $11.8 \pm 0.3$ $B^0 \to \overline{D}^0\pi^-\mu^+\nu_\mu$ and $B^0 \to \overline{D}^{*0}\pi^-\mu^+\nu_\mu$ events are expected.

The second fit component combines the $B_s^0 \to D_{s2}^*(2573)^-\mu^+\nu_\mu$ and $B_s^0 \to D_{s1}(2536)^-\mu^+\nu_\mu$ shapes using the ratio [64]

$$\frac{\mathcal{B}\left(B_s^0 \to D_{s2}^*(2573)^- X\mu^+\nu_\mu\right)}{\mathcal{B}\left(B_s^0 \to D_{s1}(2536)^- X\mu^+\nu_\mu\right)} = 0.61 \pm 0.15, \tag{4.12}$$

assumed to hold true with a nonexisting $X$, and as before accounting for the efficiencies of the reconstructed decays within $M(B_s^0) \in [5000, 5800]\,\mathrm{MeV}/c^2$. Due to the unknown branching ratios of $D_{s2}^*(2573)^- \to D^0K^-$ and $D_{s1}(2536)^- \to D^{*0}K^-$, the expected yield of the combined shape cannot be calculated and so is floated.

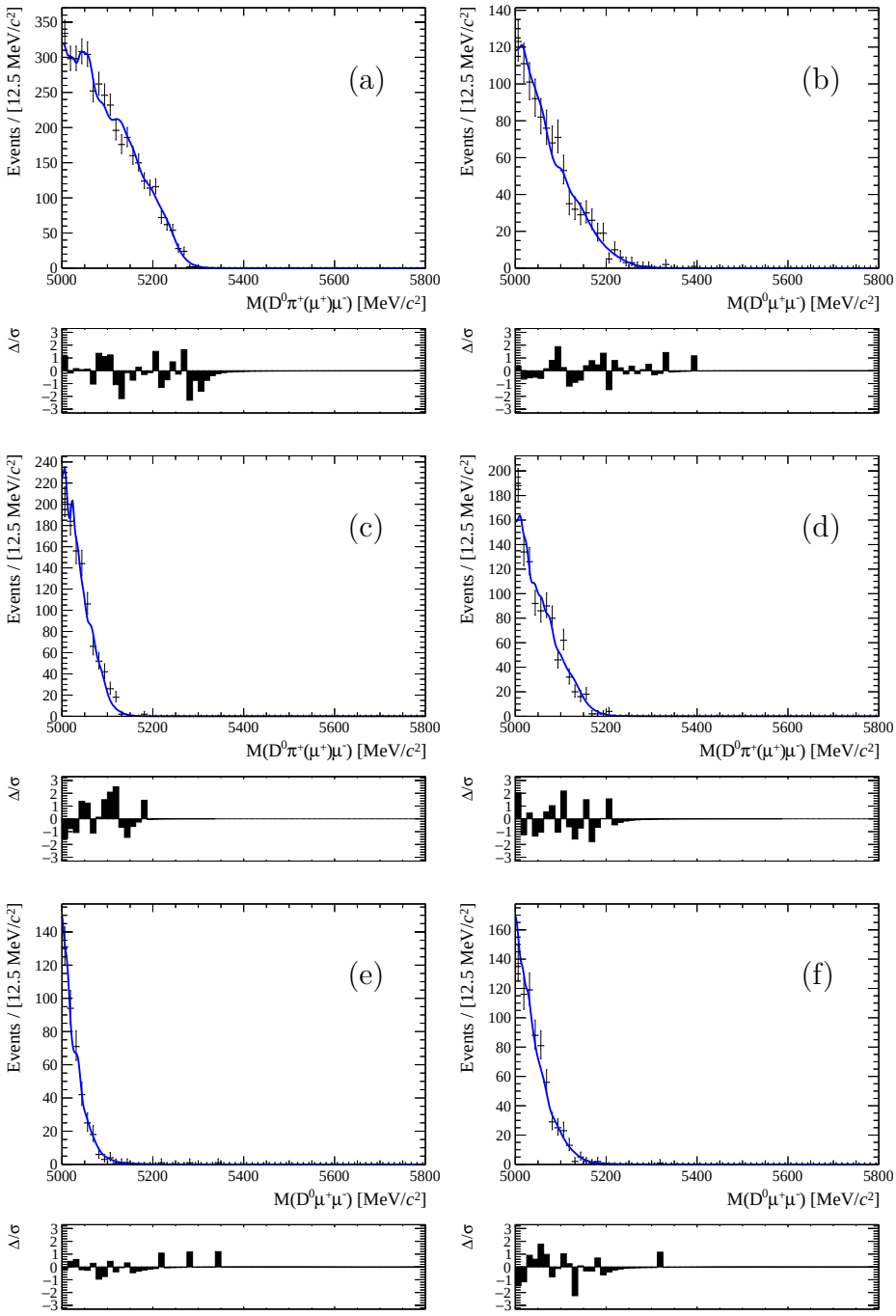



**Figure 4.17:** Simplified simulation samples [62] of (a) $B^0 \to \overline{D}^0\pi^-\mu^+\nu_\mu$, (b) $B_s^0 \to \left(D_{s2}^*(2573)^- \to D^0K^-\right)\mu^+\nu_\mu$, (c, d) $B^0 \to \overline{D}^{*0}\pi^-\mu^+\nu_\mu$, and (e, f) $B_s^0 \to \left(D_{s1}(2536)^- \to D^{*0}K^-\right)\mu^+\nu_\mu$ decays, with (c, e) $D^{*0} \to D^0\pi^0$ and (d, f) $D^{*0} \to D^0\gamma$. The $\pi^-$ particles are reconstructed as muons, the $K^-$ decays via $K^- \to \mu^-\nu_\mu$, and neutral particles are not reconstructed. The samples are fit with probability density functions generated using kernal density estimation from separate, independent samples.

### $B^0 \to K^+\pi^- J/\psi$ Background

As discussed in Subsection 4.2.9, some $B^0 \to K^+\pi^- J/\psi$ background remains in the $B^0 \to \overline{D}^0 J/\psi$ sample after the selections are applied. This is accounted for in the fit by adding a Gaussian component with a fixed yield. The mean mass is fixed to the mean of the $B^0 \to J/\psi K^{*0}$ fit, and the width to the width of the second, wider Crystal-Ball in the signal $B^0 \to \overline{D}^0 \mu^+\mu^-$ simulation fit.

Two $D$-mass sideband regions with the same width as the $D^0$-mass window, $M(D^0) \in [1785, 1835]\,\mathrm{MeV}/c^2$ and $M(D^0) \in [1900, 1950]\,\mathrm{MeV}/c^2$, are used to determine the yield of $B^0 \to K^+\pi^- J/\psi$ decays which can be expected in the $D^0$-mass signal window, and thus the yield which will appear in the final $B^0 \to \overline{D}^0 J/\psi$ mass fit. In the fit to determine the yields of $B^0 \to K^+\pi^- J/\psi$ decays, shown in Fig. 4.18, the combinatorial background is modelled with an exponential shape. The average of the two yields is used in the $B^0 \to \overline{D}^0 J/\psi$ mass fit, resulting in an expectation of $4.2 \pm 1.5$ $B^0 \to K^+\pi^- J/\psi$ decays.

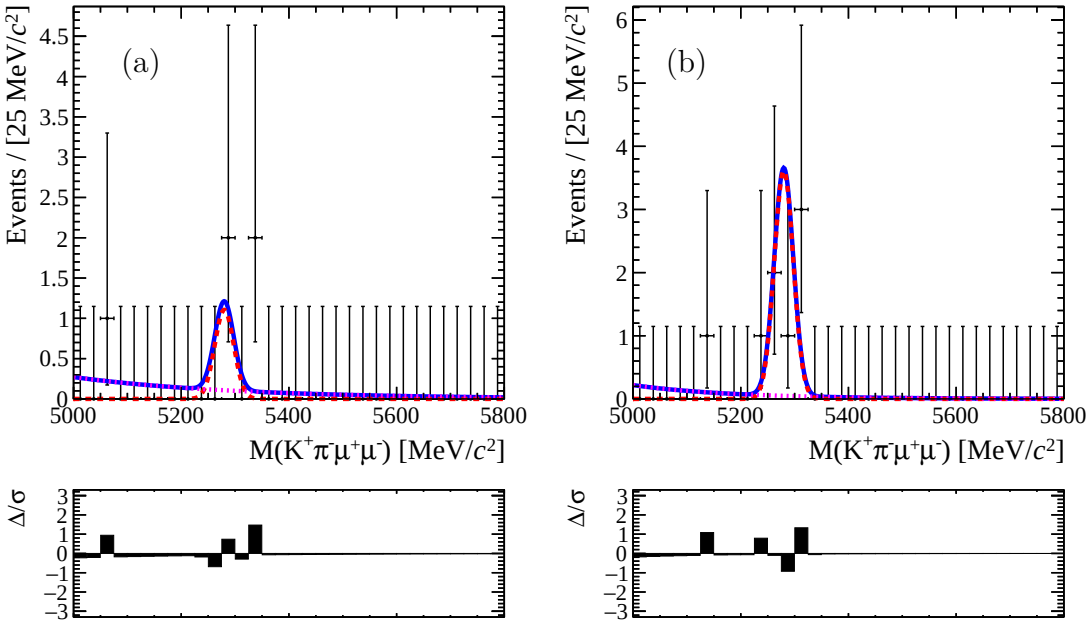



**Figure 4.18:** Fit of the $B^0$-invariant-mass spectrum in data, with no kinematic fit constraints applied (see Subsection 4.2.1). The selection for $B^0 \to \overline{D}^0 J/\psi$ decays given in Table 4.10 has been applied, except the $D^0$ mass window cut has been replaced with (a) $M(D^0) \in [1785, 1835]\,\mathrm{MeV}/c^2$ and (b) $M(D^0) \in [1900, 1950]\,\mathrm{MeV}/c^2$ to select a sideband region with the same width as the original cut. The red dashed components show the $B^0 \to K^+\pi^- J/\psi$ signal shape, the magenta dashed components show the combinatorial background, and the blue solid curves the sum of the relevant signal and background components. The measured yields are (a) $2.0 \pm 1.5$ and (b) $6.4 \pm 2.6$ $B^0 \to K^+\pi^- J/\psi$ decays.

### $B^+ \to K^+\pi^-\pi^+ J/\psi$ Background

Although the $D_s^+$ flight-significance cut is seen to remove any peaking structure in the $D_s^+$ mass sidebands for $B^+ \to (D_s^+ \to K^+K^-\pi^+)\, J/\psi$ decays (see Subsection 4.2.9), some $B^+ \to K^+\pi^-\pi^+ J/\psi$ decays, where the $\pi^-$ is misidentified as a $K^-$, remain in the sample. Similar to the $B^0 \to \overline{D}^0\pi^-\mu^+\nu_\mu$ component in the $B^0 \to \overline{D}^0\mu^+\mu^-$ mass fit, $B^+ \to K^+\pi^-\pi^+ J/\psi$ decays are accounted for with a kernal estimated fit component [63] derived from a simplified simulation sample [62]. $B^+ \to K^+\pi^-\pi^+ J/\psi$ decays were generated with a kaon mass hypothesis for the $\pi^-$ and the resultant shape, applied to a separate independent sample, is shown in Fig. 4.19.

The $B^+ \to K^+\pi^-\pi^+ J/\psi$ shape is included with a fixed yield, determined from the number of $B^+ \to K^+\pi^-\pi^+ J/\psi$ decays observed in the $D_s^+$ sideband regions of

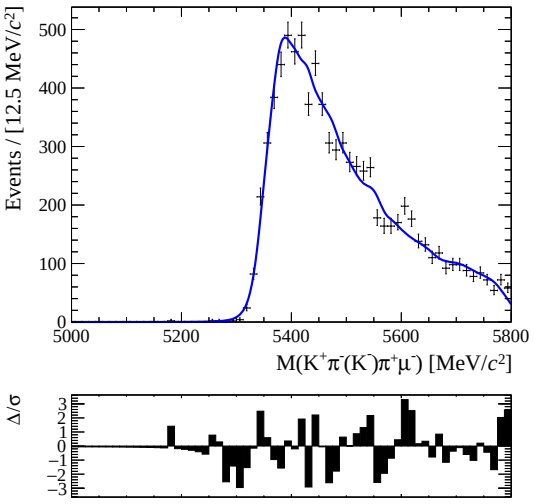



**Figure 4.19:** Simulated $B^+ \to K^+\pi^-\pi^+ J/\psi$ decays, generated with a simplified simulation [62] and with a kaon mass hypothesis for the $\pi^-$, fitted with a probability density function generated using kernal density estimation from an independent sample.

$B^+ \to (D_s^+ \to K^+K^-\pi^+)\, J/\psi$ candidates reconstructed with a pion mass hypothesis for the $K^-$. Shown in Fig. 4.20, $5.0 \pm 3.2$ events are observed in each sideband region, therefore five events can also be expected to fall within the $D_s^+$-mass window. The fraction of $B^+ \to K^+\pi^-\pi^+ J/\psi$ events within the $B^+$-mass window is $\sim 84\%$, therefore a yield of $4.2 \pm 2.0$ $B^+ \to K^+\pi^-\pi^+ J/\psi$ decays can be expected in the $B^+ \to D_s^+ J/\psi$ signal region.

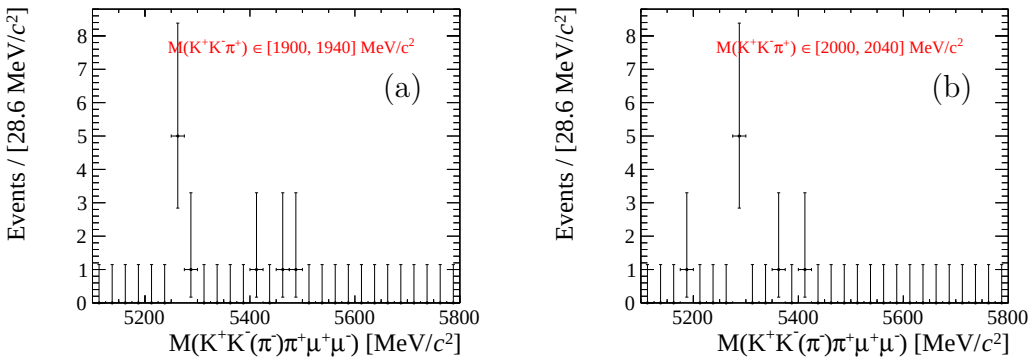



**Figure 4.20:** The unconstrained (i.e. no kinematic fit constraints) $B^+$-invariant-mass spectrum of $B^+ \to (D_s^+ \to K^+K^-\pi^+)\, J/\psi$ decay candidates reconstructed with a pion mass hypothesis for the $K^-$, in the range (a) $M(D_s^+) \in [1900, 1940]\,\mathrm{MeV}/c^2$ and (b) $M(D) \in [2000, 2040]\,\mathrm{MeV}/c^2$. The selection given in Table 4.11 for $B^+ \to D_s^+ J/\psi$ decays has been applied, except for the $D_s^+$-mass cut.

### $B_c^+ \to D_s^+ J/\psi$ Partially Reconstructed Background

The $B_c^+ \to D_s^+ J/\psi$ fit requires a partially reconstructed background component originating from $B_c^+ \to D_s^{*+} J/\psi$ decays, where $D^{*+} \to \overline{D}^0 \pi^+$ and the $\gamma$ or $\pi^0$ is not reconstructed. As a pseudoscalar particle ($B_c^+$) decays into two vector particles ($D_s^{*+}$ and $J/\psi$), the decay is described by three helicity amplitudes: $\mathcal{A}^{++}$, $\mathcal{A}^{--}$, and $\mathcal{A}^{00}$. The $D_s^+ J/\psi$ mass distribution with the missing $\gamma/\pi^0$ is identical for $\mathcal{A}^{++}$ and $\mathcal{A}^{--}$, therefore the two amplitudes can be combined into $\mathcal{A}^{\pm\pm}$. As the decay results in a $\gamma$ or $\pi^0$, this gives four combinations to model.

Three analytic probability density functions (PDFs) are used: so-called Horns, Hill and Little Horns, as described in Appendix A.7. Table 4.22 gives the shape used to model each combination of missing particle and helicity amplitude.

| Missing Particle | Analytic PDF | |
|---|---|---|
| | $\mathcal{A}^{\pm\pm}$ | $\mathcal{A}^{00}$ |
| $\gamma$ | Little Horns | Hill |
| $\pi^0$ | Hill | Horns |

**Table 4.22:** Shapes used to model the different components of the partially reconstructed $B_c^+ \to D_s^{*+} J/\psi$ background.

The shapes are primarily formed of a double Gaussian. In the fit to simulation, the relative shift between the Gaussians ($\delta$) is fixed to 0, and the kinematic endpoints of the Hill/Horns shapes ($a$ and $b$) are fixed from calculations based on the masses of the particles involved. Each helicity amplitude is treated separately, but the $\gamma$ and $\pi^0$ contributions are fitted together. This introduces an extra free parameter, the fraction of decays with a $\pi^0$ compared to a $\gamma$, which is fixed to 0.058 [28]. None of the remaining free parameters ($\xi$, $\sigma$, $R_\sigma$, and $f_G$) are shared between the shapes. The results of this fit are shown in Fig. 4.21 for a simulated $B_c^+ \to (D_s^{*+} \to D_s^+ (\gamma/\pi^0))\, J/\psi$ sample where the $\gamma/\pi^0$ is not reconstructed.

In the fit to data, all free parameters relating to the shapes are fixed from the fit to simulation, as well as the ratio of $\pi^0$ to $\gamma$ decays. The only floating parameters are the ratio between $\mathcal{A}^{\pm\pm}$ and $\mathcal{A}^{00}$ decays and the overall yield of partially reconstructed decays.

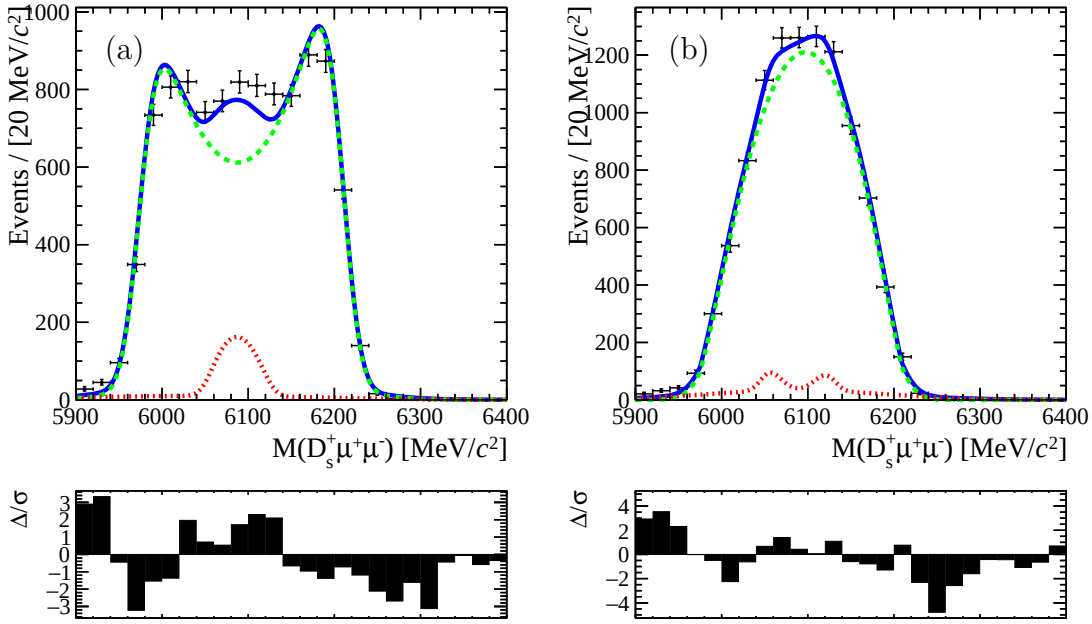



**Figure 4.21:** Fit to simulation for partially reconstructed $B_c^+ \to D_s^{*+} J/\psi$ decays, with $D_s^{*+} \to D_s^+ (\gamma/\pi^0)$ where the neutral particle is missed, for (a) $\mathcal{A}^{\pm\pm}$ and (b) $\mathcal{A}^{00}$ helicity amplitudes. The large green dashed curves show the $\gamma$ component, the smaller red dotted the $\pi^0$.

### 4.4.2 Fits to Selected Decay Candidates

After unblinding, the mass fits to selected $B \to D\mu^+\mu^-$ candidates in data, used to extract the signal yields, are shown in Figs 4.22, 4.23, and 4.24, and for the $B^0 \to J/\psi\, K^{*0}$ and $B^+ \to J/\psi\, K^+$ normalisation modes in Fig. 4.25.

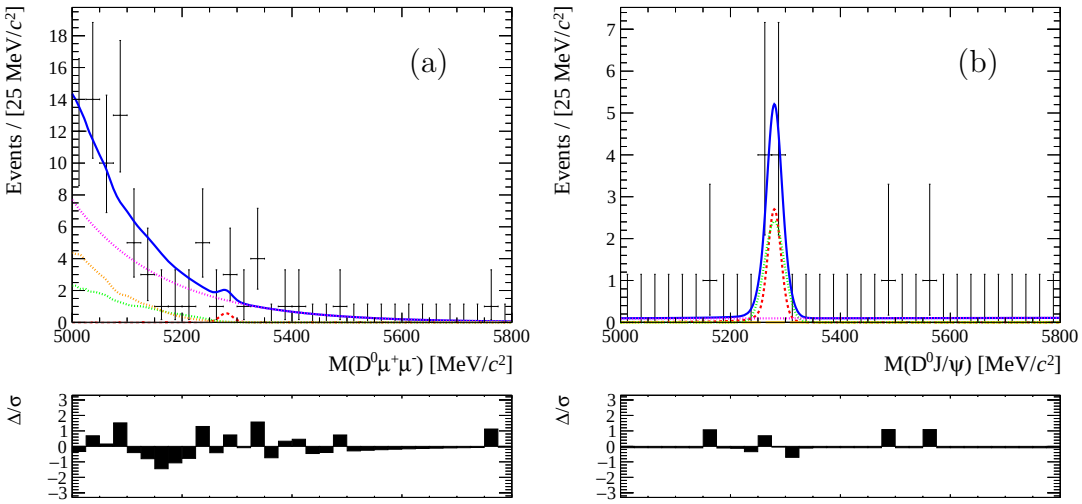



**Figure 4.22:** Fit to data for (a) $B^0 \to \overline{D}^0\mu^+\mu^-$ & $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and (b) $B^0 \to \overline{D}^0 J/\psi$ & $B_s^0 \to \overline{D}^0 J/\psi$ decay candidates. The red components show the (a) $B^0 \to \overline{D}^0\mu^+\mu^-$ and (b) $B^0 \to \overline{D}^0 J/\psi$ signal peaks, and the magenta components show the combinatorial background in both. The green component in (a) shows the $B^0 \to \overline{D}^0\pi^-\mu^+\nu_\mu$ and $B^0 \to \overline{D}^{*0}\pi^-\mu^+\nu_\mu$ background shape, and the orange component the $B_s^0 \to D_{s2}^*(2573)^-\mu^+\nu_\mu$ and $B_s^0 \to D_{s1}(2536)^-\mu^+\nu_\mu$ background shape. Both components are unique to this fit. The green component in (b) shows the $B^0 \to K^+\pi^- J/\psi$ component. The signal components for $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$ decays are not visible. The blue solid curves show the sum of all the relevant components for each fit. Yields of $0.6 \pm 2.5$ ($B^0 \to \overline{D}^0\mu^+\mu^-$), $3.5 \pm 3.0$ ($B^0 \to \overline{D}^0 J/\psi$), $0.0 \pm 3.0$ ($B_s^0 \to \overline{D}^0\mu^+\mu^-$), and $0.0 \pm 1.1$ ($B_s^0 \to \overline{D}^0 J/\psi$) are measured, with background yields over the plot range of (a) $78 \pm 9$ and (b) $3 \pm 2$.

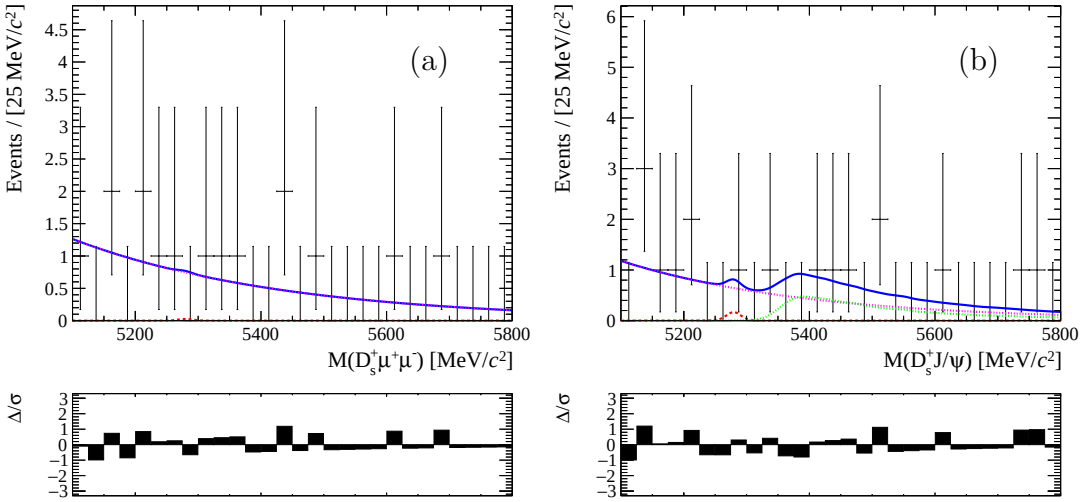



**Figure 4.23:** Fit to data for (a) $B^+ \to D_s^+\mu^+\mu^-$ and (b) $B^+ \to D_s^+ J/\psi$ decay candidates. The red components show the (a) $B^+ \to D_s^+\mu^+\mu^-$ and (b) $B^+ \to D_s^+ J/\psi$ peaks, and the magenta components show the combinatorial background in both. The green component in (b) shows the $B^+ \to K^+\pi^-\pi^+ J/\psi$ background, and is unique to this fit. The blue solid curves show the sum of all the relevant components in each fit. Yields of $0.0 \pm 0.9$ ($B^+ \to D_s^+\mu^+\mu^-$) and $0.2 \pm 1.1$ ($B^+ \to D_s^+ J/\psi$) are measured, with background yields over the plot range of (a) $15 \pm 4$ and (b) $13 \pm 4$.

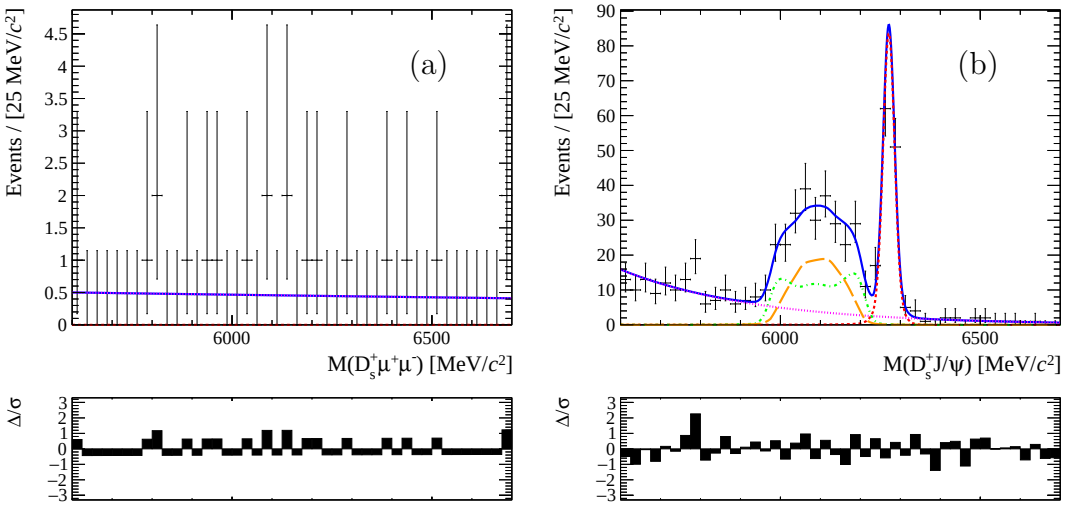



**Figure 4.24:** Fit to data for (a) $B_c^+ \to D_s^+\mu^+\mu^-$ and (b) $B_c^+ \to D_s^+ J/\psi$ decay candidates. In (a), the $B_c^+ \to D_s^+\mu^+\mu^-$ signal component is not visible. In (b), the red dashed line corresponds to the $B_c^+ \to D_s^+ J/\psi$ signal component, the green mixed line is the $\mathcal{A}^{\pm\pm}$ contribution to the partially reconstructed background, and the orange long-dashed line is the $\mathcal{A}^{00}$ contribution. In each fit, the magenta component shows the combinatorial background and the blue solid curve shows the sum of all the relevant components. Yields of $0.0 \pm 1.1$ ($B_c^+ \to D_s^+\mu^+\mu^-$) and $127 \pm 12$ ($B_c^+ \to D_s^+ J/\psi$) are measured, with background yields over the plot range of (a) $20 \pm 4$ and (b) $208 \pm 17$. The ratio of partially reconstructed $B_c^+ \to D_s^{*+} J/\psi$ decays to $B_c^+ \to D_s^+ J/\psi$ decays is $1.9 \pm 0.2$, and the fraction of $\mathcal{A}^{\pm\pm}$ helicity amplitudes out of to the total $B_c^+ \to D_s^{*+} J/\psi$ yield is $50 \pm 10$ %.

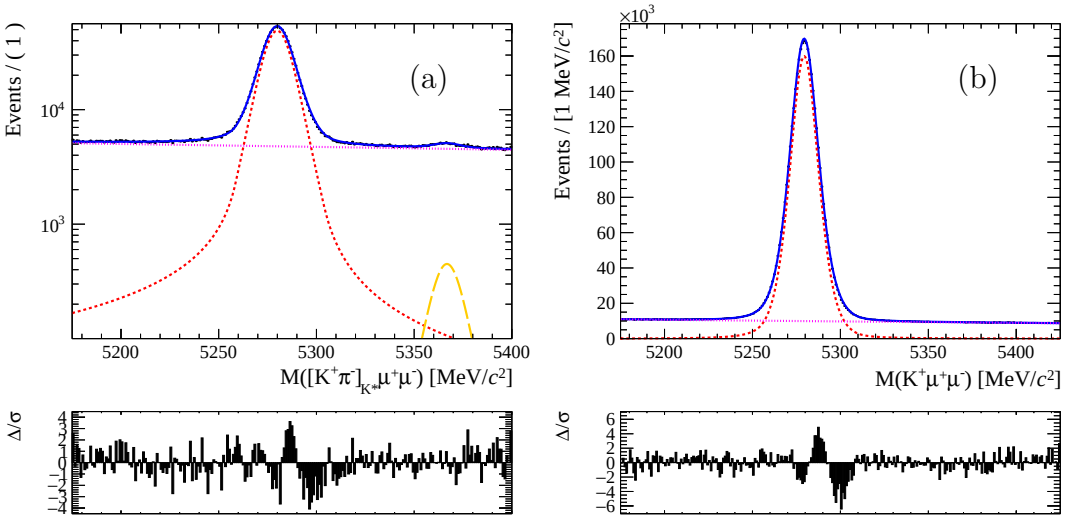



**Figure 4.25:** Fit to data for (a) $B^0 \to J/\psi K^{*0}$ and (b) $B^+ \to J/\psi K^+$ decays. The red dashed lines are the signal shapes, and the magenta dotted lines the background. In (a), the orange long-dashed line is an additional shape which accounts for $B_s^0 \to J/\psi K^{*0}$ decays. Yields of $(9.525 \pm 0.019) \times 10^5$ ($B^0 \to J/\psi K^{*0}$), $(8.70 \pm 0.49) \times 10^3$ ($B_s^0 \to J/\psi K^{*0}$), and $(3.5631 \pm 0.0032) \times 10^6$ ($B^+ \to J/\psi K^+$) are measured, with background yields over the plot range of (a) $(1.0730 \pm 0.0020) \times 10^6$ and (b) $(2.4403 \pm 0.0030) \times 10^6$.

### 4.4.3 Extracted Signal Yields

The yields extracted from the fits to data are presented in Table 4.23 [5]. In addition, a ratio of $B_c^+ \to D_s^{*+} J/\psi$ to $B_c^+ \to D_s^+ J/\psi$ decays is measured to be $1.9 \pm 0.2$, and the ratio of $\mathcal{A}^{\pm\pm}$ to $\mathcal{A}^{00}$ $B_c^+ \to D_s^{*+} J/\psi$ helicity amplitudes is $0.5 \pm 0.1$.

| Decay | Signal Yield |
|---|---|
| $B^0 \to \overline{D}^0 \mu^+ \mu^-$ | $0.8 \pm 2.7$ |
| $B^0 \to \overline{D}^0 J/\psi$ | $3.5 \pm 3.0$ |
| $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | $0.0 \pm 1.1$ |
| $B_s^0 \to \overline{D}^0 J/\psi$ | $0.0 \pm 1.1$ |
| $B^+ \to D_s^+ \mu^+ \mu^-$ | $0.0 \pm 0.9$ |
| $B^+ \to D_s^+ J/\psi$ | $0.2 \pm 1.1$ |
| $B_c^+ \to D_s^+ \mu^+ \mu^-$ | $0.0 \pm 1.1$ |
| $B_c^+ \to D_s^+ J/\psi$ | $127 \pm 12$ |
| $B^0 \to J/\psi K^{*0}$ | $(9.525 \pm 0.019) \times 10^5$ |
| $B^+ \to J/\psi K^+$ | $(3.5631 \pm 0.0032) \times 10^6$ |

**Table 4.23:** The signal yields extracted for each decay from fits to the invariant $B$-meson mass distributions.

[5] Note yield uncertainties which extend below 0 are quoted. A Bayesian analysis could be performed to prevent these from exceeding the physical bound.

## 4.5 Systematic Studies

The various sources of systematic uncertainty were investigated, detailed in this section. A summary of all the systematic uncertainties which apply to each measurement is given in Subsection 4.5.11.

### 4.5.1 Measured Error on Branching Ratios and $f_s/f_d$

The branching ratios of the normalisation modes and $D$-meson decays are used when calculating the branching ratio of the $B \to D\mu^+\mu^-$ modes (see Subsection 4.6.1). Therefore the uncertainties on these values, which are taken from Ref. [28], must be propagated into the measurement errors. The uncertainties in the $J/\psi \to \mu^+\mu^-$ and $K^{*0} \to K\pi$ branching ratios are considered to be negligible.

The error on $f_s/f_d$ must also be taken into account for the $B_s^0 \to \overline{D}^0\mu^+\mu^-$ and $B_s^0 \to \overline{D}^0 J/\psi$ modes. This is taken from Ref. [52]. The branching ratios for $B_c^+ \to D_s^+\mu^+\mu^-$ and $B_c^+ \to D_s^+ J/\psi$ are given as a function of $f_c/f_u$ and so no error needs to be included for this fraction.

### 4.5.2 Normalisation Mode Statistics

The statistics of the normalisation mode samples leads to an error on the measured branching ratios. These errors are given along with the yields in Table 4.23 for all three normalisation modes.

### 4.5.3 Simulation: Statistics

The selection efficiencies presented in Subsection 4.3.2, to be used in the branching ratio measurements, are limited in precision by the size of the available simulation samples. Additionally, the error on $R$ in Eq. (4.10) [58, 59] contributes when the individual efficiencies for Run 1 and Run 2 are combined. The uncertainty associated with each efficiency measurement was presented in Table 4.20.

### 4.5.4 Simulation: Differences with Data

As the efficiencies are measured using simulation only, any mismodelling in the simulation can result in an efficiency value which is not completely accurate for the data. Using the comparison of the efficiencies measured with $B^0 \to J/\psi K^{*0}$ simulation and ${}_s$Weighted data as a function of BDT output cut (as shown in Fig. 4.9), a systematic error is derived, calculated using

$$\text{error}\,(\%) = 100.0 \times \left| \frac{\varepsilon_{\text{sim}}}{\varepsilon_{\text{data}}} - 1 \right| , \tag{4.13}$$

where $\varepsilon_{sim}$ and $\varepsilon_{data}$ are the efficiencies measured for the simulation and ${}_s$Weighted data respectively.

The variation in the error value defined by Eq. (4.13) is shown as a function of BDT output cut in Fig. 4.26. The value at each individual BDT output cut which is used for each mode is then assigned as a systematic uncertainty.

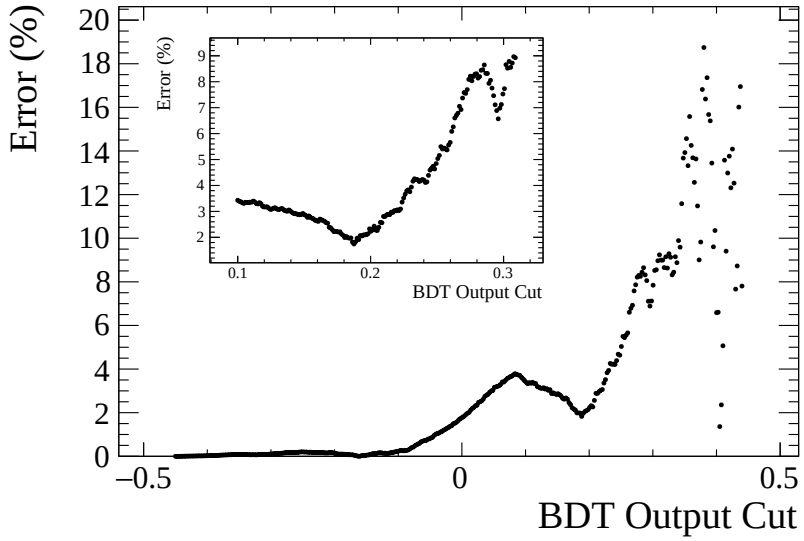



**Figure 4.26:** The systematic error, as defined by Eq. (4.13), as a function of BDT output cut, measured with $B^0 \to J/\psi K^{*0}$ simulation and ${}_s$Weighted data.

### 4.5.5 Simulation: Conditions for Different Years

Data taking conditions from 2012 (Run 1) and 2016 (Run 2) were used to generate the Run 1 and Run 2 simulation samples used in the searches, however these

conditions varied between each year of data taking. As all years of data are included in the analysis, a systematic uncertainty is determined to account for any differing conditions.

The efficiencies for $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ decays were determined using simulation generated with the conditions for each of the other years of data taking, and, using an equivalent formula to Eq. (4.13), the errors in the efficiency associated with the years of data-taking were determined. These were then combined into a single systematic uncertainty by weighting the errors according to the luminosity and ratio of $b$-quark production cross-sections for the given year. The final values are 5.0% ($B^0 \to J/\psi K^{*0}$) and 2.7% ($B^+ \to J/\psi K^+$).

### 4.5.6 Simulation: PID Correction

Depending on the simulation sample, the PID variables are resampled or transformed to better match the distribution observed in data. These procedures have associated systematic uncertainties, resulting from the size of the datasets and the choice of parameterisations used in determining the resampling distributions or variable transformations. These uncertainties vary between simulations samples; they peak at $\sim$ 3% but are most often much smaller.

### 4.5.7 Simulation: Trigger Correction

Correction factors accounting for the difference in the measured trigger efficiencies between simulation and data, presented in Subsection 4.2.2, are applied in the branching ratio calculation. The error in these correction factors are included as systematic uncertainties, and range between 3% and 7%.

### 4.5.8 Mass Fit: Choice of Background Shape

To investigate the effect of the choice of background parameterisation, the data were fit with a second order Chebychev polynomial for the background shape, instead of an exponential, and the resulting signal yields compared to that seen with the original fit. The difference in the measured signal between the two fits

is taken as the systematic uncertainty resulting from the choice of background shape, and is 2.0% or less for all modes.

### 4.5.9 Mass Fit: Fixed Parameters

Some or all of the parameters of the signal shapes in each fit are fixed from simulation, giving a source of systematic uncertainty. To evaluate this systematic, the fixed parameters for a given shape are varied (according to a Gaussian, where the width of the Gaussian is the uncertainty on the value determined from simulation) and the shape refitted to toy data. The new yield is compared to that from the original fit, and with many refits, a residual distribution is created, an example of which is shown in Fig. 4.27. The width of the distribution is assigned as the systematic uncertainty associated with fixing the corresponding parameters.

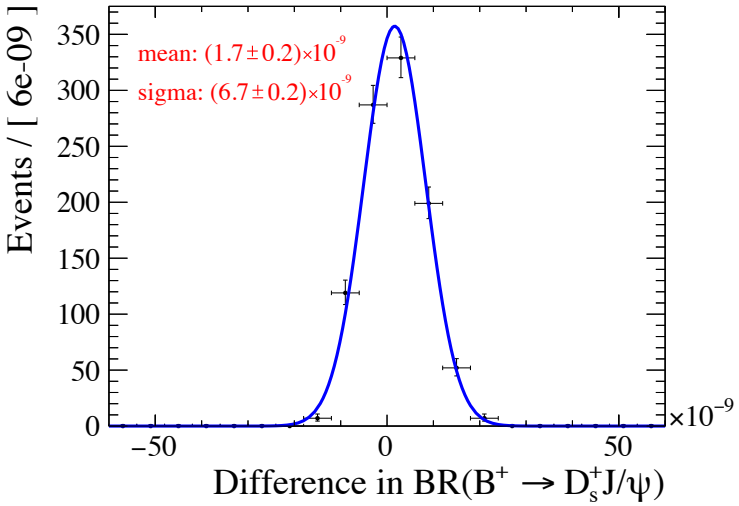



**Figure 4.27:** Variation in the measured branching ratio resulting from varying the values of the fixed parameters in the fit to toy data for $B^+ \to D_s^+ J/\psi$ candidates.

### 4.5.10 Tracking efficiency

The tracking efficiency ratios used in Subsection 4.3.1 have a systematic uncertainty of 0.8% (Run 2) and 0.4% (Run 1)[6]. To allow the systematic to be easily separated from other sources of uncertainty, 0.8% is taken as the systematic for both Run

[6]https://twiki.cern.ch/twiki/bin/view/LHCbInternal/LHCbTrackingEfficiencies, accessed 15th August 2019.

1 and Run 2 due to the larger weighting of Run 2 events compared to Run 1. This is a consequence of the relative luminosities and ratio of $b$-quark production cross-sections. The approximation leads to only a marginal overestimation of the systematic error and has a negligible effect on the overall result.

In addition to the 0.8% tracking systematic, material interactions of different hadron species impose a further systematic error of 1.4% for pions and 1.1% for kaons (taken from the previously-linked Twiki page). For the modes there are two extra tracks, one kaon and one pion, the systematic error is calculated as

$$\sigma = \sqrt{0.8^2 + 1.1^2} + \sqrt{0.8^2 + 1.4^2} = 3.0\%\,, \tag{4.14}$$

conservatively assuming full correlation between the tracks.

### 4.5.11 Total Systematic Uncertainty

A summary of all the systematic uncertainties which apply to each measurement, along with the total combined systematic uncertainties, are given in Tables 4.24 and 4.25.

| Systematic | Section | $B^0 \to \overline{D}^0 \mu^+ \mu^-$ | $B^0 \to \overline{D}^0 J/\psi$ | $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$ | $B_s^0 \to \overline{D}^0 J/\psi$ |
|---|---|---|---|---|---|
| $\mathcal{B}$ (Normalisation Mode) | 4.5.1 | 3.9% | | | |
| $\mathcal{B}(D^0 \to K^\pm \pi^\mp)$ | 4.5.1 | 1.0% | | | |
| $f_s/f_d$ | 4.5.1 | N/A | | 7.8% | |
| Normalisation Mode Statistics | 4.5.2 | 0.20% | | | |
| Simulation: Statistics | 4.5.3 | 2.1% | 2.6% | 3.1% | 3.8% |
| Simulation: Differences with data | 4.5.4 | 7.7% | | 7.7% | |
| Simulation: Other Years | 4.5.5 | 5.0% | | | |
| Simulation: PID Correction | 4.5.6 | 0.22% | 1.2% | 3.1% | 0.2% |
| Simulation: Trigger Correction | 4.5.7 | 4.1% | 6.8% | 4.5% | 7.0% |
| Fit: Bkg Shape (Sig) | 4.5.8 | 0.019% | 0.012% | 0.19% | 0.16% |
| Fit: Bkg Shape (Norm) | 4.5.8 | 0.019% | | | |
| Fit: Fixed Parameters (Signal) | 4.5.9 | 0.25% | 0.19% | 0.067% | 0.044% |
| Fit: Fixed Parameters (Charmless) | 4.5.9 | N/A | 3.8% | N/A | |
| Fit: Fixed Parameters (Normalisation) | 4.5.9 | 0.26% | | | |
| Total | - | 11.2% | 13.0% | 14.2% | 15.0% |

**Table 4.24:** Systematic uncertainties for the $B^0 \to \overline{D}^0 \mu^+ \mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$, and $B_s^0 \to \overline{D}^0 J/\psi$ modes. All values are given as errors relative to the measured branching ratio. N/A indicates the relevant systematic does not apply to that measurement.

| Systematic | Section | $B^+ \to D_s^+\mu^+\mu^-$ | $B^+ \to D_s^+ J/\psi$ | $B_c^+ \to D_s^+\mu^+\mu^-$ | $B_c^+ \to D_s^+ J/\psi$ |
|---|---|---|---|---|---|
| $\mathcal{B}$(Normalisation Mode) | 4.5.1 | 3.0% | | 7.9% | 3.0% |
| $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$ | 4.5.1 | 3.1% | | N/A | 3.1% |
| Normalisation Mode Statistics | 4.5.2 | 0.09% | | 9.3% | 0.09% |
| Simulation: Statistics | 4.5.3 | 2.8% | 3.9% | 4.7% | 2.5% |
| Simulation: Differences with data | 4.5.4 | 3.4% | 5.0% | 2.7% | 3.1% |
| Simulation: Other Years | 4.5.5 | | 2.7% | | |
| Simulation: PID Correction | 4.5.6 | 0.49% | 1.48% | 0.82% | 1.15% |
| Simulation: Trigger Correction | 4.5.7 | 6.0% | 7.0% | 6.7% | 3.2% |
| Fit: Bkg Shape (Sig) | 4.5.8 | 0.31% | 0.51% | 0.42% | 0.049% |
| Fit: Bkg Shape (Norm) | 4.5.8 | 0.0084% | | 0.049% | 0.0084% |
| Fit: Fixed Parameters (Signal) | 4.5.9 | 0.31% | 0.45% | 1.36% | 1.2% |
| Fit: Fixed Parameters (Part. Reco.) | 4.5.9 | | N/A | | $< 0.1\%$ |
| Fit: Fixed Parameters (Normalisation) | 4.5.9 | 0.28% | | 1.2% | 0.28% |
| Tracking Eff. | 4.5.10 | 3.0% | | N/A | 3.0% |
| Total | - | 9.5% | 11.2% | 15.4% | 8.0% |

**Table 4.25:** Systematic uncertainties for the $B^+ \to D_s^+\mu^+\mu^-$, $B^+ \to D_s^+ J/\psi$, $B_c^+ \to D_s^+\mu^+\mu^-$, and $B_c^+ \to D_s^+ J/\psi$ modes. All values are given as errors relative to the measured branching ratio. N/A indicates the relevant systematic does not apply to that measurement.

## 4.6 Results

### 4.6.1 Calculation of branching ratios

The branching ratios for the various decays are calculated using the ratio of yields and efficiencies between the signal and normalisation modes, as given by the following formulae:

$$\mathcal{B}\left(B^0 \to \overline{D}^0 \mu^+ \mu^-\right) = R_{\cdot} \frac{\mathcal{B}\left(J/\psi \to \mu^+ \mu^-\right) \mathcal{B}\left(K^{*0} \to K\pi\right)}{\mathcal{B}\left(\overline{D}^0 \to K^+ \pi^-\right)} \cdot \mathcal{B}\left(B^0 \to J/\psi\, K^{*0}\right), \tag{4.15}$$

$$\mathcal{B}\left(B^0 \to \overline{D}^0 J/\psi\right) = R_{\cdot} \frac{\mathcal{B}\left(K^{*0} \to K\pi\right)}{\mathcal{B}\left(\overline{D}^0 \to K^+ \pi^-\right)} \cdot \mathcal{B}\left(B^0 \to J/\psi\, K^{*0}\right), \tag{4.16}$$

$$\mathcal{B}\left(B_s^0 \to \overline{D}^0 \mu^+ \mu^-\right) = \frac{f_d}{f_s} \cdot R_{\cdot} \frac{\mathcal{B}\left(J/\psi \to \mu^+ \mu^-\right) \mathcal{B}\left(K^{*0} \to K\pi\right)}{\mathcal{B}\left(\overline{D}^0 \to K^+ \pi^-\right)} \cdot \mathcal{B}\left(B^0 \to J/\psi\, K^{*0}\right), \tag{4.17}$$

$$\mathcal{B}\left(B_s^0 \to \overline{D}^0 J/\psi\right) = \frac{f_d}{f_s} \cdot R_{\cdot} \frac{\mathcal{B}\left(K^{*0} \to K\pi\right)}{\mathcal{B}\left(\overline{D}^0 \to K^+ \pi^-\right)} \cdot \mathcal{B}\left(B^0 \to J/\psi\, K^{*0}\right), \tag{4.18}$$

$$\mathcal{B}\left(B^+ \to D_s^+ \mu^+ \mu^-\right) = R_{\cdot} \frac{\mathcal{B}\left(J/\psi \to \mu^+ \mu^-\right)}{\mathcal{B}\left(D_s^+ \to K^+ K^- \pi^+\right)} \cdot \mathcal{B}\left(B^+ \to J/\psi\, K^+\right), \tag{4.19}$$

$$\mathcal{B}\left(B^+ \to D_s^+ J/\psi\right) = R_{\cdot} \frac{1}{\mathcal{B}\left(D_s^+ \to K^+ K^- \pi^+\right)} \cdot \mathcal{B}\left(B^+ \to J/\psi\, K^+\right), \tag{4.20}$$

$$\mathcal{B}\left(B_c^+ \to D_s^+ \mu^+ \mu^-\right) = R_{\cdot} \mathcal{B}\left(J/\psi \to \mu^+ \mu^-\right) \cdot \mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right), \tag{4.21}$$

and

$$\frac{f_c}{f_u} \cdot \mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right) = R_{\cdot} \frac{1}{\mathcal{B}\left(D_s^+ \to K^+ K^- \pi^+\right)} \cdot \mathcal{B}\left(B^+ \to J/\psi\, K^+\right), \tag{4.22}$$

where the fragmentation fraction $f_c/f_u$ is the ratio of production rates for $B_c^+$ and $B^+$ mesons (as previously defined, $f_x$ is the probability that a $b$ quark will hadronize into a $B_x$ meson), and $\mathcal{B}(K^{*0} \to K\pi) = 2/3$. $R$ is defined by

$$R = \frac{N_{\text{sig}}}{N_{\text{norm}}} . \frac{\varepsilon_{\text{norm}}}{\varepsilon_{\text{sig}}} . C_{trig}^{\text{sig,norm}} , \tag{4.23}$$

where $N_x$ refers to the number of observed events for the signal or normalisation mode, $\varepsilon_x$ is the corresponding total efficiency, and $C_{trig}^{\text{sig,norm}}$ is the correction factor for the difference in trigger efficiencies between simulation and data for the signal and normalisation mode pair, given in Subsection 4.2.2.

### 4.6.2 Setting Upper Limits

As no signal is observed for any of the search modes, upper limits are set using the $\text{CL}_\text{S}$ method [65]. The $\text{CL}_\text{S}$ method tests the p-value

$$\text{CL}_S = \frac{p_{S+B}}{1 - p_B} , \tag{4.24}$$

where $p_{S+B}$ is the p-value of the signal and background (alternative) hypothesis and $p_B$ the p-value of the background only (null) hypotheses. The value of $\text{CL}_\text{S}$ is designed to be penalised when there is low sensitivity to the signal, as in such a case $1 - p_B$ becomes small, resulting in an increased limit compared to a case where only $p_{S+B}$ is used.

In the $\text{CL}_\text{S}$ method, a scan through branching ratios (or ratio of branching ratios) is performed, and the branching ratio corresponding to a $\text{CL}_\text{S}$ value of 0.05 is taken as the upper limit at the 95% confidence level. The RooStats (part of the Root framework) implementation of the $\text{CL}_\text{S}$ method was used [66]. The test statistic used to calculate $p_{S+B}$ and $p_B$ is based on the profile likelihood ratio. In a profile likelihood, the parameter of interest (in this case, the $B \to D\mu^+\mu^-$ branching ratio) takes a fixed value, and the likelihood with respect to the remaining "nuisance" parameters (*i.e.* the other fit parameters) is maximised. This "profiling" out of the nuisance parameters is convenient as it means the likelihood is maximised over the space of all possible systematic uncertainty values. From there, Wilks'

theorem [67] is used to calculate the p-values in Eq. (4.24). Note: the profile likelihood method only concerns the systematic uncertainties related to the free parameters of the fit. At this time the additional systematic uncertainties presented in Section 4.5 are not included. Upper limits with full systematic treatment will be presented in an upcoming publication [68]. Initial studies suggest increases in the upper limits of at most 6% can be expected.

The variation in $CL_S$ with branching ratio for the $B^0 \to \overline{D}^0 \mu^+ \mu^-$, $B^0 \to \overline{D}^0 J/\psi$, $B_s^0 \to \overline{D}^0 \mu^+ \mu^-$, $B_s^0 \to \overline{D}^0 J/\psi$, $B^+ \to D_s^+ \mu^+ \mu^-$, and $B^+ \to D_s^+ J/\psi$ decays is shown in Figs 4.28 and 4.29. The dotted lines give $CL_S$ taking into account the signal component of the mass fit, and the dashed line the variation only considering the background component(s) of the mass fit. With no observed signal, the two should overlap, however, as can be seen in Fig. 4.28b, the dotted line deviates from the dashed when a small signal is observed.

For $B_c^+ \to D_s^+ \mu^+ \mu^-$, the CLs method is run on the ratio $\mathcal{B}(B_c^+ \to D_s^+ \mu^+ \mu^-)$ to $\mathcal{B}(B_c^+ \to D_s^+ J/\psi)$ rather than just the $B_c^+ \to D_s^+ \mu^+ \mu^-$ branching ratio, as shown in Fig. 4.30. This is because the measured error on $\mathcal{B}(B_c^+ \to D_s^+ J/\psi)$ is quite large, and measuring the upper limit on $\mathcal{B}(B_c^+ \to D_s^+ \mu^+ \mu^-)$ as a function $\mathcal{B}(B_c^+ \to D_s^+ J/\psi)$ allows the limit to be easily updated in the future. The value of $\mathcal{B}(B_c^+ \to D_s^+ J/\psi)$ measured in Subsection 4.6.3 is used to calculate the final $\mathcal{B}(B_c^+ \to D_s^+ \mu^+ \mu^-)$ limit in this thesis.

The upper limits set for all modes are given in Table 4.26. As previously summarised in Section 3.6, only one upper limit had previously been set, $\mathcal{B}\left(B^0 \to \overline{D}^0 J/\psi\right) < 1.3 \times 10^{-5}$ [51], and therefore the limits set for all modes are world leading.

| Decay | Upper Limit |
|---|---|
| $\mathcal{B}\left(B^0 \to \overline{D}^0\mu^+\mu^-\right)$ | $< 6.1 \times 10^{-8}$ |
| $\mathcal{B}\left(B^0 \to \overline{D}^0 J/\psi\right)$ | $< 1.4 \times 10^{-6}$ |
| $\mathcal{B}\left(B_s^0 \to \overline{D}^0\mu^+\mu^-\right)$ | $< 1.8 \times 10^{-7}$ |
| $\mathcal{B}\left(B_s^0 \to \overline{D}^0 J/\psi\right)$ | $< 1.4 \times 10^{-6}$ |
| $\mathcal{B}\left(B^+ \to D_s^+\mu^+\mu^-\right)$ | $< 3.2 \times 10^{-8}$ |
| $\mathcal{B}\left(B^+ \to D_s^+ J/\psi\right)$ | $< 3.5 \times 10^{-7}$ |
| $\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+\mu^+\mu^-\right)$ | $< 9.6 \times 10^{-8}$ |

**Table 4.26:** The upper limits set with the CLs method for each mode at the 95% confidence level.

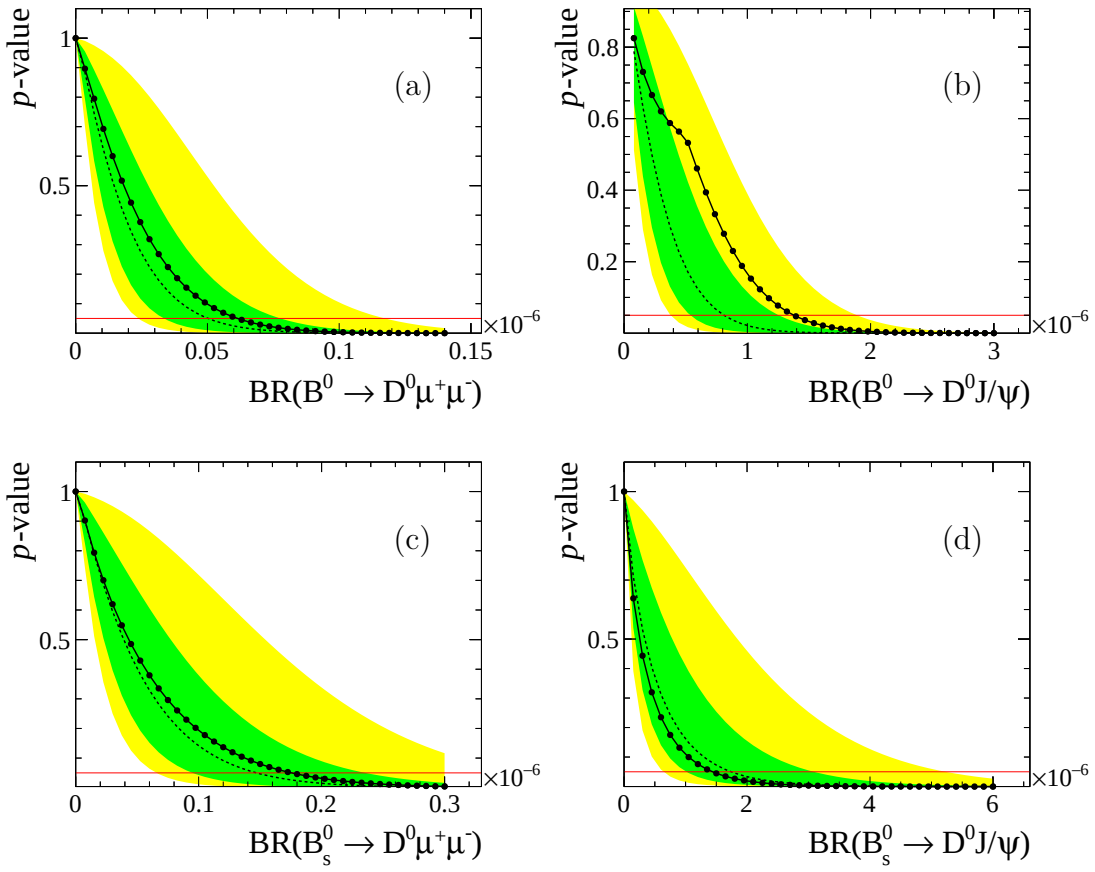



**Figure 4.28:** The variation in $p$-value with branching ratio for (a) $B^0 \to \overline{D}^0\mu^+\mu^-$, (b) $B^0 \to \overline{D}^0 J/\psi$, (c) $B_s^0 \to \overline{D}^0\mu^+\mu^-$, and (d) $B_s^0 \to \overline{D}^0 J/\psi$ decays. The dotted lines show the variation taking into account the signal component, and the dashed lines the variation only considering the background component. The green and yellow filled areas correspond to $1\sigma$ and $2\sigma$ from the background-only hypothesis. The upper limits are taken to be the values where the dotted lines cross the red line, corresponding to the 95% confidence level, and are given in Table 4.26.

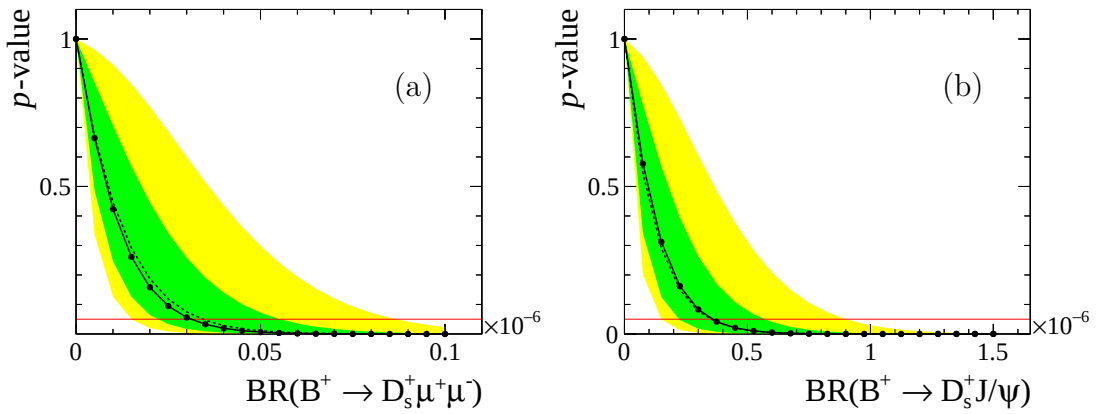



**Figure 4.29:** The variation in $p$-value with branching ratio for (a) $B^+ \to D_s^+ \mu^+ \mu^-$ and (b) $B^+ \to D_s^+ J/\psi$ decays. The dotted lines show the variation taking into account the signal component, and the dashed lines the variation only considering the background component. The green and yellow filled areas correspond to $1\sigma$ and $2\sigma$ from the background-only hypothesis. The upper limits are taken to be the values where the dotted lines cross the red line, corresponding to the 95% confidence level, and are given in Table 4.26.

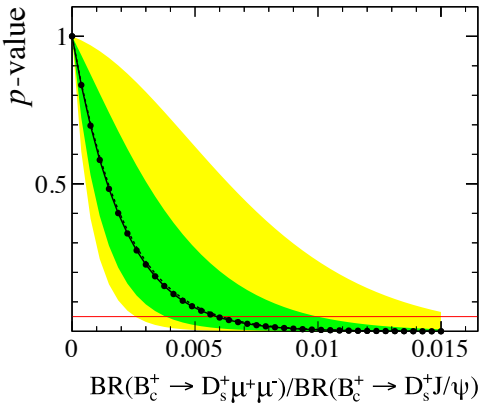



**Figure 4.30:** The variation in $p$-value with the ratio of $\mathcal{B}\left(B_c^+ \to D_s^+ \mu^+ \mu^-\right)$ to $\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right)$. The dotted line shows the variation taking into account the signal component, and the dashed line the variation only considering the background component. The green and yellow filled areas correspond to $1\sigma$ and $2\sigma$ from the background-only hypothesis. The upper limit is taken to be the value where the dotted line crosses the red line, corresponding to the 95% confidence level, and is found to be $\mathcal{B}\left(B_c^+ \to D_s^+ \mu^+ \mu^-\right) / \mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right) < 0.0059$. The product of this limit and the value of $\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right)$ measured in this thesis is given in Table 4.26.

### 4.6.3 Update of the $B_c^+ \to D_s^+ J/\psi$ Branching Ratio

Using (4.22), the branching ratio for the $B_c^+ \to D_s^+ J/\psi$ decay is measured to be

$$\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right) = (1.63 \pm 0.15 \pm 0.13) \times 10^{-5}\,,$$

where the first error is statistical and the second is the systematic uncertainty given in Table 4.25.

This new measurement is compared to previous results. Due to the choice of normalisation mode, the PDG value [28],

$$\frac{\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right)}{\mathcal{B}\left(B_c^+ \to J/\psi\,\pi^+\right)} = 3.1 \pm 0.5,$$

must be multiplied by [53],

$$\frac{f_c}{f_u}.\frac{\mathcal{B}\left(B_c^+ \to J/\psi\,\pi^+\right)}{\mathcal{B}\left(B^+ \to J/\psi\,K^+\right)} = (0.638 \pm 0.02)\,\%,$$

to give a value comparable to the one measured of

$$\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right) = (1.98 \pm 0.32)\,\% \times \mathcal{B}\left(B^+ \to J/\psi\,K^+\right) = (2.00 \pm 0.33) \times 10^{-5}\,.$$

It can therefore be seen that value measured in this thesis is consistent with the PDG average, and has a smaller uncertainty.

### 4.6.4 Measurement of $B_c^+ \to D_s^{*+} J/\psi$ decays

Two measurements relating to $B_c^+ \to D_s^{*+} J/\psi$ decays are also made by the fit shown in Fig. 4.24b. The ratio of $B_c^+ \to D_s^{*+} J/\psi$ to $B_c^+ \to D_s^+ J/\psi$ decays is measured to be $1.9 \pm 0.2$, and the ratio of $B_c^+ \to D_s^{*+} J/\psi$ decays with $\mathcal{A}^{\pm\pm}$ helicity amplitudes compared to the total is measured to be

$$f_{\mathcal{A}^{\pm\pm}} = \frac{N_{\mathcal{A}^{\pm\pm}}}{N_{\mathcal{A}^{\pm\pm}} + N_{\mathcal{A}^{00}}} = 50 \pm 10\,\%\,. \tag{4.25}$$

The errors quoted on these measurements are heavily statistics dominated and no systematic errors are quoted. Although the measured value $f_{\mathcal{A}^{\pm\pm}}$ is below the naive expectation of 2/3, $f_{\mathcal{A}^{\pm\pm}}$ and the measured ratio of $B_c^+ \to D_s^{*+} J/\psi$ to $B_c^+ \to D_s^+ J/\psi$ decays are consistent with the previous LHCb results of $52 \pm 20$ % and $2.27 \pm 0.57$ respectively [18].

# 5
# The TORCH Detector

## Contents



## 5.1 Motivation

The ability to perform accurate particle identification (PID) is a cornerstone of the design of LHCb. As detailed in Chapter 2, currently PID is performed by a pair of RICH detectors [3]. These provide strong pion-kaon separation over the momentum range 2 - 100 GeV/$c$, however they suffer at lower momenta, as shown in Fig. 5.1 [11]; there is no positive PID for kaons below $\sim 10$ GeV/$c$, and the downstream RICH1 detector is operated in "veto-mode", where the absense of Cherenkov light from a particle is used to veto a particular PID hypothesis. Additionally, proton separation efficiency rapidly falls below a value of 20 GeV/$c$.

The TORCH (Time Of interally Reflected CHerenkov light) detector is a proposed upgrade to LHCb, intended to provide complementary PID to the RICH detectors over the momentum range 2 - 10 GeV/$c$ [69].

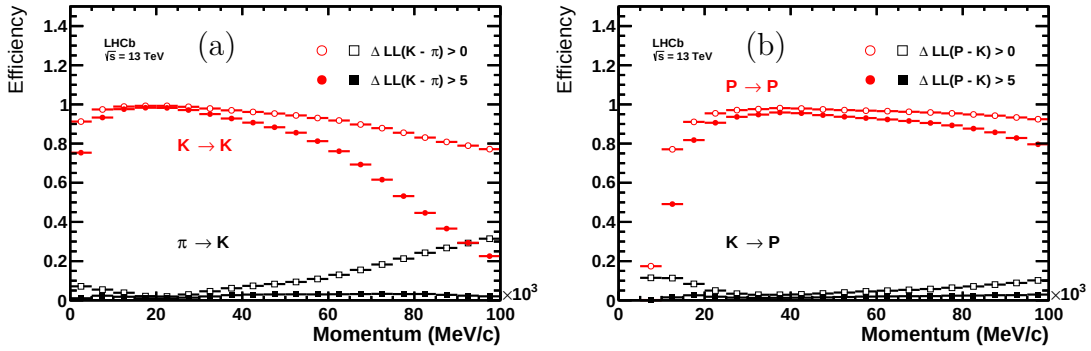



**Figure 5.1:** The (a) kaon and (b) proton PID performance of LHCb in 2015 (i.e. Run 2) as a function of momentum for two cuts on the decision-making variable (likelihood difference). Taken from Refs [11] and [70].

## 5.2 The Physics behind RICH and TORCH

Both the RICH detectors and TORCH use Cherenkov radiation to perform PID. When a charged particle passes through a dielectric medium, the medium is electrically polarised. In the case of particles with a velocity less than the phase velocity of light in that medium, the medium relaxes back faster than the particle is travelling. However, particles with a velocity greater than the phase velocity of light in the medium are decelerated by polarisation effects. The light emitted by the molecules returning to their ground state is refered to as Cherenkov radiation, named after the physicist who discovered it [71,72]. For a given refractive index, this light is emitted at a single angle, forming a coherent shock wave in the wake of the particle, the equivalent of a "sonic boom" for light. The angle of emission is given by

$$\cos\left(\theta_c\right) = \frac{1}{n\beta}, \tag{5.1}$$

where $\theta_c$ is called the Cherenkov angle, $n$ is the phase refractive index of the medium, and $\beta$ is the particle velocity relative to the speed of light.

Cherenkov radiation is prompt, and photons are emitted in a continuous spectrum of wavelengths. Although wavelengths in both the optical and ultraviolet regions are emitted, typically impurities in the radiator are absorptive in the ultraviolet region, and thus optical Cherenkov photons are usually more practical to be used for detectors [73]. The spectrum of wavelengths emitted leads to a complication when leveraging Cherenkov light for detector purposes - the refractive index of transmitting materials is wavelength dependent. This means that, even though the Cherenkov angle is the same for photons of the same wavelength, a range of angles are emitted by a single particle. The ability to account for this "chromatic dispersion" effect heavily influenced the design of TORCH.

The number of photons emitted by a particle traversing a medium depends on the particle's charge and how far it travels. This is captured by the Frank-Tamm relation [74],

$$\frac{d^2N}{dEdx} = \frac{\alpha}{\hbar c} Z^2 \left(1 - \frac{1}{n^2\beta^2}\right), \tag{5.2}$$

where $N$ is the number of photons, $E$ the photon energy, $Z$ the particle charge, and $\alpha$ the fine structure constant. Using Eq. (5.1) and integrating over the distance travelled, $x$, the relation can be written in terms of the Cherenkov angle as

$$\frac{dN}{dE} = \frac{\alpha}{\hbar c} Z^2 L \sin^2\theta_c, \tag{5.3}$$

where $L$ the radiator length.

## 5.3 The Design of TORCH

### 5.3.1 The TORCH Concept

The TORCH detector measures Time of Flight (ToF) to perform PID [75]. Over a fixed distance $x$, the time of flight, $t$, of a particle is given by

$$t = \frac{x}{c}\sqrt{1 + \left(\frac{mc}{p}\right)^2}, \tag{5.4}$$

where $p$ and $m$ are the particle momentum and mass respectively. Using (5.4), the difference in the time of flight of pions, kaons, and protons is calculated and is

shown in Fig. 5.2 over a 9.5 m flight distance. Clearly, with sufficiently accurate timing measurements, these species can be separated. For $3\sigma$ separation of pions and kaons over this distance, a $\sim 15$ ps track-time resolution is required, and this is what TORCH aims to achieve.

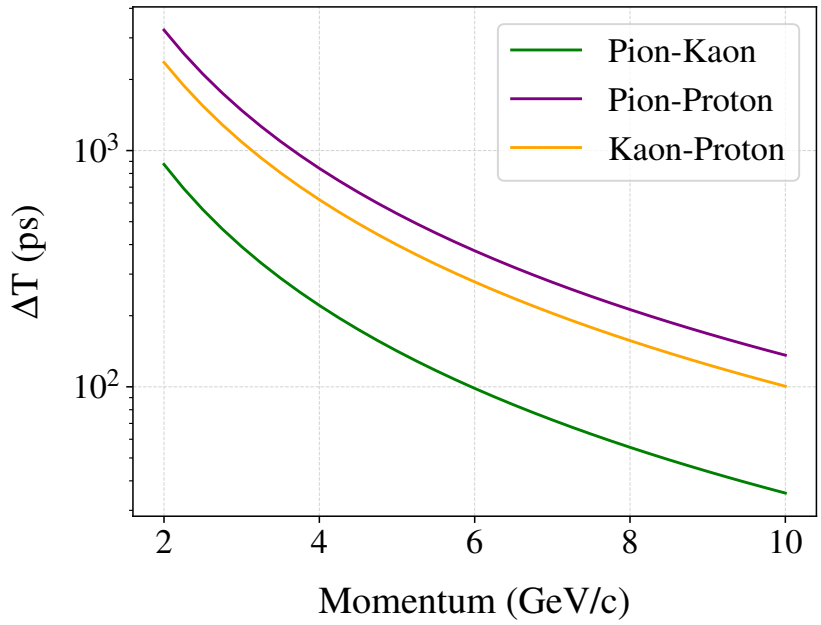



**Figure 5.2:** The difference in the time of flight ($\Delta$T) over a 9.5 m flight path for pions, kaons, and protons over the momentum range 2 - 10 GeV/$c$.

TORCH is a DIRC-style detector, employing the same reconstruction principles as the BaBar DIRC [76] and Belle II TOP [77, 78] detectors. Consisting of a quartz radiator plate and focusing optics, shown in Fig. 5.3, TORCH exploits Cherenkov radiation emitted by particles traversing the radiator plate. These photons are trapped by total internal reflection and travel to the top of the plate (or bottom, depending on orientation) where they are focused onto an array of Micro-Channel Plate PhotoMultiplier Tube (MCP-PMT) detectors. The focusing optics are designed to map the photon angle in the $y$-$z$-plane (defined in Fig. 5.3) to position on the detector plane. For convenience, an additional axis is defined, the $y'$-axis, which is tilted by 36° from the $y$ axis in the $y$-$z$-plane to run parallel to the detector plane.

The mapping of photon angle to position is essential to the TORCH operating principle. Combining the angle in the $y$-$z$-plane and the angle in the $x$-$z$-plane

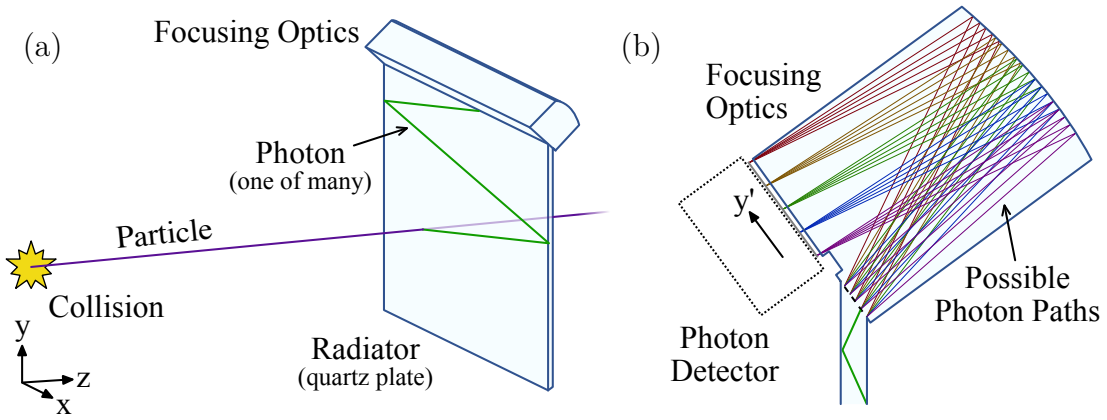



**Figure 5.3:** Schematics demonstrating the TORCH principle. (a) Particles passing through the radiator plate produce Cherenkov photons which are trapped by total internal reflection and travel to the focusing optics. (b) The focusing optics map photon angle to $y'$-position on the photon detector plane, allowing the Cherenkov angle to be determined.

(which can be determined geometrically) with knowledge of the direction of the emitting particle, the Cherenkov angle can be determined. Given the particle momentum, and given a mass hypothesis, the photon wavelength can be found, meaning chromatic dispersion can be accounted for when calculating the time it takes for photons to reach the detector plane after emission, the so called "time of propagation". This must be subtracted to find the time at which the emitting particle was incident upon TORCH, the desired "time of flight".

To determine the time of propagation of a single photon with $\sim 50\,\text{ps}$ precision[1], the angles in the $x$-$z$- and $y$-$z$-planes must be measured with roughly 1 mrad precision [75], providing restrictions on the granularity of the photon detectors used. The width of the radiator plate provides a long lever arm, and so in the $x$-direction a pixelisation of roughly 6 mm pitch is needed. This requirement is far more stringent in the $y'$-direction, requiring $\sim 0.4\,\text{mm}$ pitch or better. As will be shown later, a dedicated programme was undertaken to produce MCP-PMTs which met these requirements.

[1]This requirement is justified in Section 5.4.

### 5.3.2 Optics

Ideally TORCH would be formed of a single piece of quartz, however practicalities dictate it be divided into two primary pieces: a radiator plate and a set of focusing optics. Further subdivisions can be made; for example, in the final LHCb detector, multiple modules will be used (see Section 5.4), however the primary distinction between radiator plate and focusing optics remains valid.

The radiator plate is a large sheet of fused silica (quartz), 10 mm thick, within which Cherenkov photons are generated by traversing charged particles. The final TORCH design will be formed of multiple $660 \times 2500\,\mathrm{mm}^2$ plates (width $\times$ length), however plates with smaller $x$ and $y$ dimensions have been used in prototypes. The top edge of the plate which connects to the focusing optics has a 36° bevel, as seen in Fig. 5.3b.

The focusing optics consist of a block with a mirrored back spherical surface to reflect photons onto the detector plane. As previously mentioned, the block maps angle to position in the $y'$-direction, allowing the differing Cherenkov angles over the spread of wavelengths to be determined.

### 5.3.3 TORCH MCPs

MCP-PMTs are employed by TORCH due to their fast response, potential for high spacial resolution, and resistance to high magnetic fields [79–81]. The working principle of an MCP-PMT is shown in Fig. 5.4. When a photon strikes the photocathode, there is a chance it liberates an electron (the "quantum efficiency"). This electron is attracted by positive high voltage into a micro-channel plate, where it triggers a charge avalanche, following into a second plane of plates in a so-called "chevron arrangement". The resulting avalanche is collected at the anode. Because of the compact design, MCP-PMTs typically have a very fast response (order of tens of ps). The gain of the MCP-PMT depends on the high-voltage bias across the plates.

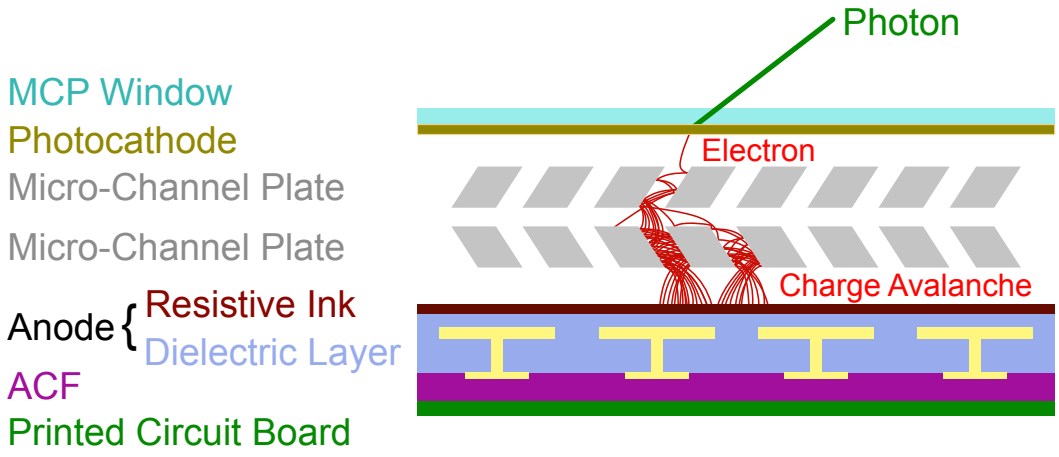



**Figure 5.4:** A schematic of a TORCH MCP-PMT. The yellow bars within the dielectric layer are the anode pads, with vias which are connected by Anisotropic Conductive Film (ACF) to an external printed circuit board. Note elements within the diagram are not to scale.

To establish an MCP-PMT suitable for TORCH, a development program has been carried out in collaboration with Photek UK (Ltd.) [81, 82]. Three TORCH-specific requirements were specified for the MCP-PMTs:

1. to endure $5\,\mathrm{C/cm^2}$ of accumulated anode charge without a significant loss of sensitivity;

2. to acheive 6 mm and 0.4 mm spatial granularity ($6/\sqrt{12}$ mm and $0.4/\sqrt{12}$ mm resolutions) in the $x$ and $y'$ directions respectively;

3. to allow linearly close packing with an 88% active area in the $x$-axis.

The development programme was split into three phases; Phase-I aimed to deliver a tube to meet requirement (1), the Phase-II tube would meet requirement (2), and the Phase-III would be a final tube to meet all three conditions simultaneously [83, 84]. The completion of this programme delivered a square MCP-PMT design which employs a combination of novel approaches to meet the requirements, hereon referred to as the TORCH MCP-PMT.

As outlined above, the TORCH MCP-PMTs are of "chevron" type, meaning they contain two micro-channel plates stacked such that the pore angles are in opposite

directions (shown in Fig. 5.4). Secondary electrons in the charge avalanche can librerate ions from the MCP pores. Gas can also be desorbed from the pore surface, or residual gas in the detector vacuum can be ionised. Ions from either source are accelerated by the electric field and can bombard the photocathode, damaging it, and leading to a reduction in the tube quantum efficiency. As the ion production mechanisms are tied to the charge avalanche, they are most prevalent at the bottom of the second MCP; hence, the chevron pattern is desirable as it eliminates a direct path from the greatest source of production to the photocathode [85].

The chevron pattern alone is not sufficient to meet the high lifetime requirement of the TORCH tubes. To further suplement this, Atomic Layer Deposition (ALD) [86] is applied to the MCPs. The ALD process involves repeatedly layering a substrate (in this case, $Al_2O_3$) onto a surface (the MCPs) in layers of single atom thickness. This has two advantages. Firstly, the yield of secondary electrons is increased, allowing a lower voltage to produce the same gain. This reduces the chance of an ion being liberated as the energies of the produced secondary electrons are lower. Secondly, it acts as a sealant on the MCP pore surface, preventing the outgassing of contaminants [83, 85]. Both these effects combine to reduce the ion bombardment on the photocathode and improve the MCP-PMT lifetime.

To achieve the desired granularity, a novel anode design is employed. Shown in Fig. 5.4, this anode is a hybrid of a directly connected multi-anode detector and a charge-sharing imaging device. Charge from the avalanche is collected on a resistive ink layer, and is capacitively coupled to anode pads embedded in a dielectric layer. This indirect coupling allows charge to be disributed over several pads, resulting in charge sharing which can be exploited to improve the granularity beyond that implied by the number of pads themselves [81].

Anisotropic Conductive Film (ACF) is used to connect the MCP-PMT anode to an external Printed Circuit Board (PCB), which subsequently interfaces with the readout electronics. ACF is a $\sim 100\,\mu$m elastomer film with embedded wires. As the wires all have the same pitch, conduction is only allowed in a single direction, making

ACF ideal for the TORCH MCP-PMTs as it allows a high density of connections between the anode and PCB, an essential quality for the high granularity readout.

The final MCP-PMT has a physical area $59 \times 59\,\mathrm{mm}^2$ in $x \times y'$, with an active area of $53 \times 53\,\mathrm{mm}^2$ to meet the close packing requirement. The basic design employs $64 \times 64$ pixels, however, as a reduced granularity for the TORCH MCP-PMT is required in the $x$-direction, neighbouring sets of 8 channels are ganged together on the PCB to form a single channel [81]. This leads to an $8 \times 64$ ($x \times y'$) pixel layout in the final MCP-PMT, with charge-sharing increasing the $y'$ granularity to an effective 128 pixels.

### 5.3.4 Readout Electronics & DAQ

Custom electronics have been designed to read out the TORCH MCP-PMTs, employing the NINO and HPTDC chipsets and consisting of four electronic units [81,87–89]:

- the NINO board;
- the HPTDC board;
- the backplane;
- the readout board.

The NINO board performs amplification and signal discrimination. Employing the eponymous NINO chip [90,91], a time-over-threshold (TOT) measurement of the input charge is made, setting a user-defined threshold value. Signals from the NINO board are then digitized by the HPTDC board. Employing the HPTDC chip [92,93], the leading and trailing edges of the NINO signal are measured and stored in a buffer on an on-board FPGA [89]. The backplane then connects the HPTDC board to the readout board, through which the digitised signals are transmitted to a DAQ PC.

## 5.4 TORCH as an LHCb sub-detector

TORCH is proposed for the Upgrades Ib and II of the LHCb experiment [94]. As a subdetector system in LHCb, TORCH would be ideally located in front of RICH2, as shown in Fig. 5.5. This provides a good trade-off between having a sufficiently long path length over which to time the particles, and not suffering a large effect of multiple Coulomb scattering which would occur were TORCH placed behind the more material-heavy subdetectors. In this position, TORCH would be 9.5 m from the interaction point.

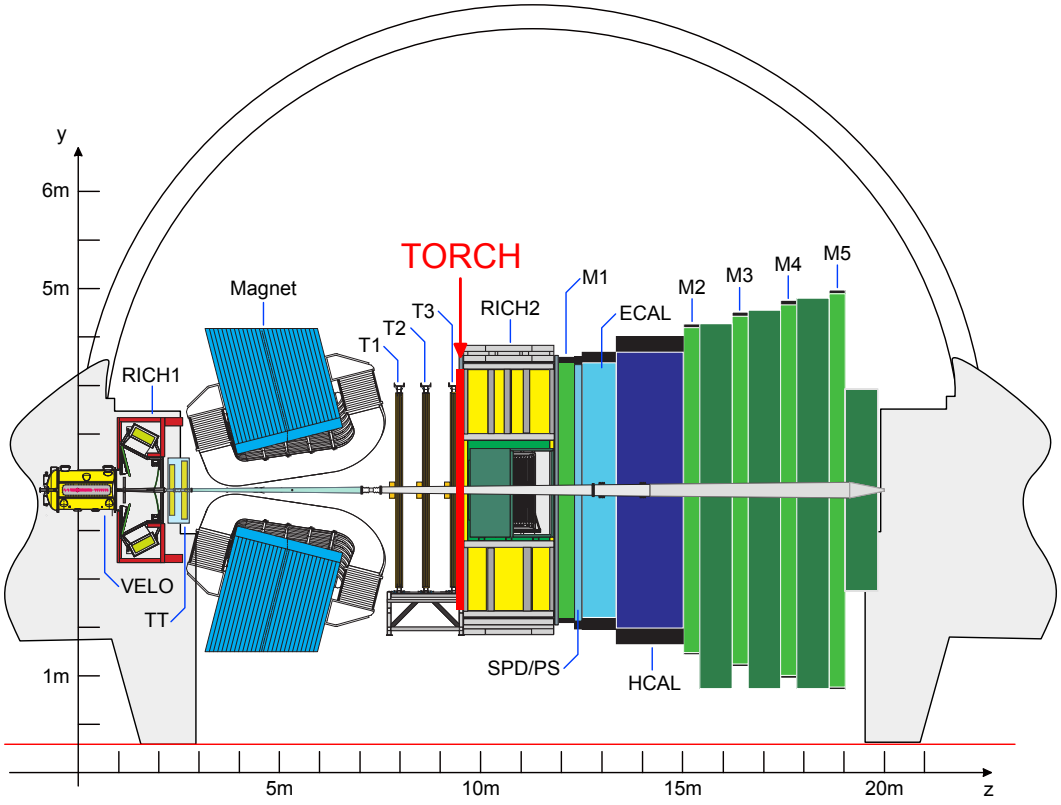



**Figure 5.5:** The proposed placement of TORCH within LHCb.

At a momentum of 10 GeV/$c$, the difference in time of flight between pions and kaons over a 9.5 m flight path is 35 ps, therefore to distinguish the two species with $3\sigma$ significance, a per-track timing resolution of approximately 15 ps is desired. With 30 detected photons per particle anticipated, this requires a single-photon timing resolution of 70 ps, which can be broken down in to a 50 ps contribution

from the MCP-PMT and readout electronics, and another 50 ps from the accuracy with which the time of propagation can be determined.

In LHCb, the full TORCH detector would be composed of eighteen $660 \times 2500 \times 10\,\text{mm}^3$ (width $\times$ height $\times$ depth) modules, arranged in a $9 \times 2$ grid with the focusing blocks and detectors along the top and bottom edges. The modular layout in LHCb is shown in Fig. 5.6, along with the density of charged paricles passing through TORCH, determined using the LHCb simulation [95]. The central modules are seen to receive the greatest flux of particles, with a large reduction at the sides. Additionally, the largest flux is seen to be around the vertical middle, furthest from the photon detectors. This highlights the importance of being able to accurately measure the time of arrival of photons emitted at the bottom of the plate, which subsequently must travel to the top outer edge to be detected.

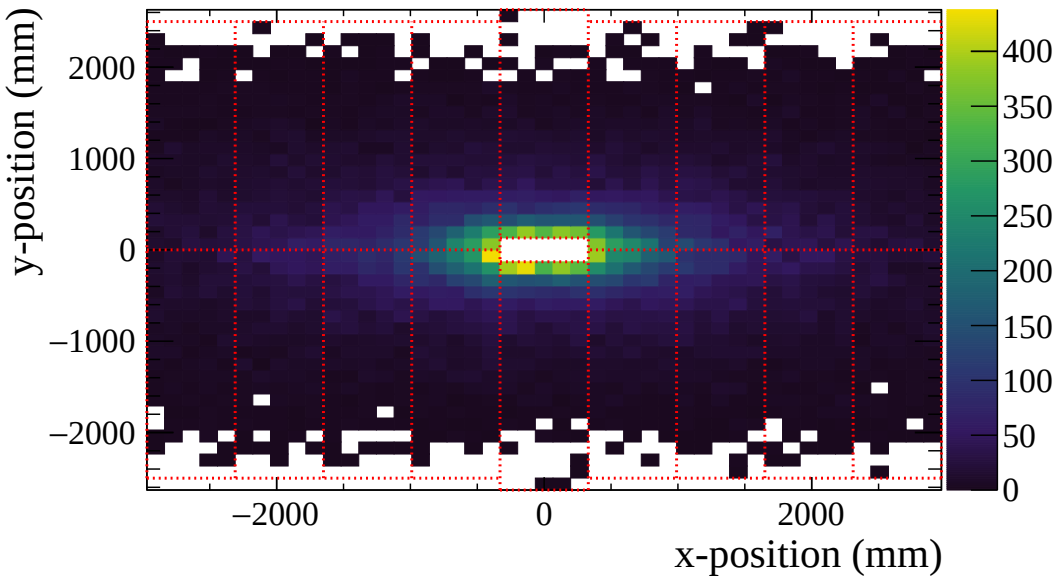



**Figure 5.6:** The density of charged particles passing through TORCH in the LHCb environment as a function of $x$-$y$ position, determined from simulation with 7 TeV per beam and a luminosity of $2 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$ (i.e. LHC Run 3 conditions). The red dotted lines show the outline of the 18 TORCH modules. The central plates are displaced 260 mm vertically apart to leave room for the beam pipe.

### 5.4.1 PID Performance

The expected PID performance of the full TORCH detector has been measured using the LHCb simulation. To achieve this, an algorithm for performing PID was developed[2]. Based on the algorithm used by the current LHCb RICH detectors [96], a pattern of photon hits in TORCH is generated for each incident particle and for each PID hypothesis being considered. These patterns are then converted into probability density functions. The detected hits are compared to each combination of PID hypotheses and the difference in log-likelihood calculated. The combination with the largest log-likelihood difference is taken to be the "correct" hypothesis.

Figure 5.7 shows the PID performance expected by applying the PID algorithm to Monte-Carlo data [97]. Excellent pion-kaon separation is seen for momenta up to 10 GeV/$c$, with the performance tailing off as the momentum grows and the difference in time of flight becomes small. For protons and kaons, strong separation continues to be seen to 20 GeV/$c$ and beyond. Both these plots show promising performance for TORCH as an LHCb subdetector and, at time of writing, work is ongoing to improve these efficiencies still further.

[2]While the author was involved in the development of the PID algorithm and made preliminary PID performance measurements, the work shown here is the result of successive iterations and is not the author's own.

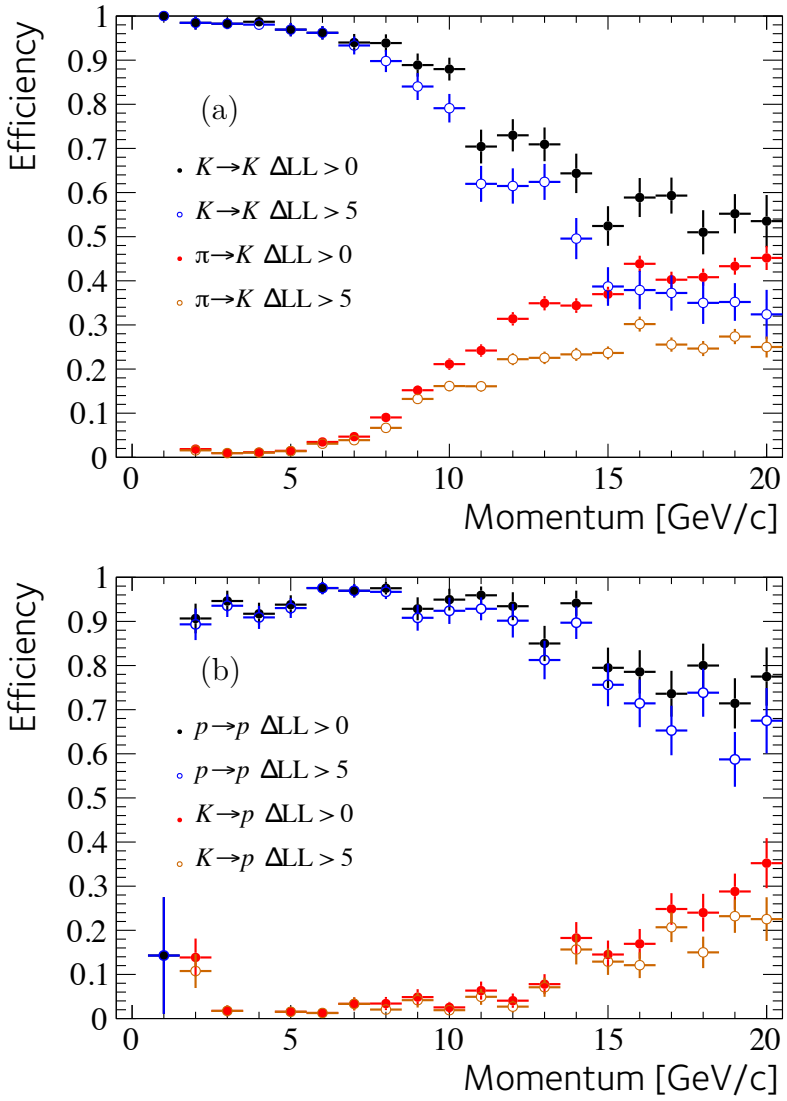



**Figure 5.7:** The simulated PID efficiencies and misidentification fractions of the TORCH detector in LHCb for (a) kaon-pion separation and (b) proton-kaon separation. The simulation was run with pile-up conditions corresponding to a luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Taken from Ref. [97].

# 6

# Test-beam Studies of TORCH Detector Prototypes

## Contents

## 6.1 Introduction

Throughout the TORCH project, test-beam campaigns have been carried out to test new developments and drive the design towards one which can meet the requirements for operation in LHCb. This chapter details the three most recent test-beam campaigns, referred to by the month in which they occurred:

- November 2017 ($30^{\text{th}}$ October to $20^{\text{th}}$ November 2017);
- June 2018 ($6^{\text{th}}$ to $27^{\text{th}}$ June 2018);
- October 2018 ($22^{\text{nd}}$ October to $12^{\text{th}}$ November 2018).

The goal of each test-beam run is to measure the single-photon time resolution achievable with the latest TORCH demonstrator, MCP-PMTs and readout electronics, as well as to investigate the photon counting efficiency. Presented here are the single-photon time resolutions measured in the November 2017 and June 2018 test-beam periods. Additionally, preliminary measurements of time

resolutions are presented for the October 2018 test-beam run, along with the photon counting efficiency compared to simulation.

Due to the many similarities between the test-beam runs, especially in terms of methodology, the three are presented together. Data from all three are used as appropriate to describe the data analysis procedures, before the single-photon time-resolution results from each are presented chronologically. In addition to being presented here, the results for the November 2017 and June 2018 test-beam campaigns have been published in Ref. [98], while preliminary results for the October 2018 test-beam campaign are published in Ref. [99]. Operational details and results for previous TORCH test-beam campaigns can be found in Refs [69, 100].

## 6.2 The TORCH Prototypes

### 6.2.1 MCP-PMTs

Three TORCH MCP-PMTs were used across the three test-beam periods, the properties of which are summarised in Table 6.1. The measured quantum efficiencies[1] are shown in Fig. 6.1 as a function of wavelength.

| Name | Serial Number | Pixelisation $(x, y')$ | Dielectric Thickness (mm) | PSF (mm) | Gain (electrons) |
|---|---|---|---|---|---|
| Z | A3170412 | $4 \times 64$ | 0.5 | 1.8 | $1.82 \times 10^6$ |
| A | A2170331 | $8 \times 64$ | 0.3 | 1.54 | $1.0 \times 10^6$ * |
| B | A3170517 | $8 \times 64$ | 0.5 | – | $1.0 \times 10^6$ * |

**Table 6.1:** Serial numbers and properties of the MCP-PMTs used in the test-beam campaigns. The quoted gains correspond to the values of MCP-PMT voltages used in the test-beam runs. PSF stands for Point Spread Function. An asterisk (*) denotes values taken from Photek data sheets which have not been otherwise measured. A dash (−) indicates the value has not been measured and is not available on the Photek data sheet. The MCP-PMT names A and B were selected to match Ref. [99], with the $4 \times 64$ being designated Z for "zero" as it preceded the other two.

[1] These measurements were made by colleagues at CERN, not by the author.

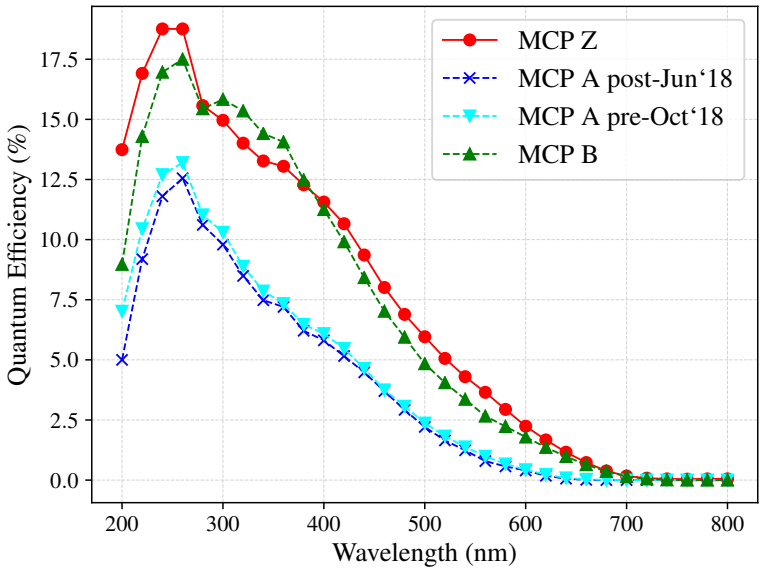



**Figure 6.1:** The measured quantum efficiencies of the MCP-PMTs used in the test-beam campaigns as a function of wavelength, measured at CERN [98]. Two measurements are displayed for MCP-PMT A, one taken just after the June 2018 beam test and the second measured on 9$^{\text{th}}$ October 2018, just before the October 2018 test-beam operation.

### 6.2.2 Mini-TORCH

Mini-TORCH is a small-scale prototype of TORCH, with a $120 \times 350 \times 10\,\text{mm}^3$ (width $\times$ height $\times$ depth) radiator plate. The focusing block is the same size as it would be for a full-size TORCH detector in the $y$-$z$-plane, but is scaled in the $x$ dimension to match the width of the reduced-size radiator plate. The optical components for Mini-TORCH are made from fused silica (Corning 7980 [69]) and are shown in Fig. 6.2. The width allows Mini-TORCH to be instrumented with two TORCH MCP-PMTs, however only a single tube was used for the two Mini-TORCH test-beam campaigns. In both cases, the MCP-PMT was mounted between one side edge and the centre of the focusing block. This layout is shown subsequently in Figs 6.8 and 6.9. In the November 2017 campaign MCP-PMT Z was instrumented, and for June 2018 MCP-PMT A was used. Due to the doubling of the coarse-direction granularity of tube A, new electronics were instrumented in June 2018, replacing those of November 2017. This required the implementation of new mounting mechanics. A 5 mm offset in the positive $y'$-direction (see Fig. 5.3) relative to the November 2017 test-beam was introduced as a result.

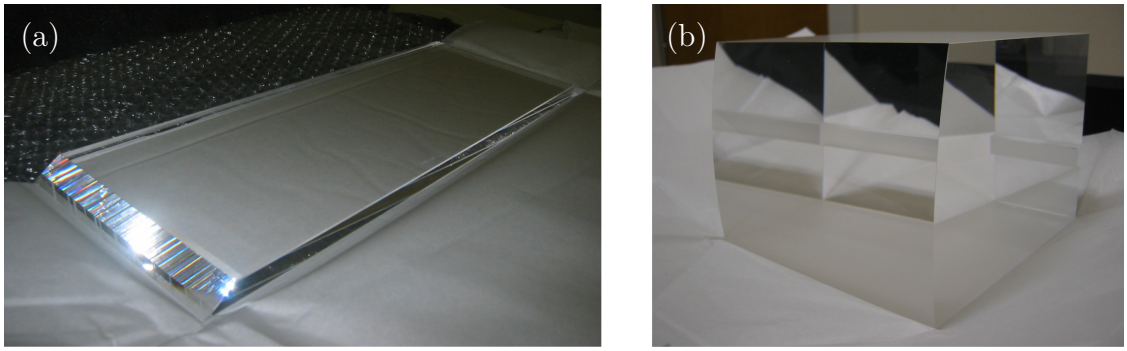



**Figure 6.2:** The (a) radiator plate and (b) focusing block for Mini-TORCH, taken from Ref. [69].

### 6.2.3 Proto-TORCH

Proto-TORCH is a large-scale demonstrator with a $660 \times 1250 \times 10\,\text{mm}^3$ (width $\times$ height $\times$ depth) radiator plate, chosen to be half the height and the full width of an LHCb module. With a width 5.5 times larger than Mini-TORCH, Proto-TORCH can be instrumented with up to eleven MCP-PMTs. In the October 2018 campaign, Proto-TORCH was instrumented with MCP-PMTs A and B, positioned beside each other at the positive-$x$ edge of the focusing block. This layout is shown later in Fig. 6.10.

## 6.3 The Test-beam Setup

### 6.3.1 Overview

The beamline and infrastructure for the three test-beam campaigns remained largely unaltered between each. Figure 6.3 shows the items of infrastructure as well as their relative positions along the beam line, with the labels corresponding to:

- a Device Under Test (DUT) - either Mini-TORCH (November 2017 and June 2018) or Proto-TORCH (October 2018);
- two threshold Cherenkov counters (C1 & C2, Subsection 6.3.3);
- a large scintillator counter ($T_b$, Subsection 6.3.4);
- two timing stations (T1 & T2, Subsection 6.3.5);

- a beam telescope (Subsection 6.3.6);
- a Trigger Logic Unit (TLU, Subsection 6.3.7).

Details of each item of infrastructure are given in the following sections.

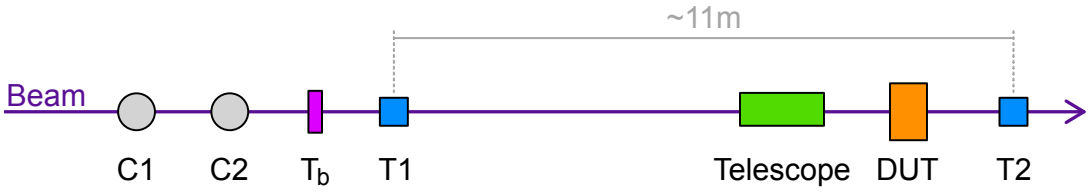



**Figure 6.3:** The test-beam infrastructure used in the three test-beam campaigns. C1 & C2 are threshold Cherenkov counters, $T_b$ a large scintillator, T1 & T2 timing stations, and DUT either Mini-TORCH or Proto-TORCH depending on the test-beam campaign. The position of each item of infrastructure relative to the others is shown, although not to scale.

### 6.3.2 The Proton Synchrotron T9 Beam Line

All three test-beam campaigns were carried out in the East Hall T9 facility at CERN, where proton/pion beams of up to 15 GeV/$c$ are provided by the Proton Synchrotron (PS) (see Fig. 2.1) [101]. The facility allows the momentum to be set by the user, allowing a range of momenta to be investigated. Additionally a set of collimators are provided which allow the range of momenta accepted around the nominal setting to be varied. Several targets are available; for the TORCH test-beams, the "hadron" aluminium target was used.

The beam arrives in spills, with approximately two per minute under regular running conditions. Each spill lasts for roughly 350 ms and typically contains $\mathcal{O}(10^5)$ particles, however this number depends on the momentum and collimator settings, as well as the target used [102].

### 6.3.3 Cherenkov Counters

The two Cherenkov counters, C1 and C2, consist of large barrels filled with $CO_2$ pressurised to 2.5 bar [99]. When a low-mass particle (e.g. a pion) passes through, it generates Cherenkov light which is detected by a single-channel PMT. Higher

mass particles (e.g. kaons and protons) are below the Cherenkov threshold and so no signal is seen. Thus the Cherenkov counters can be used for independent PID.

### 6.3.4 Large Scintillator $T_b$

$T_b$ consists of a large rectangular scintillator optically coupled to a single channel PMT. In the test-beams, signals from both C1 and C2 were required in coincidence with a signal from $T_b$. This was to reduce the effect of random noise creating a pion-like signal. Both C2 and $T_b$ can be seen in Fig. 6.4.

**Figure 6.4:** The C2 Cherenkov counter installed in the T9 beam test zone. $T_b$ can also be seen on the right within the purple holding mechanics; the scintillator is encased in a black velvet cover, and below extends the single-channel PMT.

### 6.3.5 Timing Stations

The timing stations T1 and T2 were used to provide a start time, $t_0$, for the DUT [99]. They consist of two parts:

- a timing finger;
- a pair of crossed scintillators.

The timing finger is composed of an $8 \times 8 \times 100\,\mathrm{mm}^3$ borosilicate bar optically coupled to a single-channel MCP-PMT (Photonis PP0365G), shown in Fig. 6.5. The timing finger is positioned with the long axis 49° from the beam axis so that Cherenkov light generated by passing particles is directed straight towards the MCP-PMT. Unlike in Ref. [100], the bar is not blackened; this increases photon yield and has negligible effect on the resolution of the timing signal as this is primarily influenced by the first-arriving photons which reach the MCP-PMT.

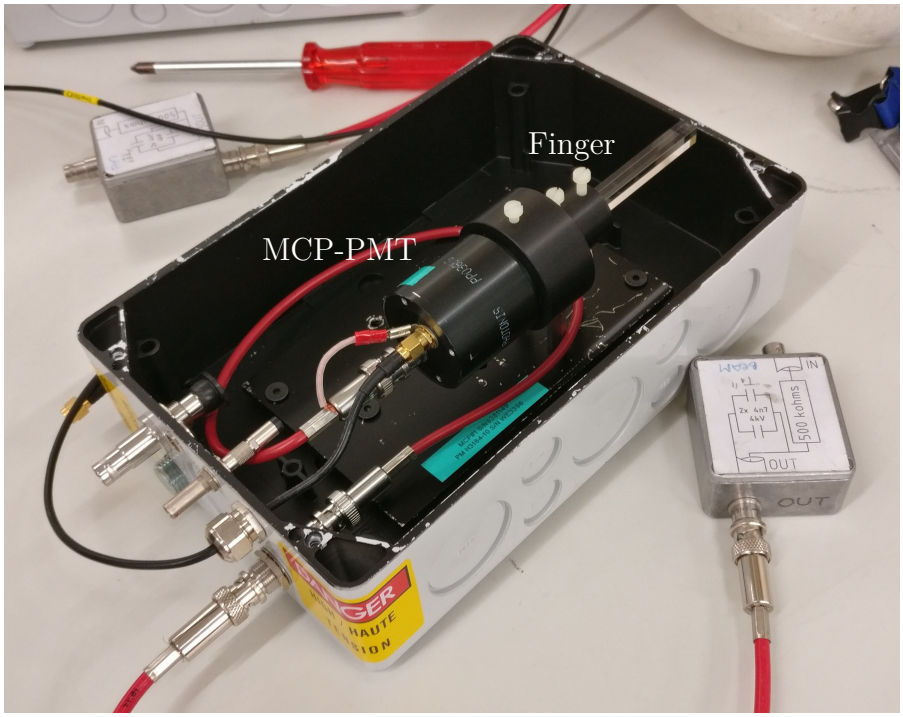



**Figure 6.5:** The timing element of one of the timing stations, consisting of an $8 \times 8 \times 100\,\mathrm{mm}^3$ borosilicate finger optically coupled to a single channel MCP-PMT (Photonis PP0365G). The finger and MCP-PMT are housed in a light-tight box.

The second part of the timing station consists of a pair of scintillator bars, crossed at right-angles to present an $8 \times 8\,\mathrm{mm}^2$ overlap perpendicular to the beam. Each bar is connected to a PMT. By requiring a signal from both PMTs in the trigger, the beam definition can be narrowed, and it is ensured that the triggered particles pass through the borosilicate finger.

In the test-beam setup, the two timing stations are positioned roughly 11 m apart, with one just downstream of the Cherenkov counters and the other behind the DUT, as shown in Fig. 6.3. Over this distance there is a significant difference in the ToF of protons and pions at 5 GeV/$c$, therefore by combining the signal from both timing stations, another independent source of PID can be performed [99].

Signals from the time reference stations are processed in two different ways. The first is with a commercial system produced by ORTEC[2], where the signals are digitized using 6.31 ps wide bins. In the second, the signals are fed into the TORCH readout electronics [100]. This has a pair of minor drawbacks: firstly, the MCP-PMT pixels that have been replaced by the reference station readout are effectively endered as dead channels, and second, the signals are then digitised by the HPTDC chip, which has a 100 ps bin width instead of 6.31 ps. However, this injection procedure has the significant benefit of removing the need to synchronise the TORCH data with those from the reference stations, as the latter signals are then present in the data-stream. For this reason this signal-injection method is used almost exclusively in the analysis presented in this Chapter.

### 6.3.6 Beam Telescope

An EUDET/AIDA pixel beam telescope [103] was installed in the beam-line, consisting of six Mimosa26 18.4 μm pitch sensors [104]. The original purpose of the telescope was to provide individual tracking information for particles passing through the DUT, however synchronization between the telescope and the TLU proved impossible. Instead, the beam profile was measured and the contribution statistically acounted for using simulation (see Subsection 6.8.5).

In the October 2018 analysis, due to issues with the telescope reconstruction software, the telescope data were unable to be reconstructed. To minimise the effect this had on the analysis, the beam conditions were set to match those used in June 2018, so that the beam definition measured previously could be used.

[2]Model 9327 from ORTEC, Oak Ridge, TN37831-0895, U.S.A.

### 6.3.7 Trigger Logic Unit

The clock and trigger was provided by an AIDA-2020 TLU [105]. The unit accepts up to six asynchronous input signals, and when pre-configured logical combinations of inputs are seen, a global trigger is issued. The TLU was used in conjunction with the EUDAQ2 framework [106].

### 6.3.8 Combining the Infrastructure

#### Maximising Rate

The infrastructure was combined and used in such a way as to optimise the data rate without harming the analysis potential. In particular, the Cherenkov counters and telescope were added beyond the test-beams detailed in Ref. [69], and proved key to achieving the increased data rate. First, through providing an independent source of PID to the timing stations, the Cherenkov counters allowed T1 to be removed from the trigger. Without the need for coincidence between the two timing stations, a large increase in rate was achieved, but with the trade-off of a larger beam profile. However, as this profile could be measured by the telescope, it could be accounted for in the later analysis.

#### Synchronization of Data

Data were taken in three independent streams:

- one from the DUT, containing detected photons and the time reference signals;
- one from the TLU, containing the Cherenov counter signals;
- one from the beam telescope, containing tracking information.

Before analysis could be performed, the data in the streams had to be synchronized. The TORCH and TLU data streams were synchronised by reconstructing and comparing the spill structure of the beam. Once spills reconstructed in each stream had been correctly matched to each other, the events within the spill were synchronised based on their timestamps.

To validate the TORCH-TLU synchronization, the ToF distribution produced by comparing the T1 and T2 timing signals (a distribution drawn entirely from TORCH data) was compared to the signal from the Cherenkov counters (recorded solely by the TLU). If an improper synchronization had been determined, there would be no correlation between the ToF distribution and the Cherenkov signal. Instead, a Cherenkov signal is observed for the pion peak and not the proton peak, indicating the correct synchronization has been found. Examples of such distributions are given later, in Subsection 6.7.3.

The telescope data stream was initially envisioned to be synchronised by matching the assigned event numbers with those recorded by the TLU. Unfortunately, it was found that the telescope was not recording all events, randomly missing roughly 1%. The event numbering was found to be continuous and independent in the telescope and TLU, meaning no relation between the two could be drawn, and due to the random distribution of missing events, a verifiable synchronization was not possible.

## 6.4 Simulating TORCH

A stand-alone simulation of TORCH has been created using the GEANT4 toolkit [107]. This simulation is used in determining both the single-photon time resoluion and the photon counting efficiency, and provides a convenent alternative to the full simulation of TORCH implemented within the LHCb simulation framework. The stand-alone simulation was first presented in Ref. [100] and has since been updated and adapted with further features.

The main improvement over the version presented in Ref. [100] is the implementation of a rudimentary charge-sharing simulation of the MCP-PMT. Using estimates of the MCP-PMT gain and point spread function, a Gaussian charge distribution is created for each photon incident on the detector. Pixels register a hit where the value of this Gaussian is greater than a pre-set threshold, thus simulating the spread of charge across multiple pixels.

The charge-sharing simulation lacks a pair of important features. There is no random element to the size of the Gaussian distribution; every photon is assumed to deposit the same amount of charge and simply the position of charge distribution varies. Further, each photon is treated independently, meaning the summation of charge resulting from two photons close in space and time is not accounted for. Further developments would aim to address both these points, creating an improved simulation of charge-sharing effects.

In addition to the charge sharing, the recording of information relating to the photon paths within the radiator plate has been added, as well as any reflection and scattering processes the photon has undergone. This was essential for validating the path-length derivation given in Appendix B. The same simulation was used for Mini-TORCH and Proto-TORCH.

## 6.5 Nomenclature and definitions

This section sets out some names and definitions which will be used throughout the rest of the Chapter.

### 6.5.1 Names

- *Hit*: a signal on a single MCP-PMT channel (pixel) resulting from an incident photon. Due to the charge-sharing nature of the MCP-PMT, a single photon should create multiple hits.
- *Cluster*: a collection of hits identified as originating from the same incident photon (more details are given in Subsection 6.7.7).

### 6.5.2 Coordinates

The incidence of the beam on the radiator plate is defined relative to the centre of the plate, with the displacement in the $x$-direction labelled $x_b$ and $y_b$ in the $y$-direction, shown in Fig. 6.6.

In all three test-beam campaigns, Mini-TORCH or Proto-TORCH were tilted back by 5° relative to the $y$-axis in the $y$-$z$-plane to increase the number of photons

captured by total internal reflection. As a result, the axis of the beam is tilted 5° in the $y$-$z$-plane from the local $z$-axis of the plate. The distance the beam covers from the T1 time reference to the front of the radiator plate along this tilted axis is labelled $l_b$. This is shown graphically in Fig. 6.6.

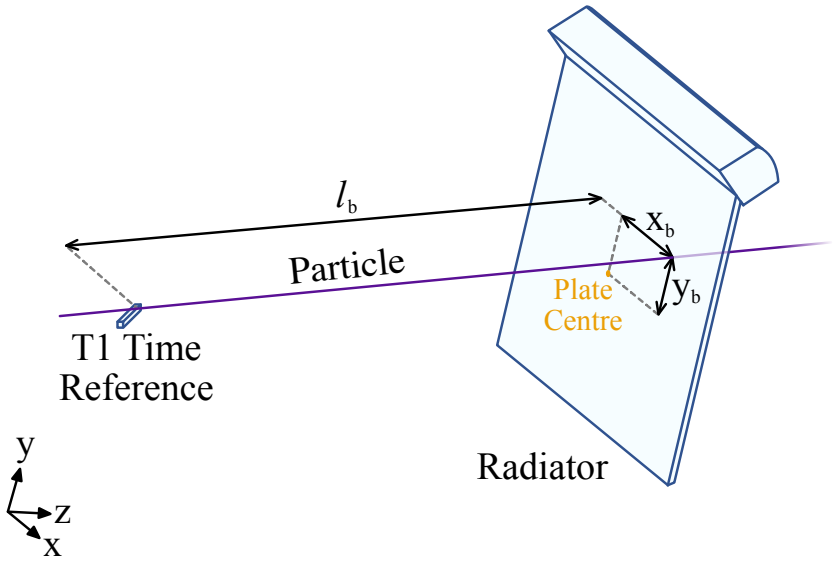



**Figure 6.6:** The definitions of $x_b$, $y_b$, and $l_b$. Note that, because of the 5° tilt of the radiator plate, $l_b$ is not parallel to the $z$-axis.

### 6.5.3 Photon paths within the radiator plate

The paths that photons take through the radiator plate are categorised, based on the number of reflections they make off the side faces of the plate. As such, it is helpful to have a system for uniquely identifying a specific path.

When viewed as projections in the $x$-$y$-plane, reflections off the front and back faces of the radiator plate are projected to leave simple paths which can be distinguished by three properties:

1. the number of times the photon reflects off a side edge of the radiator plate;
2. the side edge (if any) the photon reflects off first;
3. if the photon reflects off the bottom of the radiator plate.

The second two properties can be thought of alternatively as

2. if the photon makes a side reflection, whether it is emitted in the positive or negative $x$-direction;

3. if the photon is emitted in the positive or negative $y$-direction,

Examples of different hypothetical paths are shown in Fig. 6.7. Each path is labelled with a number which records how many side edge reflections the photon makes (property 1). The second property is tracked with a superscript $+$ or $-$ to signify whether the path reflected first off the side edge with a positive or negative $x$-position (path 0 has no reflection off a side edge, therefore it has no superscript). To track the third property, paths which reflect of the bottom are marked with a prime[3].

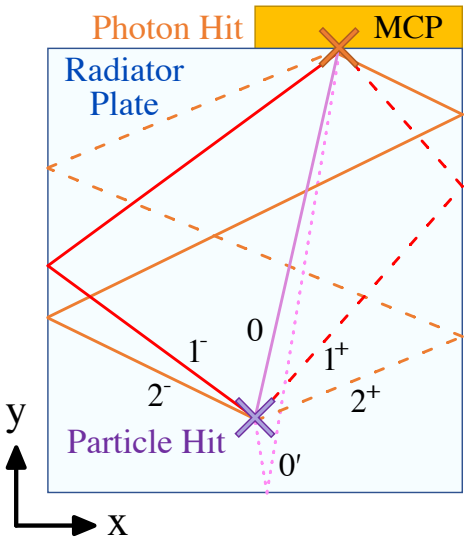



**Figure 6.7:** Examples of hypothetical photon paths within the radiator plate, viewed in the $x$-$y$-plane. Each is colour-coded based on the number of reflections. Paths reflecting first off the positive-$x$ side edge are dashed, and paths reflecting off the bottom are dotted.

[3] Note the labelling system used in this thesis differs from that used in Ref. [98] due to the need to differentiate paths which reflect off the bottom of the radiator plate as well as those which reflect off the sides.

## 6.6 Test-beam Running Conditions

### 6.6.1 November 2017

In November 2017 Mini-TORCH was instrumented with MCP-PMT Z. The columns are numbered from 1 to 4 as shown in Fig. 6.8. Although data were taken at several positions and with varying beam momenta, the data used in the analysis were taken at 5 GeV/$c$, at the positions displayed in Fig. 6.8 and given in Table 6.2. Position 1 is used for the time resolution measurement, and was chosen such that the 0 and $1^+$ paths almost overlap in space and time. This configuration allowed a good acceptence of photon hits across the MCP-PMT surface and a wide range of visible patterns of higher order reflections.

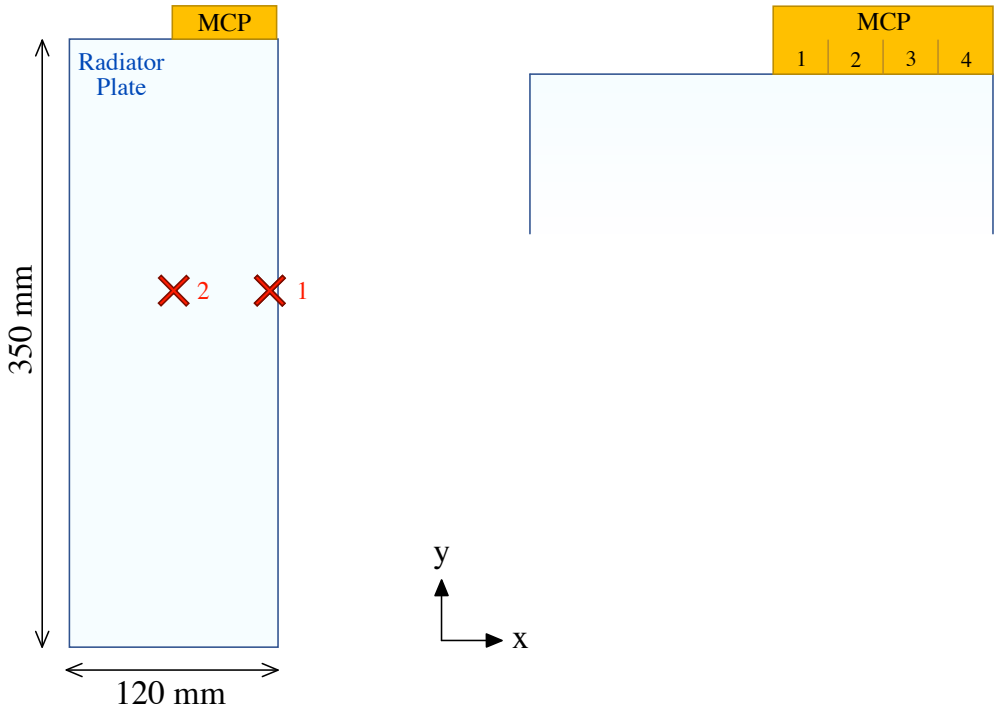



**Figure 6.8:** The position of the beam relative to Mini-TORCH in the November 2017 test-beam run. The values of the positions are given in Table 6.2.

| Position | $x_b$ ±0.3 (mm) | $y_b$ ±0.3 (mm) | $l_b$ ±2.0 (mm) |
|---|---|---|---|
| 1 | 54.9 | 13.9 | 10015 |
| 2 | 0.0 | 13.9 | 10015 |

**Table 6.2:** Beam positions for the November 2017 test-beam run (with the origin at the centre of the plate).

### 6.6.2 June 2018

In the June 2018 test-beam run Mini-TORCH was again instrumented, but now with MCP-PMT Z replaced by MCP-PMT A. The columns of MCP-PMT A are numbered 1 through 8, as shown in Fig. 6.9. The beam impinged close to the plate edge as in November 2017, but also now nominally into the vertical centre of the plate. A $3 \times 3$ grid of points was covered, displayed in Fig. 6.9 and given in Table 6.3, and with nominal beam momenta of 3.5, 5 and 8 GeV/$c$. As in November 2017, position 1 was chosen for the time-resolution analysis with a 5 GeV/$c$ beam, the lower momentum preferred as this gives a clean separation of protons and pions using the difference in time of flight.

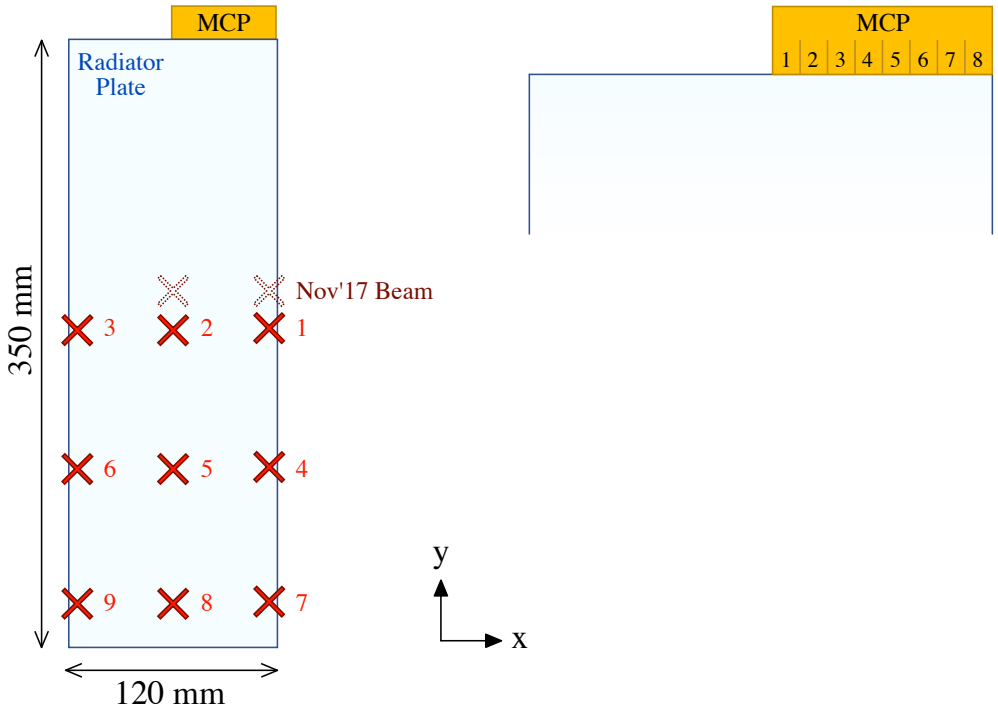



**Figure 6.9:** The position of the beam relative to Mini-TORCH in the June 2018 test-beam run. The values of the positions are given in Table 6.3. The positions used in November 2017 are shown for comparison as dotted crosses.

| Position | $x_b$ ±0.3 (mm) | $y_b$ ±0.3 (mm) | $l_b$ ±2.0 (mm) |
|---|---|---|---|
| 1 | 55.1 | −0.1 | 10017 |
| 2 | 0.0 | −0.1 | 10017 |
| 3 | −55.0 | −0.1 | 10017 |
| 4 | 55.1 | −75.1 | 10011 |
| 5 | 0.0 | −75.1 | 10011 |
| 6 | −55.0 | −75.1 | 10011 |
| 7 | 55.1 | −150.1 | 10003 |
| 8 | 0.0 | −150.1 | 10003 |
| 9 | −55.0 | −150.1 | 10003 |

**Table 6.3:** Beam positions for the June 2018 test-beam run (with the origin at the centre of the plate).

### 6.6.3 October 2018

The final test-beam period utilized Proto-TORCH, instrumented with MCP-PMTs A and B. The columns across the two MCP-PMTs are numbered 1 to 16, as shown in Fig. 6.10. Note that due to their relative positions, the columns on MCP-PMT A have been renumbered compared to June 2018.

Figure 6.10 also shows the beam positions used in the test-beam, with the coordinates of the positions given in Table 6.4. To facilitate comparisons between Proto-TORCH and Mini-TORCH, positions 1 and 2 were chosen to match the equivalent locations in the June 2018 test-beam relative to the MCP-PMTs. To allow a proper investigation of large times of propagation, a vertical column of points was chosen beneath the MCP-PMTs. Positions 6 and 7 were taken to provide additional datasets with alternative geometry and acceptance.

Nominal beam momenta of 5 and 8 GeV/$c$ were used, however only data taken with the 8 GeV/$c$ beam instrumented with MCP-PMT B were considered for the time-resolution analysis. This was because MCP-PMT B provided significantly more data due to its larger quantum efficiency (see Fig. 6.1). The decision to use 8 GeV/$c$ was made as the beam contains more particles per spill at higher momenta, and thus more data could be recorded in the limited time available.

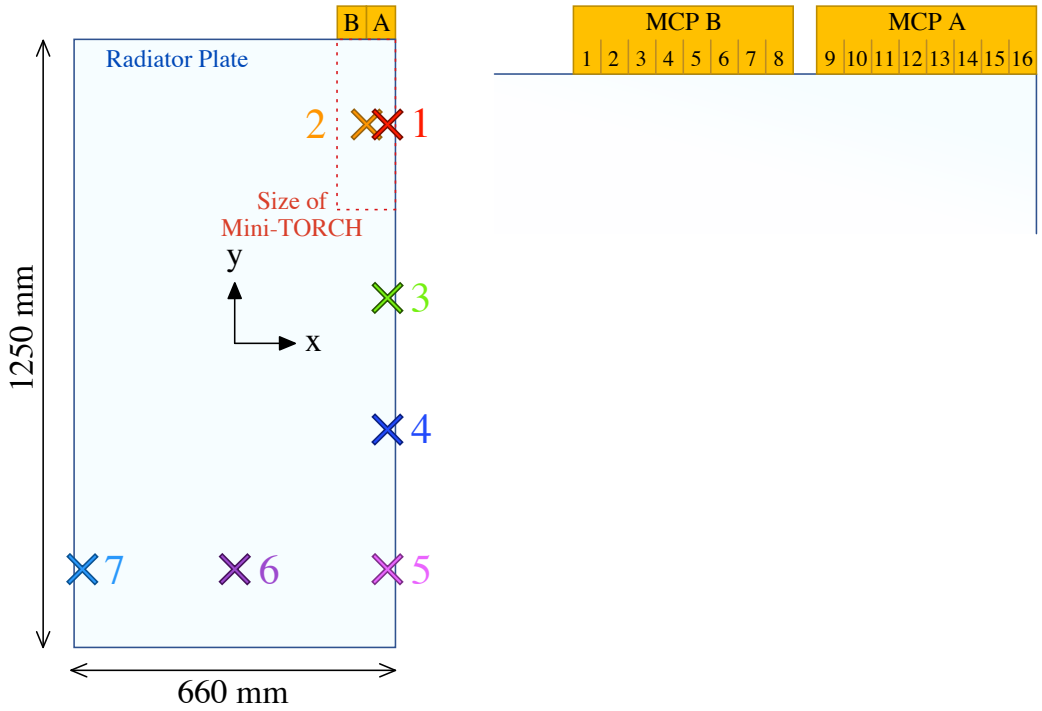



**Figure 6.10:** The positions of the beam relative to Proto-TORCH in the October 2018 test-beam run. The values of the positions are given in Table 6.4. The dotted red box indicates the size of Mini-TORCH compared to Proto-TORCH, and shows positions 1 and 2, which correspond to the equivalent positions in the June 2018 test-beam.

| Position | $x_b$ ±0.3 (mm) | $y_b$ ±0.3 (mm) | $l_b$ ±2.0 (mm) |
|---|---|---|---|
| 1 | 325.0 | 450.0 | 9415 |
| 2 | 270.0 | 450.0 | 9415 |
| 3 | 325.0 | 136.4 | 9388 |
| 4 | 325.0 | −176.9 | 9360 |
| 5 | 325.0 | −490.4 | 9333 |
| 6 | 0.0 | −490.4 | 9333 |
| 7 | −325.0 | −490.4 | 9333 |

**Table 6.4:** Beam positions for the October 2018 test-beam run (with the origin at the centre of the plate).

## 6.7 Preliminary Analysis and Corrections

### 6.7.1 Overview

Several steps must be carried out before the single-photon time resolution or photon counting efficiency can be measured. Broadly, the analysis flow is as follows:

- the data are validated via hitmaps of detected raw hits;
- the time reference signals are extracted from the data;
- an INL (Integral Non-Linearity) correction is generated and applied;
- a clustering algorithm is applied to reconstruct incident photons;
- a correction for timewalk is generated and applied.

### 6.7.2 TORCH patterns

In traditional RICH detectors, Cherenkov rings are observed and reconstructed. In TORCH, the Cherenkov rings are transformed into parabolas by the multiple internal photon reflections. Figure 6.11 shows two examples from simulation of a fully-instrumented Mini-TORCH, the first with pions incident at position 1 from the November 2017 test-beam, and the second at position 2. Distinct wings can be seen stretching across the width of the photo-detector area. The bands are seen to cover several $y'$ bins. This is a result of the angle to $y'$-position mapping of the focusing block - by design photons with different wavelengths are mapped to different $y'$-coordintes, resulting in the observed spread. The wings structure occurs due to side-reflections within the Mini-TORCH radiator plate. Considering Fig. 6.11b specifically, the top band at $y' \sim 22 - 25\,\text{mm}$ consists of photons from orders 0 and $1^+$, the band below from orders $1^-$ and $2^+$, and the bottom band from orders $2^-$ and $3^+$. These pairs of orders overlap themselves due to the beam being positioned at the side of the radiator plate. In Fig. 6.11a these orders no longer overlap and can be seen individually - the top band is order 0, the pair of bands crossing below are orders $1^\pm$, and the pair crossing below this are orders

$2^{\pm}$. Below the $2^{\pm}$ orders, from $y' \sim 5\,\text{mm}$ or less, the order 3 photons overlap with those which reflect off the bottom of the radiator plate. Here the orders rely on the time of arrival to be distinguished.

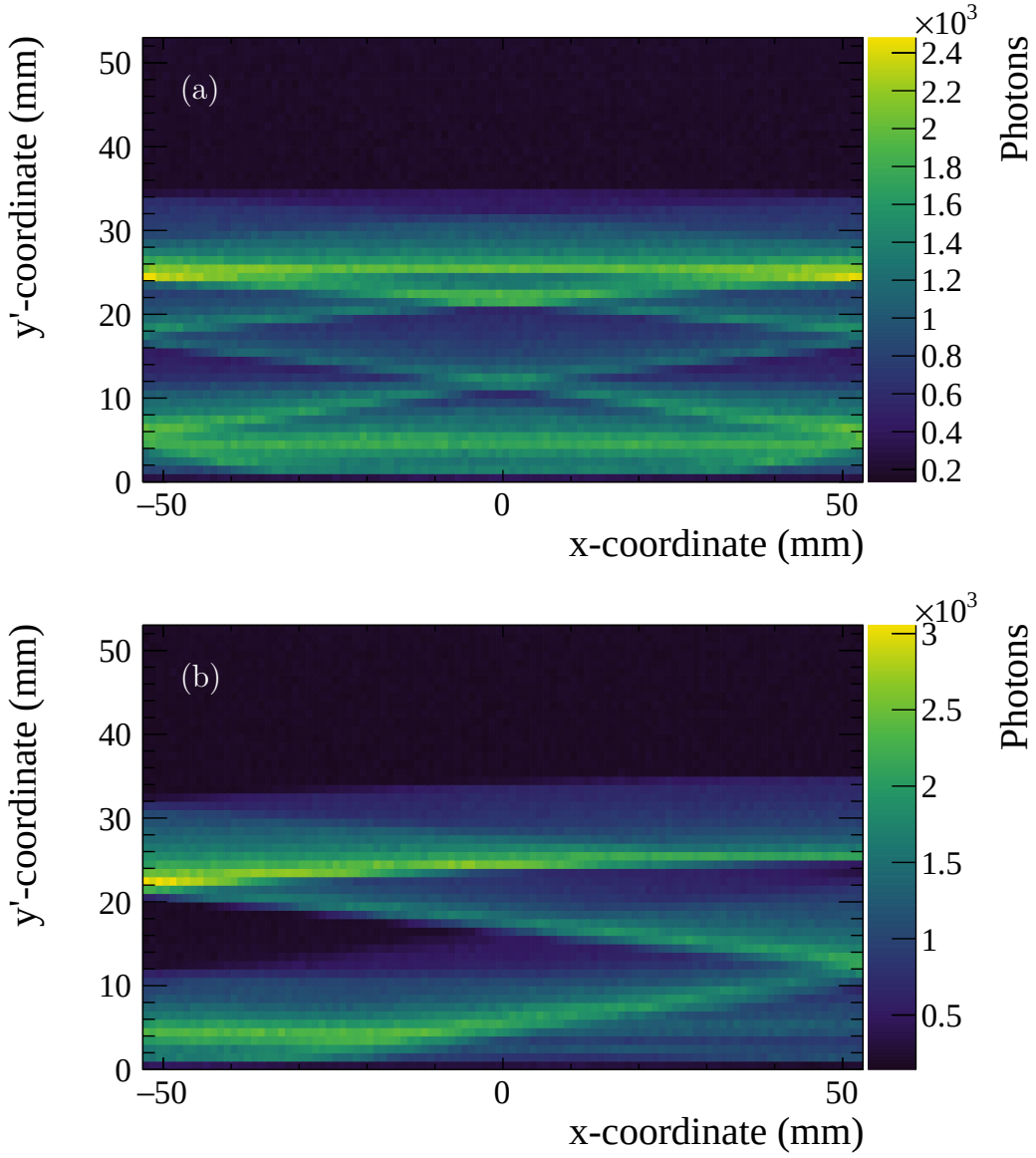



**Figure 6.11:** A hitmap of photons in a simulation of a fully-instrumented Mini-TORCH, with a beam of pions incident on the radiator plate at positions (a) 2 and (b) 1 from the November 2017 test-beam, as defined in Table 6.2.

While clean patterns are seen in Fig. 6.11, moving to data introduces the pixelisation of the MCP-PMT detector. A hitmap from the November 2017 beam test can be seen beside a comparable hitmap from simulation in Fig. 6.12. The simulation shown is the same as for Fig. 6.11b, except only photons with $x > 0$

are retained in the former case. This is because only half the detector area was instrumented in the November 2017 beam test. The pixelisation of MCP-PMT Z has also been applied, and a comparable pattern is observed.

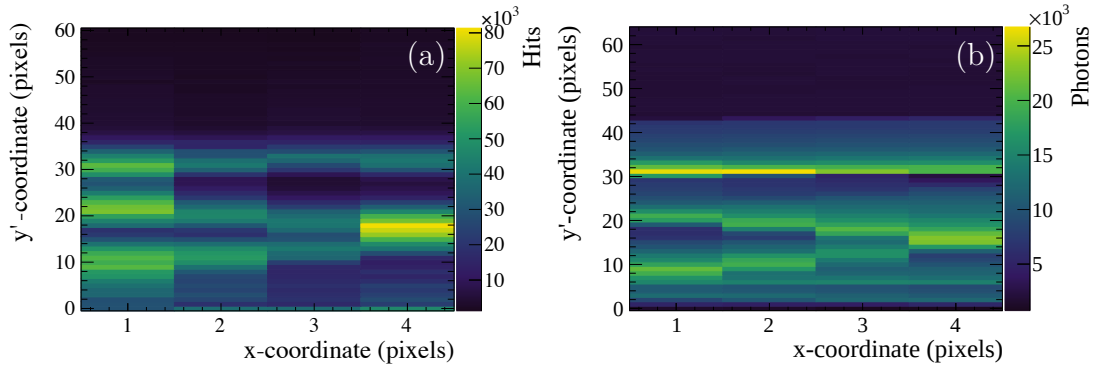



**Figure 6.12:** The position of detected hits in (x, $y'$) for incident pions in (a) data taken at position 1 in November 2017 and (b) simulation. The simulation sample is the same as shown in Fig. 6.11b, however with photons with negative $x$-coordinate removed, and the pixelisation of MCP-PMT Z applied.

Figure 6.13 shows the hitmaps observed for each beam position of the June 2018 beam test. Comparing hitmaps with the same $x$-coordinate for the beam (vertical columns in Fig. 6.13), it can be seen that moving the relative beam position downwards compresses the pattern.

Due to only one MCP-PMT being instrumented, moving the opposite side of the plate into the beam produces the "other half" of the pattern previously observed, i.e. a mirrored image. This can be intuited by flipping the hitmaps on the right of Fig. 6.13 in the $x$-direction and placing them beside those on the left - a comparable pattern to that seen in Fig. 6.11b is produced.

In theory, the hitmaps in the centre of Fig. 6.13 should match the $x > 0$ half of Fig. 6.11a. However, it is seen that even with the fine-grain pixelisation, the individual orders of reflection are impossible to visually distinguish for this small-scale module. For this reason, positions with the beam at the side of the plate are chosen for the time-resolution analysis, with the difference in propagation time of the order 0 and $1^+$ photons degrading the resolution only minimally.

A feature common to all the June 2018 hitmaps is the presence of blacked-out pixels. The pixels on even numbered columns in row 32 correspond to the

channels on which the time reference signals are injected. Other blacked out channels are due to broken wire-bonds on the NINO board, a problem which was largely rectified in future studies.

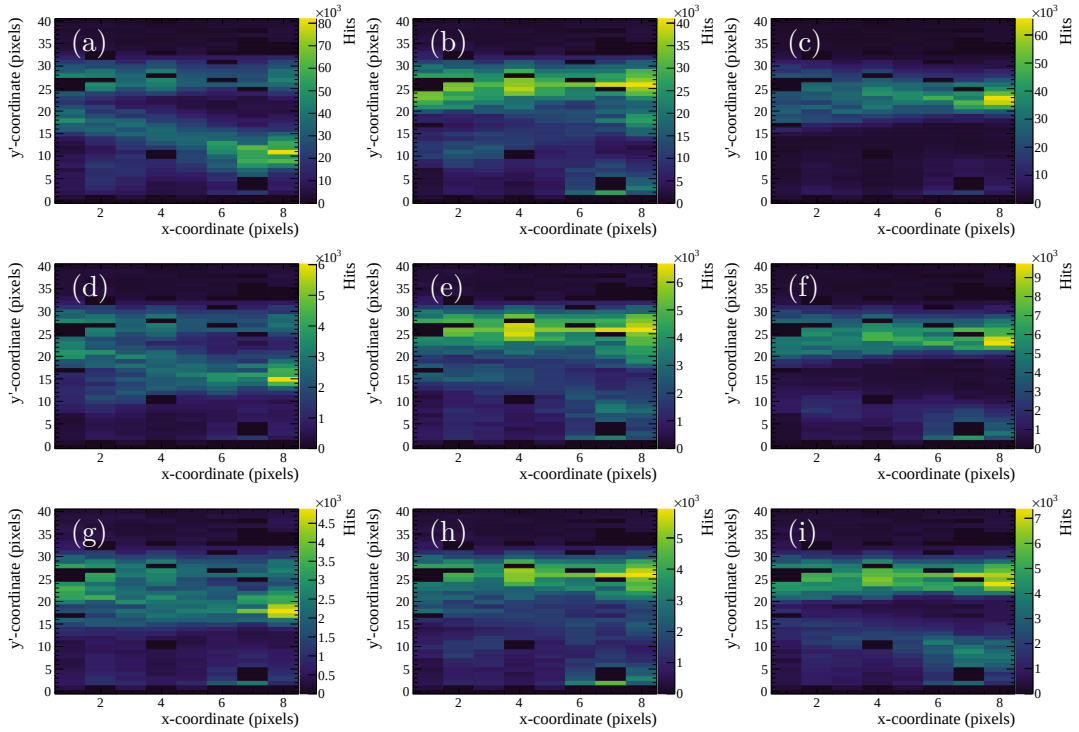



**Figure 6.13:** The position of detected hits in (x, $y'$) for incident pions in the June 2018 test-beam, with plots (a) - (i) corresponding to positions 1 - 9 respectively. Note the top 24 pixels in $y'$ have been truncated from the plots.

Finally, Fig. 6.14 shows a hitmap for position 1 of the October 2018 beam test. Comparing to the equivalent position for the November 2017 and June 2018 beam tests, Figs. 6.12a and 6.13a, the effect of the larger plate can be clearly seen; with the opposite side edge now much further away, the hypothetical path for photons which form orders $1^-$ and $2^+$ requires angles which no longer meet the total internal reflection criteria, and so these additional orders are not observed.

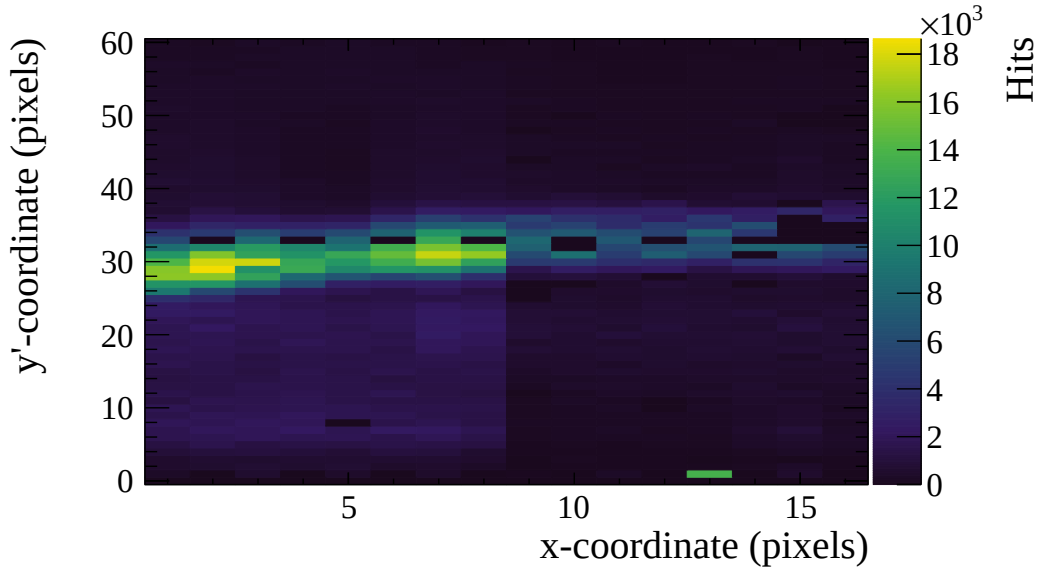



**Figure 6.14:** The position of detected hits in (x, $y'$) for incident pions in October 2018 at position 1. The left half of the hitmap (columns 1 - 8) correspond to MCP-PMT B, with the right half (columns 9 - 16) MCP-PMT A. The green pixel at low $y'$ on column 13 is a noisy pixel. The blacked-out pixels correspond to either channels which were used for time-reference signal injection or broken wire bonds on the NINO board.

### 6.7.3 Time Reference Signals

For a single triggered beam particle, identical time reference data were sent to all 16 reference channels. For November 2017 and June 2018 data, after digitisation, all these signals were averaged to provide a more precise value [100]. For October 2018 only the 4 time references corresponding to channels on MCP-PMT B were used.

Due to T1 not being required in the trigger, not all events have data for both time reference stations, however events where the incident particle did pass through T1 can be selected. This allows the time of flight of the particles between the two time reference stations to be calculated. Figure 6.15 shows the time of flight for pions and protons for three different beam momenta. The Cherenkov counters have additionally been used to separate the two species. An excellent correlation is seen.

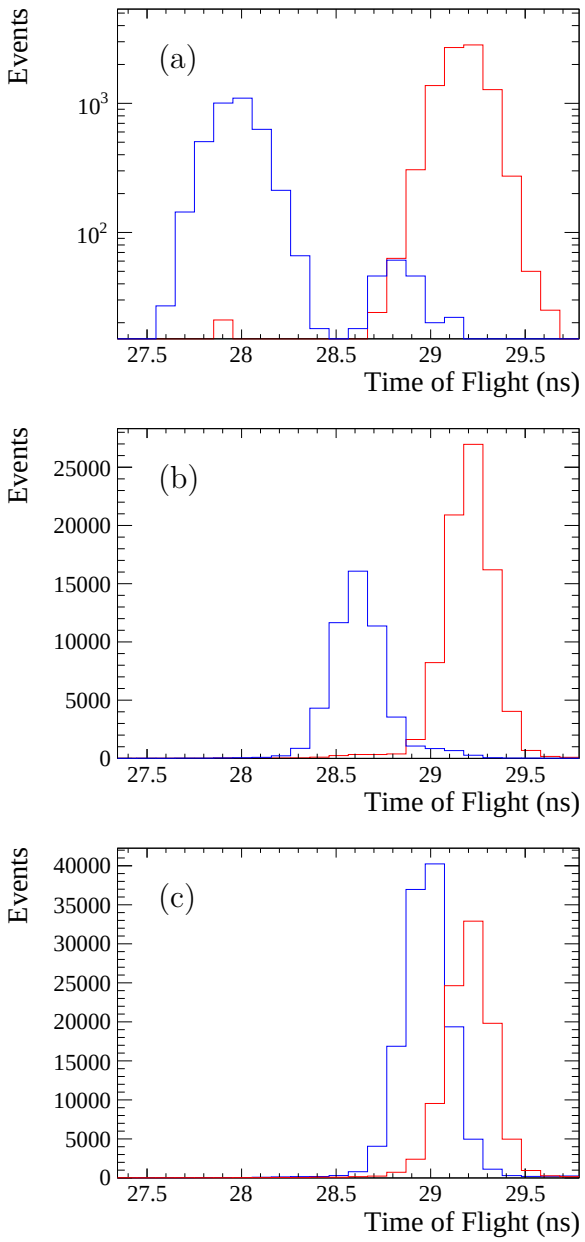



**Figure 6.15:** Particle times of flight measured in June 2018 at the nominal beam momenta of (a) 3.5 GeV/$c$, (b) 5 GeV/$c$, and (c) 8 GeV/$c$, coloured according to whether or not a signal was seen in the Cherenkov counters. Red indicates a signal was seen, and hence corresponds to pions; blue indicates no signal was seen, corresponding to protons. In (a), a second blue peak can be seen at 28.8 ns, corresponding to a small kaon signal. In (b), part of the upper tail of the proton peak may be attributed to kaons. In (c), any kaon signal is impossible to distinguish. Note the $x$-axis range is the same for all three plots for ease of comparison, however only (a) has a logarithmic $y$-axis.

### 6.7.4 Measurement of Beam Properties

Using the ToF distributions produced from both time reference signals, the beam composition can be determined from the relative size of the peaks. The beam momentum is determined using the expression

$$p_{beam} \approx \sqrt{\frac{m^2_{proton} - m^2_{pion}}{2}} \frac{Lc}{\Delta T}, \tag{6.1}$$

where $p_{beam}$ is the beam momentum, $m_X$ is the mass of species $X$, and $\Delta T$ is the mean time of flight difference between protons and pions measured over a distance $L$. To extract the relative size of the peaks and accurately measure the mean time of flight, the peaks are fitted with a pair of Gaussian distributions, shown in Fig. 6.16 for the November 2017 test-beam dataset[4].

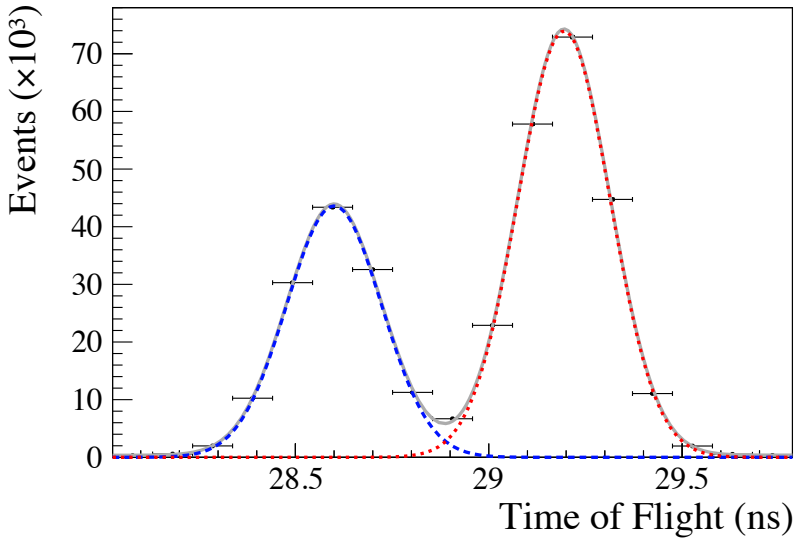



**Figure 6.16:** The ToF for protons (blue fit) and pions (red fit) in the November 2017 test-beam dataset, digitised with the ORTEC system rather than the TORCH electronics. Each peak is fit with a Gaussian distribution, with the sum of the shapes shown by the solid grey line. Reproduced from Ref. [98].

The measured beam momenta and pion fractions for the three test-beam datasets are shown in Table 6.5. Contributions from kaons are taken to be negligible. The difference in momentum and composition between November 2017 and June 2018

[4]The proton peak having a smaller time of flight compared to the pion peak is counter to expectation, and is a result of the electronics configuration.

is due to other users changing beam parameters between the test-beam periods; the magnets which set the beam momentum have hysteresis effects which prevent them from being set to identical conditions. The beam composition naturally varies with momentum, and hence the October 2018 pion fraction is significantly lower than for the two 5 GeV/$c$ test-beam periods.

| Test-Beam | Beam Momentum (GeV/$c$) | | Pion |
|---|---|---|---|
| | Nominal | Measured | Fraction (%) |
| November 2017 | 5 | $5.090 \pm 0.006$ | $64.5 \pm 0.2$ |
| June 2018 | 5 | $5.131 \pm 0.013$ | $66.2 \pm 0.4$ |
| October 2018 | 8 | $8.651 \pm 0.029$ | $47.3 \pm 1.5$ |

**Table 6.5:** The measured momenta and pion fractions of the beams for each dataset.

The sample purities, measured by comparing the Cherenkov counter decisions with the ToF distribution, are given in Table 6.6. Due to the large impurity in the November 2017 proton sample, the time-resolution analysis is carried out only on the pion sample. The sample purity cannot be measured for the October 2018 data due to the similarity in the ToF between the two species at 8 GeV/$c$.

| Beam Test | Pion Purity (%) | Proton Purity (%) |
|---|---|---|
| November 2017 | 94 | 82 |
| June 2018 | 98 | 96 |

**Table 6.6:** The purity of the pion and proton samples used in the time-resolution analysis of the November 2017 and June 2018 test-beam datasets. Note that due to the large impurity measured in the November 2017 proton sample, these data are not used in the analysis (see Subsection 6.8.6).

The beam profile is measured using the beam telescope and is defined by the cross-section of the T2 scintillators used in the trigger, as shown in Fig. 6.17a. Fitting the distribution gives an RMS spot size of $2.73 \pm 0.02$ mm in $x$ and $2.01 \pm 0.02$ mm in $y$ at the point of incidence in TORCH. The beam divergence is also measured, giving $5.8 \pm 0.2$ mrad in $x$ and $2.6 \pm 0.2$ mrad in $y$ [98]. As these values are primarily dependent on the scintillator coincidence, the same values are assumed to hold for all three test-beam periods.

Modifying the trigger results in different beam definitions. In Fig. 6.17b, the T2 scintillator requirement has been removed, triggering instead on the signal

from the borosilicate bar. In Fig. 6.17c, the borosilicate finger has been replaced with a Phase-1 TORCH MCP-PMT. Although the MCP-PMT is circular, it was positioned at an angle of 49° relative to the beam, resulting in the circle-shape being compressed horizontally.

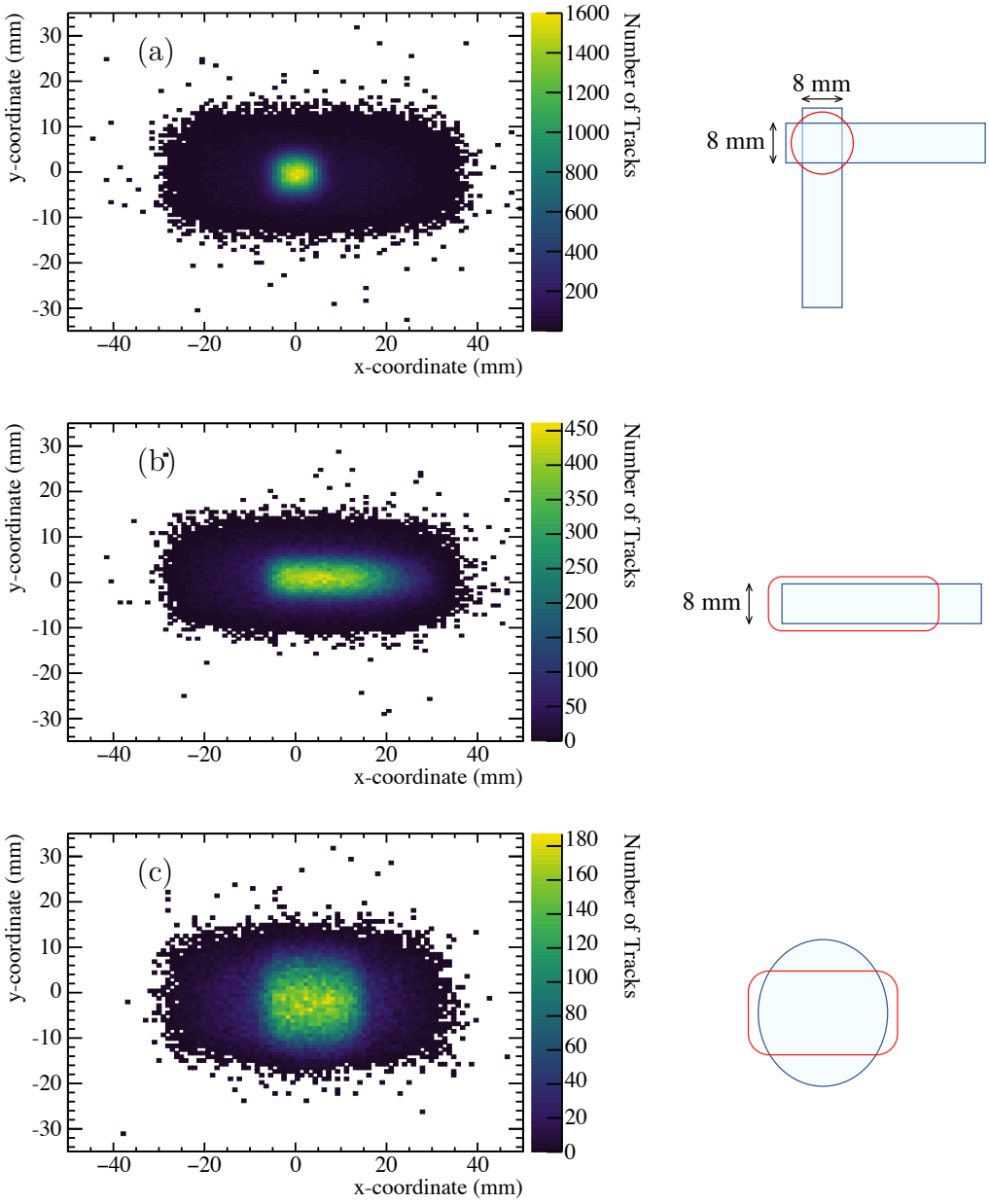



**Figure 6.17:** The distribution of tracks extrapolated from the beam telescope to the TORCH radiator plate, triggering on (a) the coincidence of the T2 scintillators, (b) the T2 borosilicate finger, and (c) a Phase-1 TORCH MCP inserted in place of the T2 borosilicate finger. In all cases, a silhouette defined by the trigger is seen, as indicated by the red boxes on the right-hand schematics. The data used were taken in November 2017.

### 6.7.5 Integral Non-Linearity Correction

The resolution which can be achieved with the time reference station is heavily influenced by the HPTDC chip. The bins used in the digitisation are not precisely spaced equally in time, leading to Integral Non-Linearity (INL), and this must be corrected for.

While the time references provide a convenient way to benchmark the INL correction, corrections must be applied to all channels. A code density test is used to perform the correction [108]. The full range in time is uniformly sampled using diffuse light and the number of hits in each bin, $P_i$, counted. $P_i$ is then compared to what would be expected if the bins were uniformly distributed in time, $P_{exp}$, to give the Differential Non-linearity ($DNL_i$):

$$DNL_i = \frac{P_i - P_{exp}}{P_{exp}} \,. \tag{6.2}$$

The INL correction for each bin, $INL_i$, is then calculated by summing the preceding DNL values:

$$INL_i = \sum_{j=0}^{i} DNL_j \,. \tag{6.3}$$

In essence, $INL_i$ gives a shift in the mean of bin $i$ which corrects for the irregular spacing of the bins.

It has been determined that approximately $10^6$ hits are needed on each channel to provide a correction where the error in the correction itself does not dominate the time-resolution measurement [100]. Therefore, during each test-beam period several dedicated runs were carried out, triggering only on the Cherenkov counters and $\mathrm{T_b}$ in order to maximise the rate. Alongside this, dedicated calibration runs for the time reference channels were taken in the lab by injecting signals from a pulse generator to complement the data taken in the test-beams.

Figure 6.18 shows an example ToF distribution before and after the INL correction is applied for the time reference signals. As shown in Table 6.7, an $\mathcal{O}(20\,\mathrm{ps})$ improvement in the width of the distributions in seen, showing the INL corrections have a significant effect on the accuracy of the measurement.

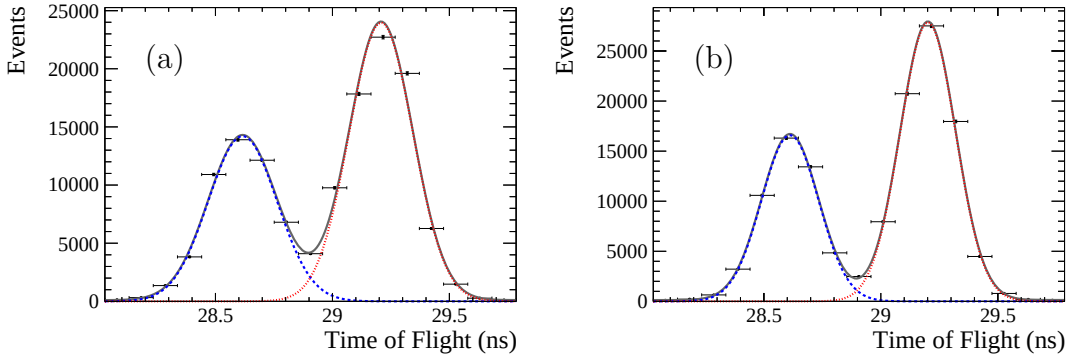



**Figure 6.18:** The time of flight for protons and pions with (a) no INL correction applied and (b) with the INL correction applied. The data have been fit with a pair of Gaussian distributions, with the blue dashed peak modelling protons and the red dotted peak pions. The solid grey line shows the sum of the two shapes. The measured widths of each Gaussian are given in Table 6.7. The data were taken in the June 2018 test-beam, with comparable behaviour seen in the other two test-beam periods.

| Peak | Gaussian $\sigma$ (ps) | | Difference in |
|---|---|---|---|
| | without INL | with INL | quadrature (ps) |
| Pions | $136.1 \pm 0.5$ | $118.6 \pm 0.4$ | 67.8 |
| Protons | $147.2 \pm 0.7$ | $122.5 \pm 0.5$ | 81.6 |

**Table 6.7:** The measured widths ($\sigma$ of the fitted Gaussian) of the proton and pion peaks shown in Fig. 6.18. Note the pion data without INL (the right Gaussian in Fig. 6.18a) does not follow a Gaussian distribution, and as a result the measured width is smaller than expected.

### 6.7.6 The Time Reference Resolution

With the INL correction applied, the time resolution of the time reference stations can be measured. Two assumptions are made:

1. the intrinsic (detector specific) time resolution of both the T1 and T2 timing stations is the same;

2. any observed difference is due to signal degradation in the 11 m length cable from the upstream station.

With these assumptions, it is possible to utilize the presence of the crossed scintillators in each station to vary the resolution by narrowing the beam profile (see Fig. 6.17). This is a result of the smaller beam profile reducing the range of possible photon propagation times (the time from photon generation to detection, which

cannot be accounted for without tracking). Through testing multiple configurations, the resolution under the conditions used in the test-beam analysis can be extracted. Three such configurations are used:

1. the requirement of scintillator signals in both T2 and T1;
2. the requirement of scintillator signals in T2 but not T1;
3. the requirement of scintillator signals in neither T2 nor T1.

A fourth set of conditions would have been possible, requiring a scintillator signal in T1 but not T2, however this significantly reduces the rate of data taking.

The combined resolution of the two timing stations can be measured using the time of flight distribution for protons and pions, which is the combination of the two stations added in quadrature,

$$\sigma_{ToF}^2 = \sigma_{T1}^2 + \sigma_{T2}^2 \,, \tag{6.4}$$

where $\sigma_{ToF}$ is the width of the proton/pion peak, and $\sigma_{T1/T2}$ is the resolution of T1/T2.

From equation 6.4, three equations can be written:

$$\sigma_{both}^2 = (\sigma'_{T1})^2 + (\sigma'_{T2})^2 \,, \tag{6.5}$$

$$\sigma_{T2\,only}^2 = (\sigma_{T1})^2 + (\sigma'_{T2})^2 \,, \tag{6.6}$$

$$\sigma_{neither}^2 = (\sigma_{T1})^2 + (\sigma_{T2})^2 \,, \tag{6.7}$$

where a prime denotes the requirement of a scintillator signal for that timing station.

To account for signal degradation between the two stations, a factor $\phi$ is introduced. This provides a link between the two station resolutions, resulting in

$$\sigma_{T2} = \phi\sigma_{T1} \tag{6.8}$$

and

$$\sigma'_{T2} = \phi\sigma'_{T1} \,, \tag{6.9}$$

where it is assumed the amount of degradation is independent of requiring a scintillator signal.

Using equations 6.5 to 6.9, the resolution of each station and the factor $\phi$ can be determined from the width of the proton/pion peak through the following expressions:

$$\phi^2 = \frac{\sigma^2_{both} - \sigma^2_{T2\,only}}{\sigma^2_{T2\,only} - \sigma^2_{neither}}\,, \tag{6.10}$$

$$(\sigma'_{T2})^2 = \frac{\sigma^2_{both}}{1+\phi^2}\,, \tag{6.11}$$

and

$$(\sigma_{T2})^2 = \frac{\sigma^2_{neither}}{1+\phi^2}\,. \tag{6.12}$$

The ToF distributions used to measure the resolutions for each set of conditions are shown in Fig. 6.19, with the measured width values given in Table 6.8. The derived time resolutions are presented in Table 6.9.

| Scintillators required | Figure | Proton peak Width (ps) | Pion peak Width (ps) |
|---|---|---|---|
| Neither | 6.19a | $132.2 \pm 0.7$ | $128.2 \pm 0.5$ |
| T2 only | 6.19b | $122.7 \pm 0.3$ | $119.2 \pm 0.2$ |
| Both | 6.19c | $86.8 \pm 1.2$ | $86.9 \pm 0.8$ |

**Table 6.8:** The widths of the proton and pion peaks shown in Fig. 6.19.

| Variable | Value (Protons) | Value (Pions) |
|---|---|---|
| $\phi$ | $1.76 \pm 0.08$ | $1.73 \pm 0.06$ |
| $\sigma'_{T2}$ | $42.8 \pm 2.0$ ps | $43.4 \pm 1.5$ ps |
| $\sigma_{T2}$ | $65.2 \pm 3.1$ ps | $64.1 \pm 2.2$ ps |

**Table 6.9:** The degradation factor and resolutions of the downstream timing station derived from the data shown in Fig. 6.19.

The resolutions for T2 are consistent when measured for pions and protons, and also with the previous measurement made in [69]. Requiring the scintillator signal is seen to improve the resolution by $\sim 20$ ps. Because of this improvement, and to recap, T2 was required in the trigger when measuring the single-photon time resolution. Although this resolution measurement was made with June 2018 data

only, the same resolution is assumed to be valid for November 2017 and October 2018 as the infrastructure was identical apart from the DUT. The value of $\sigma'_{T2}$ ($\sim 43\,\text{ps}$) is an important measurement, since this provides the start time reference ($t_0$) for TORCH timing resolution measurements in all the analyses.

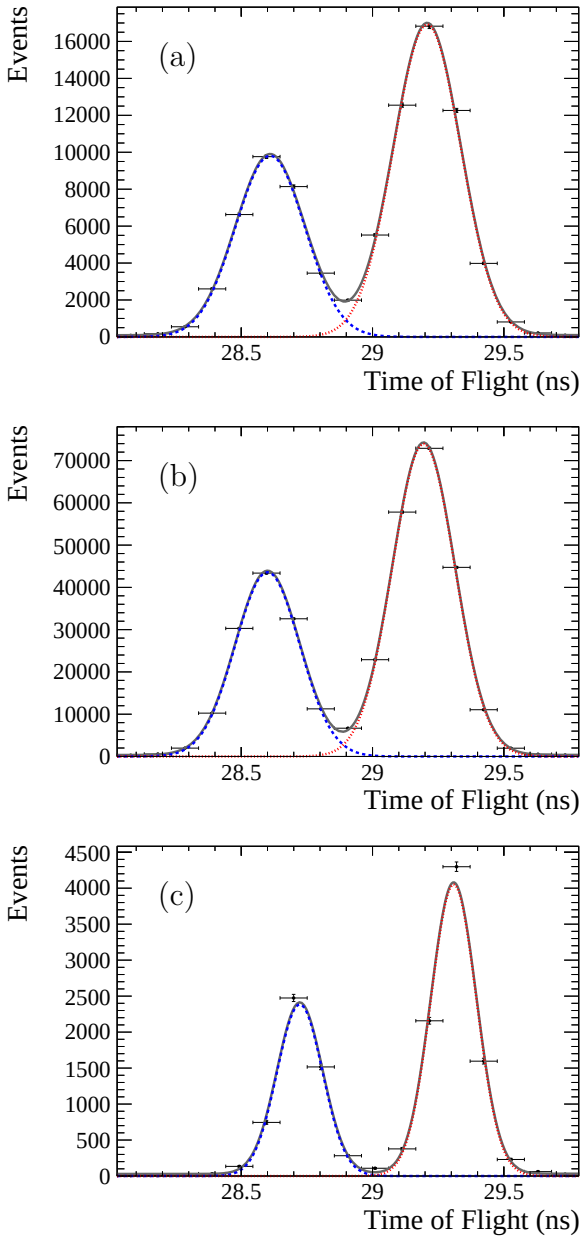



**Figure 6.19:** The time of flight for protons and pions with (a) no scintillator signals required, (b) only the T2 scintillator required, and (c) signals in both T1 and T2 required. The blue dashed peak corresponds to signals from protons and the red dotted peak signals from pions. The peaks can be seen to narrow from (a) through (c), and the measured widths are given in Table 6.8. With both scintillators required, it can be seen the peaks begin to become non-Gaussian, and the fit is unable to fully model the data. The data used was taken in the June 2018 test-beam.

### 6.7.7 Clustering

As detailed in Subsection 5.3.3, photons incident on the MCP-PMT window strike a photocathode which converts them into electrons. The electron signal is amplified by the MCP, creating electron cascades which are collected by an anode. As the cascades can fall across multiple pixels, detected hits must be clustered across pixels to reconstruct the individual photons.

Hits are clustered if they meet all the following criteria:

- they have the same $x$ coordinate (coarse pixel direction);
- they are neighbouring in the $y'$-coordinate;
- they are within 1 ns of their neighbour in time.

For the June 2018 and October 2018 datasets, the criteria were slightly modified to account for dead channels. If two clusters fall on either side of a known dead channel and the neighbouring hits are recorded within 2 ns of each other, then the clusters are merged.

The distribution of cluster sizes in the June 2018 data is shown in Fig. 6.20. Incident photons are expected to generate clusters of between two and four hits, and so only clusters of these sizes are selected for further analysis.

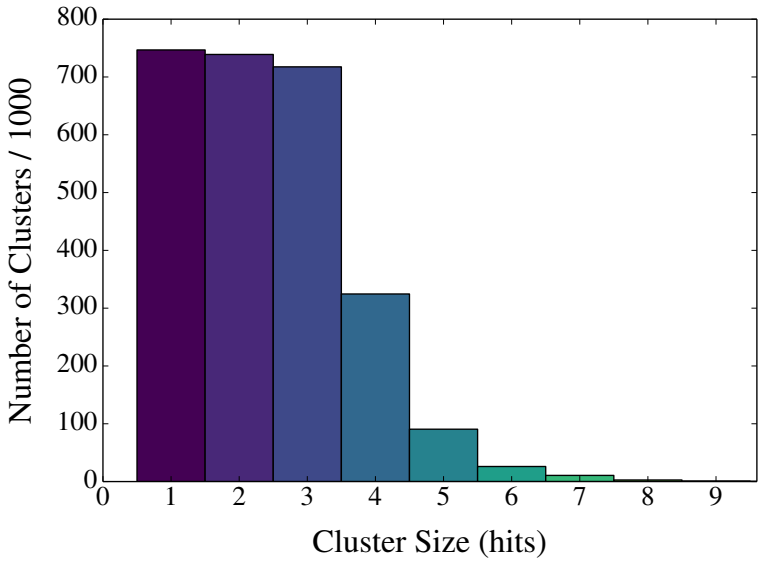



**Figure 6.20:** The distribution of cluster sizes observed in the June 2018 dataset. The bar colour is simply to make the bars easier to visially distinguish and has no relevance to the counts.

### 6.7.8 Timewalk Correction

Once clusters have been formed, another correction is applied to account for varying pulse amplitudes (time slew) which have been digitized by the NINO chip. As the NINO employs time-over-threshold, the timestamp recorded by the rising and falling edges of the input pulse crossing the threshold can be dependent on the pulse size. To account for this, a calibration of the charge-to-width correlation for each NINO channel should in principle be performed. Unfortunately, laboratory calibration was not available at this time, and so instead a data-driven approach was adopted.

As each cluster corresponds to one incident photon, it is assumed the hits within that cluster are simultaneous in time, and that, once any constant channel-to-channel effects have been subtracted, any difference is due to time slew introduced by differing pulse amplitudes. Taking $T_{Ch0}$ and $T_{Ch1}$ to be the times of two hits within the cluster, it can be implied from Fig. 6.21 that the difference between them, $\Delta T$, is proportional to the widths of the pulses on each channel. Thus,

$$T_{Ch0} - T_{Ch1} = \Delta T \left( w_{Ch0}, w_{Ch1} \right), \tag{6.13}$$

where $w_{ChX}$ is the width of the pulse corresponding to the hit for $T_{ChX}$. A two-dimensional second-order polynomial is chosen to model $\Delta T\,(w_{Ch0}, w_{Ch1})$ for each pair of channels and, through fitting the $\Delta T$ distribution as a function of $w_{Ch0}$ and $w_{Ch1}$, the coefficients are determined. Using this polynomial correction, the time of each pixel hit can be offset for the variation in input pulse height[5]. Figure 6.22 shows the variation in offset with pulse width for four pixels on different MCP-PMT columns.

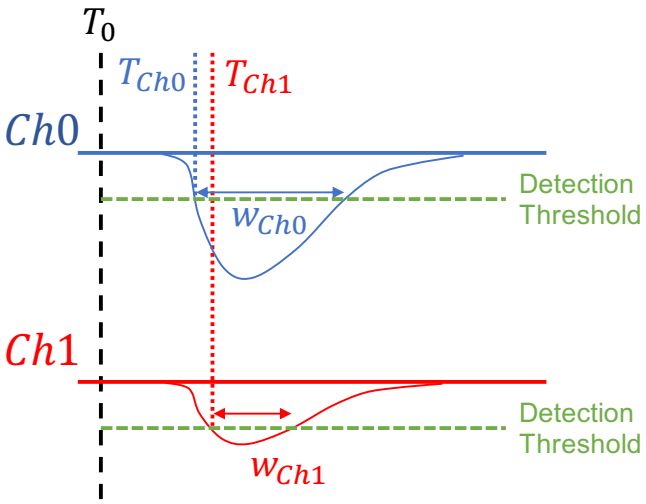



**Figure 6.21:** A schematic of the time-over-threshold performed by the NINO chip. The pulse on Channel 0 is drawn purposely larger to show the relationship between pulse width, $w_{ChX}$, and the resulting leading time measurement, $T_{ChX}$. Although not labelled, the same behaviour is seen for the trailing edge.

[5]The timewalk corrections for the November 2017 and June 2018 datasets were determined by a TORCH collaborator, but applied to the data by the author. For October 2018 they were created and applied by the author.

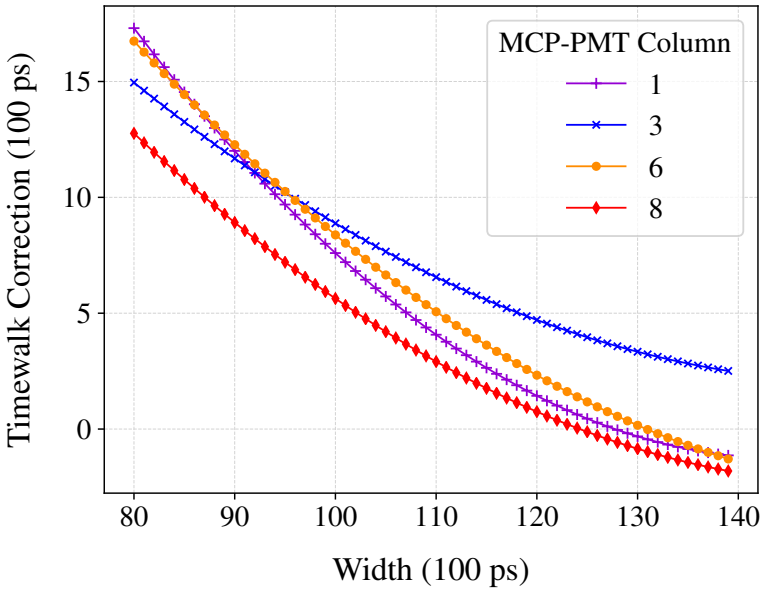



**Figure 6.22:** The variation in the timewalk correction with pulse width for single pixels on four different MCP-PMT columns in the June 2018 test-beam period. The 100 ps units correspond to the bin size of the HPTDC chip.

## 6.8 Single-Photon Time-Resolution Measurement

### 6.8.1 Overview

To measure the single-photon time resolution, the time of arrival for each photon is predicted and compared to that measured. The differences for all observed photons are combined into a distribution of residuals, the width of which gives the spread in time of photons from the prediction. This spread, $\sigma_{Total}$, is a combination of several factors which must be subtracted in quadrature to give the single-photon timing resolution being measured, $\sigma_{TORCH}$,

$$\sigma^2_{Total} = \sigma^2_{TORCH} + \sigma^2_{Beam} + \sigma^2_{TimeRef}\,, \tag{6.14}$$

where $\sigma_{Beam}$ is the spread in the distribution resulting from the profile of the beam, and $\sigma_{TimeRef}$ is the resolution of the T2 time-reference station, measured as described in Subsection 6.7.6.

### 6.8.2 Photon Reconstruction

To predict the time of arrival for a given photon, the photon path through the radiator plate and focusing block must be determined. For a specific incident particle hypothesis, given an order of reflection, the photon path length can be calculated using the procedure described in Appendix B. With the photon path known, the Cherenkov angle, $\theta_c$, can be found from the position of the photon hit and by finding the angle between the photon trajectory and the track of the emitting particle. Using the momentum and mass of the emitting particle, the phase refractive index, $n_p$, can be calculated via the expression

$$n_p = \frac{1}{\beta \cos(\theta_c)}\,, \tag{6.15}$$

where $\beta$ is the velocity of the primary particle hypothesis relative to the speed of light in vacuum.

The group velocity determines the time of propagation for a photon over a fixed distance, so the group refractive index, $n_g$, must be determined from $n_p$. The two are related by the dispersion relation [109]

$$n_g = n_p - \lambda \frac{\partial n_p}{\partial \lambda}\,, \tag{6.16}$$

and so, by considering a range of wavelengths covering the typical energies of photons within TORCH (roughly 2 - 7 eV), a look-up table relating $n_g$ and $n_p$ can be made. To calculate the differential in (6.16), the Sellmeier equation is used,

$$n_p^2 = 1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3}\,, \tag{6.17}$$

where $\lambda$ is the wavelength of the photon, and $B_{1,2,3}$ and $C_{1,2,3}$ are experimentally determined coefficients dependent on the media in which the photon is propagating [110].

Finally, the group velocity is calculated using

$$v_g = \frac{c}{n_g}\,, \tag{6.18}$$

and hence the time for the photon to propagate from emission to the detector can be found by combining the group velocity with the path length.

While in principle the above procedure can unambiguously calculate the time of propagation for an individual photon, this requires exact knowledge of the geometric path. Within TORCH, two ambiguities exist: the point of emission within the plate, and the point where the photon enters the focusing block. If either one were known, the other could be calculated, but in combination they represent an unsolvable system.

To account for the ambiguities, when considering the possible photon paths in the radiator plate, many different positions of entry to the focusing block from the radiator plate are considered. For each one, the path is tracked back to the emitting particle, and the physicality of the path validated (i.e. by ensuring the Cherenkov angle is physically possible). All hypotheses with non-physical paths are discarded, and the times of propagation of the remaining paths are averaged to give a final prediction.

### 6.8.3 Separating Reflections with Time Projections

By considering the time of arrival of a photon hit and the $y'$-coordinate of the corresponding MCP-PMT column, the orders of reflection in the radiator plate can be separated. Figure 6.23 shows an example from the June 2018 data. Three distinct bands are observed, corresponding to the various orders. The lines overlaid on Fig. 6.23a show the predicted position of each order of reflection according to Fig. 6.23b, showing how the position of the beam at the side of the radiator plate results in overlapping orders.

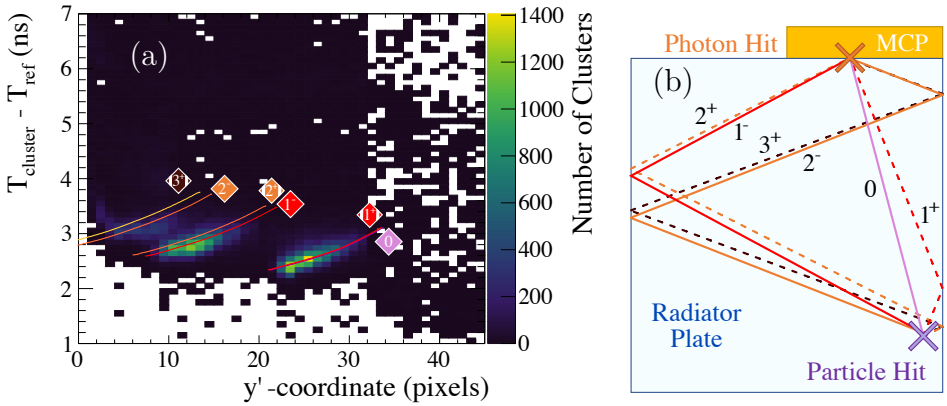



**Figure 6.23:** (a) Average time of detected hits within a cluster relative to the T2 time reference versus the average $y'$ coordinate of detected hits within the cluster for column 4 of the MCP-PMT in the June 2018 dataset. The distinct bands observed correspond to different orders of reflection in the radiator plate. The overlaid lines show where a particular reflection is expected, assuming the paths shown in (b). Channels with $y' \geq 32$ are read out by a different NINO chip to those with $y' < 32$, leading to the discontinuity observed in the background.

Corresponding distributions for different MCP-PMT columns are shown for the November 2017 test-beam in Fig. 6.24. It is seen that the higher orders become closer overlapping when moving from column 1 to 4. Additionally, photons which have reflected off the bottom of the radiator plate can be seen in the upper-left areas of the figures. While some of these match well the expectations from the reconstruction, in Fig. 6.12d, especially, there is a clear offset. This is most likely due to the timewalk calibration not accurately correcting the data at low $y'$ values. As the correction is data-driven, calculating an accurate correction in this sparsely populated region has not been possible due to the limited time available to take data.

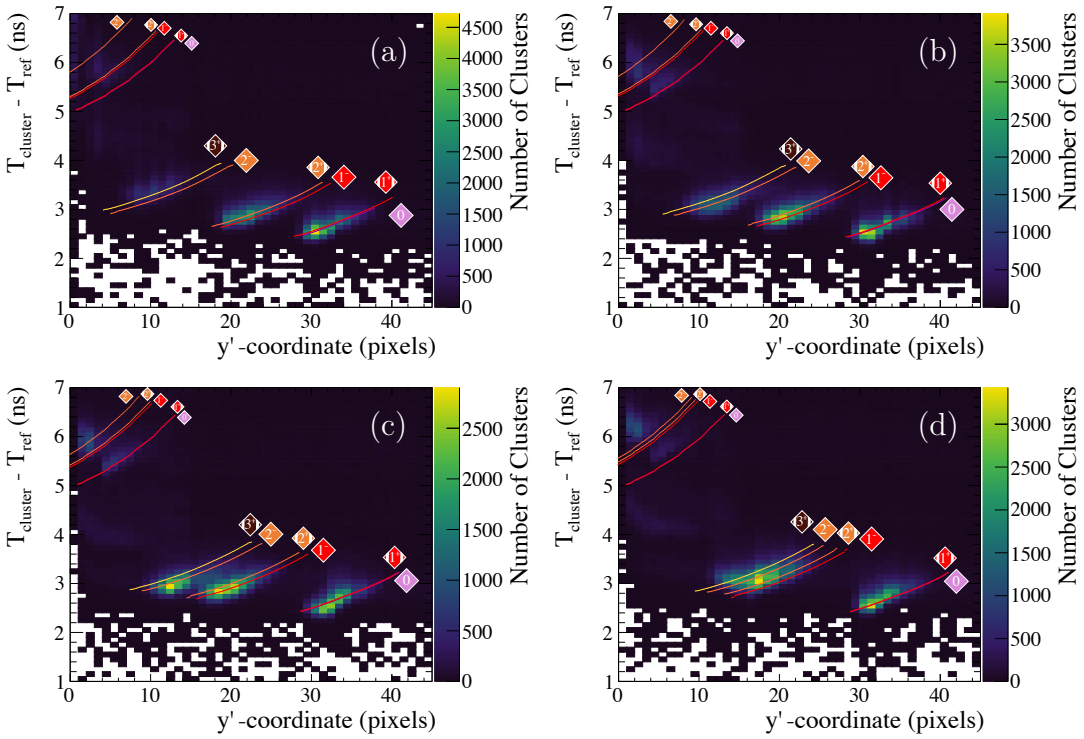



**Figure 6.24:** The measured time versus vertical position for clusters detected in the November 2017 test-beam data, with (a) - (d) showing MCP-PMT columns 1 - 4 respectively. The distinct bands observed correspond to different orders of reflection in the radiator plate, with the overlaid lines showing the predictions.

The effect of the increased time of propagation can be investigated by comparing the projections in time as a function of beam position from the MCP-PMT plane. This is best seen in the October 2018 data with the large radiator plate where the path length is longest. The time projections are shown for positions 1, 3, 4, and 5 in Fig. 6.25. The pattern is seen to stretch in time as the beam is incident further down the radiator plate. This is because the range of propagation times for photons within the angular acceptance of a given $y'$ pixel increases with the displacement of the beam from the MCP-PMTs. At the lowest position, Fig. 6.25d, reflections off the bottom of the radiator plate become visible in the top left of the plot, as a result of the beam being only a few centimetres above the bottom face.

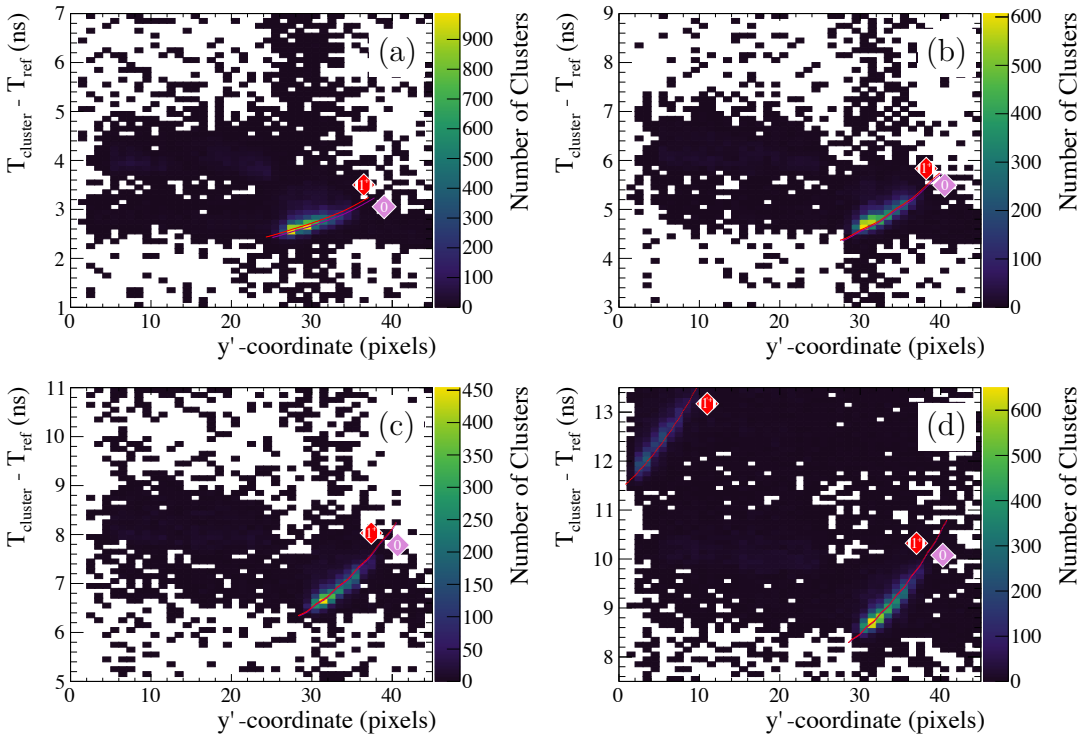



**Figure 6.25:** The measured time versus vertical position for clusters in the October 2018 dataset detected on MCP-PMT column 1 with the beam at positions (a) 1, (b) 3, (c) 4, and (d) 5.

### 6.8.4 Measuring the Time Resolution

With the photon orders of reflection separated, the time resolution can be measured by creating a residual distribution of the difference between the predicted and measured times of arrival for each detected single photon. When making this measurement, only the 0 and $1^+$ orders are used (in this geometrical configuration, these are the only orders which are cleanly separated from the others for all MCP-PMT columns)[6]. Due to the overlap of the 0 and $1^+$ reflections, those orders are treated as one, with the 0 order prediction used for the comparison.

Figure 6.26 shows examples of residual distributions from the November 2017 test-beam period, corresponding to the data and predictions shown in Fig. 6.24. A fit with a Crystal-Ball function [111] (see Appendix A) is used to measure the width of the distribution to determine $\sigma_{Total}$ in Eq. (6.14). The tail of the Crystal-Ball

[6]Analysis is in progress to measure higher orders, however is outside the scope of this thesis.

accounts for backscattering of primary photoelectrons in the MCP-PMT [109] and contributions from residual time-walk effects.

When creating the residual distributions, a correction is applied to account for energy-dependent inaccuracies in the time of arrival prediction. Within a given order of reflection, the $y'$ position of the photon hit is directly correlated to the photon energy, therefore, as photons of all energies are included in the residual distribution, any systematic offset in relation to energy artificially broadens the distribution. To correct for this, a residual distribution is calculated for each pixel $y'$ bin and fitted to determine the mean. The photon arrival times within this distribution are then offset by this mean, before being recombined to give the final residual distribution.

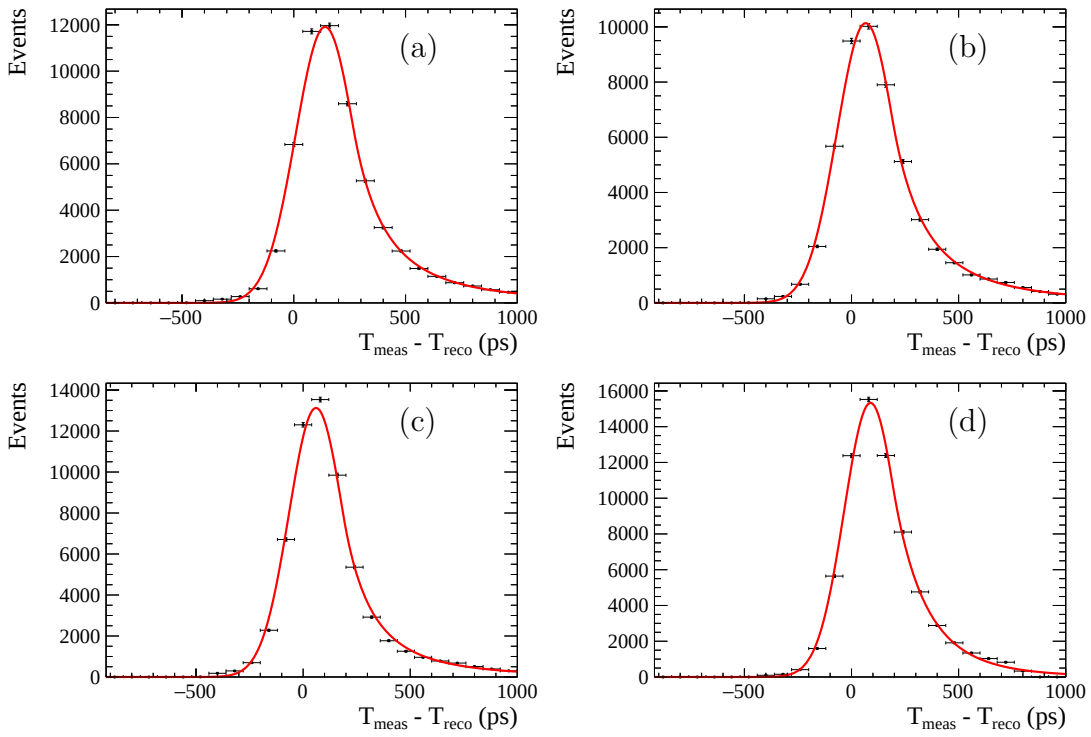



**Figure 6.26:** Residual distributions for the November 2017 test-beam data, showing the difference of the measured time, $T_{\text{meas}}$, and predicted time, $T_{\text{reco}}$, for the order 0 reflection. Plots (a) - (d) correspond to the MCP-PMT columns 1 - 4.

### 6.8.5 The Beam-Spread Contribution

To calculate $\sigma_{TORCH}$ in Eq. (6.14), the contribution to the residual distribution due to the uncertainty in timing from the beam spread must be subtracted. The nominal beam position and direction are used when calculating the time of arrival prediction, where in actuality the beam has finite size and particles enter the radiator plate over a range of angles.

The stand-alone simulations of Mini-TORCH and Proto-TORCH are used to estimate $\sigma_{Beam}$ as a function of beam position and MCP-PMT column number. First, a pencil beam, where all the particles are incident at the same point and have the same direction of travel, is used to build a residual distribution. A Gaussian shape is fit to this distribution and the width extracted. Second, the simulation is run using input from real tracks measured by the beam telescope in position and angle of incidence. The resulting residual distribution is fit with a pair of convolved Gaussians, one of which has the width fixed to the width of the Gaussian distribution from the pencil beam simulation. The width of the second Gaussian in the convolution therefore gives the contribution to the width resulting from the spread of the beam. An example of residual distributions built from simulation with a pencil beam and realistic beam are shown in Fig. 6.27. Note no MCP-PMT smearing has been included.

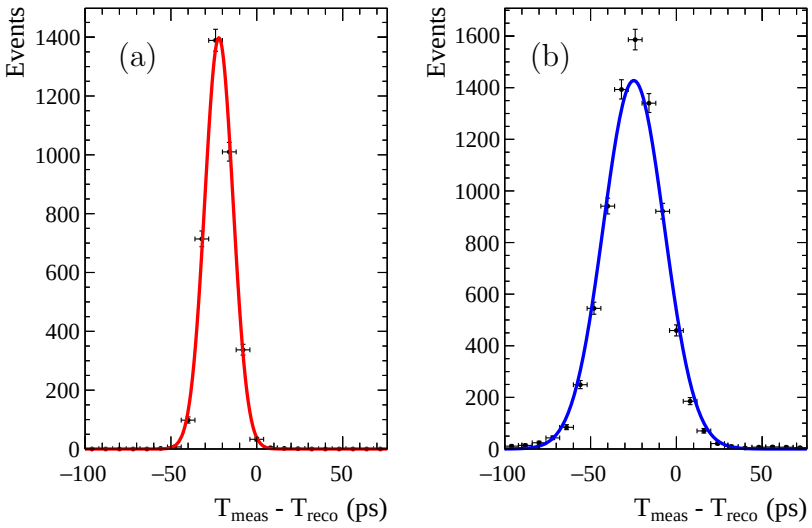



**Figure 6.27:** The time spread measured using the Mini-TORCH simulation with (a) a pencil beam and (b) the measured beam profile from November 2017. A simultaneous fit has been performed, (a) with a single Gaussian, and (b) with a pair of convolved Gaussians, where one Gaussian has the same width as in (a), and the width of the other gives the contribution to the time spread due to the beam profile.

The beam-spread contributions for each test-beam period are given in Tables 6.10, 6.11 and 6.12. The contributions for November 2017 are larger than June 2018 and position 1 of October 2018 due to different beam parameters being used. Consistent results are seen with simulated pions and protons.

| MCP Column | $\sigma_{Beam}$ (ps) |
|---|---|
| 1 | $15.6 \pm 0.1$ |
| 2 | $15.0 \pm 0.1$ |
| 3 | $15.7 \pm 0.2$ |
| 4 | $16.8 \pm 0.2$ |

**Table 6.10:** The contributions to the timing resolution due to the beam spread for the November 2017 test-beam period, determined from a comparison of simulation samples generated with a pencil beam and the measured beam spread.

| MCP Column | $\sigma_{Beam}$ (ps) |
|---|---|
| 1 | $13.0 \pm 0.1$ |
| 2 | $12.7 \pm 0.1$ |
| 3 | $12.8 \pm 0.1$ |
| 4 | $13.2 \pm 0.1$ |
| 5 | $13.3 \pm 0.2$ |
| 6 | $13.5 \pm 0.2$ |
| 7 | $14.3 \pm 0.2$ |
| 8 | $14.6 \pm 0.2$ |

**Table 6.11:** The contributions to the timing resolution due to the beam spread for the June 2018 test-beam period, determined from a comparison of simulation samples generated with a pencil beam and the measured beam spread.

| MCP Column | $\sigma_{Beam}$ Pos. 1 (ps) | $\sigma_{Beam}$ Pos. 3 (ps) | $\sigma_{Beam}$ Pos. 4 (ps) | $\sigma_{Beam}$ Pos. 5 (ps) |
|---|---|---|---|---|
| 1 | $13.0 \pm 0.2$ | $19.3 \pm 0.0$ | $20.0 \pm 1.0$ | $30.2 \pm 1.0$ |
| 2 | $14.3 \pm 0.3$ | $19.6 \pm 0.5$ | $23.0 \pm 1.2$ | $32.5 \pm 0.9$ |
| 3 | $15.5 \pm 0.3$ | $16.1 \pm 0.5$ | $22.4 \pm 1.0$ | $31.9 \pm 1.0$ |
| 4 | $14.4 \pm 1.4$ | $18.3 \pm 0.7$ | $23.7 \pm 1.1$ | $29.7 \pm 1.1$ |
| 5 | $15.3 \pm 0.3$ | $19.5 \pm 0.5$ | $23.2 \pm 1.3$ | $29.5 \pm 1.0$ |
| 6 | $15.1 \pm 0.3$ | $18.4 \pm 0.5$ | $23.8 \pm 1.6$ | $30.5 \pm 1.0$ |
| 7 | $15.4 \pm 0.3$ | $20.1 \pm 0.5$ | $24.0 \pm 4.9$ | $29.4 \pm 1.0$ |
| 8 | $15.4 \pm 0.3$ | $18.8 \pm 0.7$ | $22.3 \pm 2.0$ | $28.9 \pm 1.3$ |

**Table 6.12:** The contributions to the timing resolution due to the beam spread, measured from a comparison of simulation samples generated with a pencil beam and the measured beam spread, for each MCP-PMT column and position down the side of the radiator plate in the October 2018 test-beam period.

A strong dependence on the position of beam incidence is predicted in the October 2018 data, more clearly seen in Fig. 6.28. The beam profile remains constant at each position, and the range of photon angles within the radiator plate is also the same.

However, because of the increased photon path length, there is an increase in the range of propagation times for photons with the same range of Cherenkov angles.

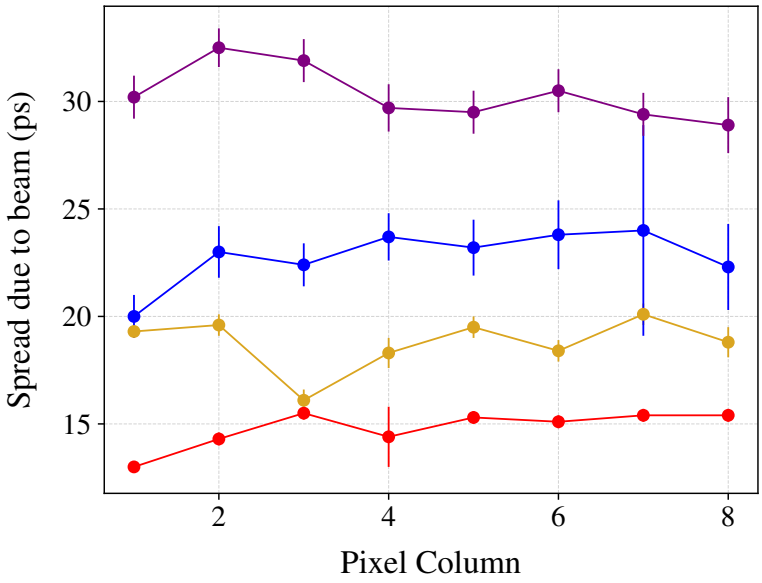



**Figure 6.28:** The simulated values of $\sigma_{Beam}$ as a function of pixel column number, with beam parameters measured for the October 2018 test-beam, as presented in Table 6.12. From bottom to top, the red points show position 1, yellow position 3, blue position 4 and purple position 5.

### 6.8.6 Final Results

With $\sigma_{Beam}$ and $\sigma_{TimeRef}$ known, $\sigma_{TORCH}$ can be calculated. Tables 6.13 and 6.14 present the measured values of $\sigma_{TORCH}$, with $\sigma_{Beam}$ and $\sigma_{TimeRef}$ subtracted in quadrature, for the November 2017 and June 2018 test-beam periods, respectively. These results are also reported in Ref. [98]. A general improvement is observed between the November 2017 and June 2018 resolutions, attributed to the improved $x$-direction granularity of MCP-PMT A compared to MCP-PMT Z. In June 2018 specifically, columns 5 and 7 are seen to exhibit better resolutions than the others as the timewalk correction appears to perform better on these columns. Six of the eight columns are also seen to have marginally better resolutions when measured with the pion sample than the proton sample. This effect is attributed to the better purity of the pion sample compared to the proton sample (see Table 6.6). Note

that when the incident particle is misidentified, the wrong mass will be used in the reconstruction, hence providing an incorrect predicted time.

| MCP Column | $\sigma_{TORCH}$ (ps) |
|---|---|
| 1 | $112.1 \pm 1.4$ |
| 2 | $114.8 \pm 1.4$ |
| 3 | $104.3 \pm 1.4$ |
| 4 | $111.4 \pm 1.4$ |

**Table 6.13:** The single-photon time resolutions for pions measured in the November 2017 test-beam. Due to the high level of pion pollution observed in the proton sample, only results measured with the pion sample are presented. The quoted uncertainties are purely statistical.

| MCP Column | $\sigma_{TORCH}$ Pions (ps) | $\sigma_{TORCH}$ Protons (ps) |
|---|---|---|
| 1 | $110.6 \pm 1.2$ | $112.7 \pm 1.4$ |
| 2 | $101.7 \pm 1.2$ | $110.6 \pm 1.4$ |
| 3 | $101.5 \pm 1.2$ | $110.6 \pm 1.4$ |
| 4 | $105.5 \pm 1.2$ | $106.2 \pm 1.4$ |
| 5 | $83.8 \pm 1.3$ | $91.0 \pm 1.4$ |
| 6 | $101.3 \pm 1.2$ | $103.4 \pm 1.2$ |
| 7 | $90.3 \pm 1.2$ | $87.5 \pm 1.4$ |
| 8 | $112.4 \pm 1.1$ | $102.8 \pm 1.4$ |

**Table 6.14:** The single-photon time resolutions for pions and protons measured in the June 2018 test-beam. The quoted uncertainties are purely statistical.

Preliminary time-resolution measurements are also made for the October 2018 test-beam data, presented in Tables 6.15 and 6.16 for the pion and proton samples respectively, and shown together in Fig. 6.29. Due to the poor quantum efficiency of MCP-PMT A, measurements are only made for MCP-PMT B. These initial results are also reported in Ref. [99].

For all beam positions, the proton samples are seen to provide better resolutions than the pion samples, attributed to pion pollution in the proton sample. Seen in Fig. 6.29, for example at columns 5 and 7, the measured resolutions show a slight variation between MCP-PMT columns. This is suspected to be due to the presence of the time reference channels on the odd-numbered columns. From Fig. 6.25, the range of $y'$ coordinates covered by the 0 and $1^+$ order reflections increases as the

| MCP Column | $\sigma_{TORCH}$ Pos. 1 (ps) | $\sigma_{TORCH}$ Pos. 3 (ps) | $\sigma_{TORCH}$ Pos. 4 (ps) | $\sigma_{TORCH}$ Pos. 5 (ps) |
|---|---|---|---|---|
| 1 | $105.3 \pm 1.4$ | $119.1 \pm 1.5$ | $120.3 \pm 1.6$ | $175.4 \pm 1.6$ |
| 2 | $99.4 \pm 1.3$ | $114.9 \pm 1.8$ | $125.4 \pm 1.6$ | $148.8 \pm 1.5$ |
| 3 | $102.2 \pm 1.4$ | $107.3 \pm 1.7$ | $120.8 \pm 1.6$ | $134.7 \pm 1.5$ |
| 4 | $104.0 \pm 1.4$ | $117.1 \pm 1.5$ | $119.3 \pm 1.7$ | $130.7 \pm 1.5$ |
| 5 | $95.0 \pm 1.4$ | $105.3 \pm 1.8$ | $103.7 \pm 2.0$ | $166.2 \pm 1.6$ |
| 6 | $96.5 \pm 1.4$ | $103.7 \pm 1.6$ | $114.1 \pm 1.6$ | $130.7 \pm 1.4$ |
| 7 | $93.8 \pm 1.3$ | $104.0 \pm 1.5$ | $126.8 \pm 1.8$ | $147.8 \pm 1.5$ |
| 8 | $93.0 \pm 1.3$ | $107.1 \pm 1.7$ | $111.5 \pm 1.6$ | $126.5 \pm 1.4$ |

**Table 6.15:** The single-photon time resolutions measured in the October 2018 test-beam with pion samples, given for the four beam positions close to the side of the radiator plate. Positions 1, 3, 4, and 5 are at distances −175 mm, −488 mm, −801 mm, and −1115 mm from the top of the radiator plate respectively, as per Fig. 6.10.

| MCP Column | $\sigma_{TORCH}$ Pos. 1 (ps) | $\sigma_{TORCH}$ Pos. 3 (ps) | $\sigma_{TORCH}$ Pos. 4 (ps) | $\sigma_{TORCH}$ Pos. 5 (ps) |
|---|---|---|---|---|
| 1 | $103.0 \pm 1.3$ | $114.8 \pm 1.4$ | $116.6 \pm 1.5$ | $170.2 \pm 1.4$ |
| 2 | $95.2 \pm 1.3$ | $101.2 \pm 1.5$ | $122.2 \pm 1.5$ | $138.5 \pm 1.3$ |
| 3 | $95.6 \pm 1.3$ | $102.0 \pm 1.5$ | $114.4 \pm 1.5$ | $126.3 \pm 1.3$ |
| 4 | $99.2 \pm 1.4$ | $110.6 \pm 1.4$ | $115.7 \pm 1.6$ | $123.6 \pm 1.4$ |
| 5 | $108.9 \pm 1.2$ | $100.2 \pm 1.6$ | $95.2 \pm 1.7$ | $163.7 \pm 1.5$ |
| 6 | $94.5 \pm 1.3$ | $97.8 \pm 1.5$ | $108.8 \pm 1.5$ | $124.7 \pm 1.4$ |
| 7 | $89.1 \pm 1.3$ | $97.8 \pm 1.4$ | $115.1 \pm 1.7$ | $146.2 \pm 1.3$ |
| 8 | $90.3 \pm 1.3$ | $108.9 \pm 1.3$ | $98.3 \pm 1.7$ | $115.4 \pm 1.3$ |

**Table 6.16:** The single-photon time resolutions measured in the October 2018 test-beam with proton samples, given for the four beam positions close to the side of the radiator plate. Positions 1, 3, 4, and 5 are at distances −175 mm, −488 mm, −801 mm, and −1115 mm from the top of the radiator plate respectively, as per Fig. 6.10.

beam moves further from the top of the radiator plate. At lower positions, the time reference channel at a $y'$ coordinate of 32 overlaps and interferes with the pattern, hence potentially causing problems when clustering the hits.

The resulting time resolutions are plotted as a function of beam displacement in Fig. 6.30, showing a degradation in measured resolution with increased displacement. Again referring to Fig. 6.25, as expected from chromatic dispersion effects, at lower beam positions a larger range of times are seen to fall across a single $y'$ pixel. This leads to a naturally larger spread in time, and hence a worsening resolution.

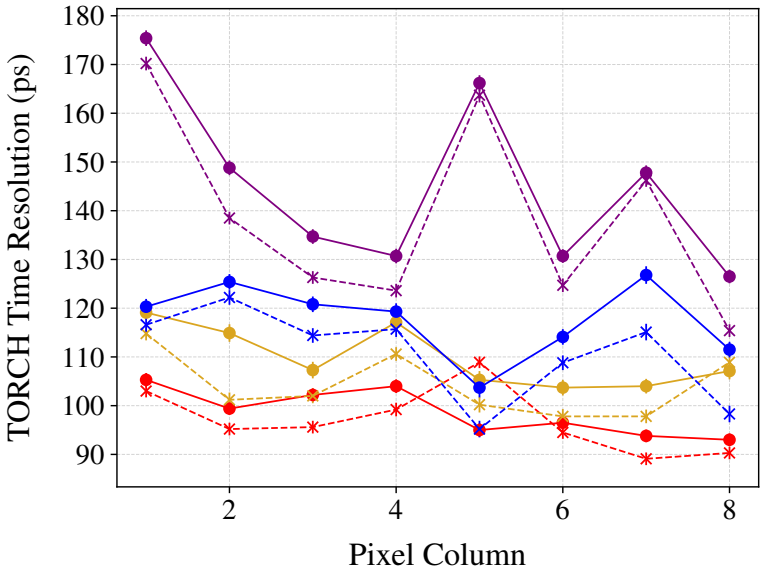



**Figure 6.29:** The single-photon time resolutions measured in the October 2018 test-beam as a function of MCP-PMT column number. The red points show position 1, yellow position 3, blue position 4 and purple position 5. The crossed points connected with dotted lines are measured using proton samples, the dotted points connected with solid lines with pion samples. Positions 1, 3, 4, and 5 are at distances −175 mm, −488 mm, −801 mm, and −1115 mm from the top of the radiator plate respectively, as per Fig. 6.10.

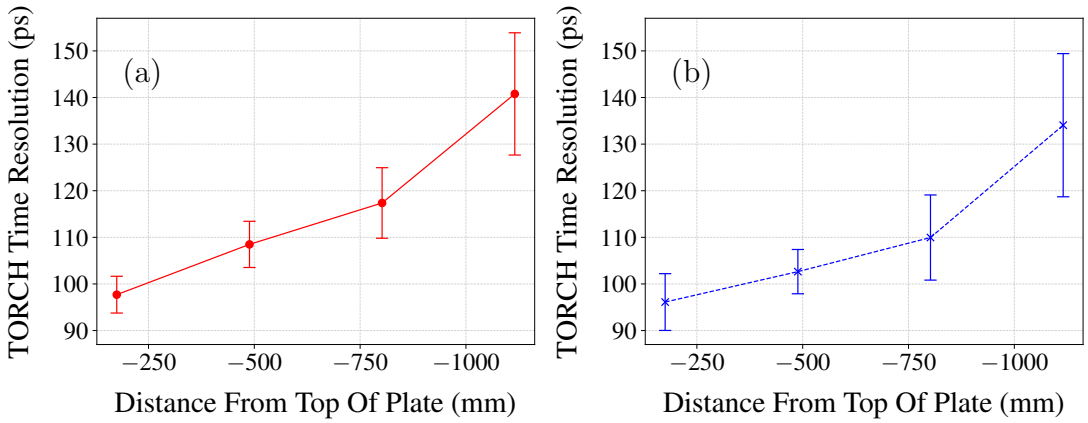



**Figure 6.30:** The average measured single-photon time resolution as a function of displacement from the top of the radiator plate, presented separately for (a) pion and (b) proton samples in the October 2018 test-beam data. The final TORCH modules will be 2500 mm in height, therefore extrapolating indicates time resolutions of $\sim 180$ ps could be expected for photons which travel the entire height of a module.

## 6.9 Photon Counting

Photon yields have been presented elsewhere for the November 2017 and June 2018 data, and can be found in Ref. [98], with further details provided in Ref. [112]. For the October 2018 data, the number of photons observed is compared to the number expected from the simulation described in Section 6.4. In data, photons detected by both MCP-PMT A and B are used, and the quantum efficiency and instrumented area is set in the simulation accordingly. An identical timing window is also applied to both data and simulation to ensure only photons falling in reconstructable reflections are counted, and rejecting photons which appear later than those from the order $0'$ reflection. This cut is an update compared to the results presented in Ref. [99].

The number of photons observed per incident particle is shown in Fig. 6.31 for various beam positions, and the corresponding mean values in Table 6.17. With the exception of position 3, a $\sim 20\,\%$ loss is observed for all positions. The measured yields have a strong dependence on the MCP-PMT gain, which has not been measured for either MCP-PMT A or B, and so the nominal data sheet values are used. Similarly, the NINO thresholds have a large effect on the yields, and although best estimates were used in the simulation, they are not accurately known.

| Position | Data Mean (Photons) | Simulation Mean (Photons) | Ratio |
|---|---|---|---|
| 1 | $2.751 \pm 0.008$ | $3.26 \pm 0.11$ | $0.84 \pm 0.03$ |
| 3 | $1.661 \pm 0.006$ | $2.81 \pm 0.10$ | $0.59 \pm 0.02$ |
| 4 | $1.220 \pm 0.005$ | $1.57 \pm 0.08$ | $0.78 \pm 0.04$ |
| 5 | $1.183 \pm 0.003$ | $1.50 \pm 0.08$ | $0.79 \pm 0.04$ |
| 6 | $1.282 \pm 0.003$ | $1.61 \pm 0.08$ | $0.80 \pm 0.04$ |
| 7 | $1.230 \pm 0.004$ | $1.65 \pm 0.08$ | $0.75 \pm 0.04$ |

**Table 6.17:** The mean number of photons observed in data and simulation for different beam positions in October 2018. The ratio of means is also given. The quoted errors are purely statistical.

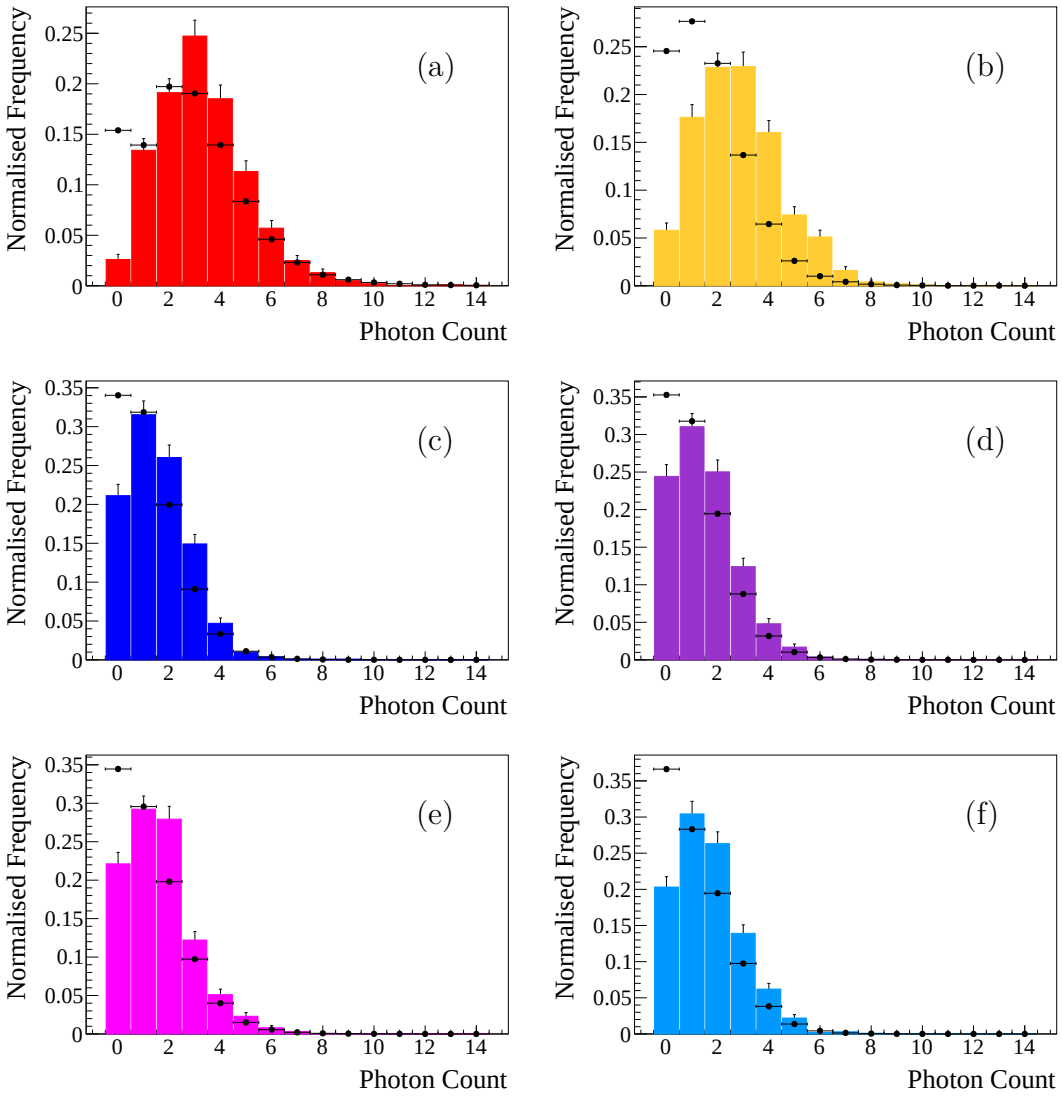



**Figure 6.31:** The number of photons observed in data (points) and simulation (histograms) for position (a) 1, (b) 3, (c) 4, (d) 5, (e) 6, and (f) 7. The corresponding vertical displacements from the MCP-PMT are (a) −175.0 mm, (b) −488.6 mm, (c) −801.9 mm, and (d)–(f) −1115.4 mm. The means of the distributions are given in Table 6.17.

Figure 6.32 shows the mean number of photons as a function of beam displacement from the top of the plate. Given the same shape of data and simulation, this would suggest the decrease is not primarily due to attenuation resulting from the additional propagation distance, but rather caused by acceptance effects from the range of photon angles and energies incident on the MCP-PMTs. With a complete array of detectors, this phenomenon would not be present, however as only two MCP-PMTs are instrumented, photons emitted at lower positions fall outside the detector area.

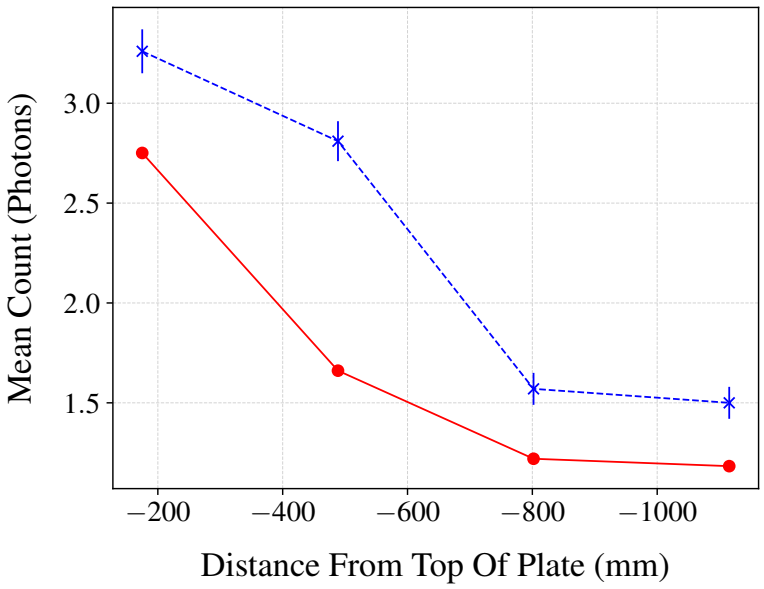



**Figure 6.32:** The mean number of photons observed per event as a function of vertical displacement from the top of the radiator plate, for data (red, dots) and simulation (blue, crosses) corresponding to the October 2018 test-beam.

## 6.10 Potential Improvements

### 6.10.1 Single-Photon Time Resolution

Single-photon time resolutions below 100 ps have been measured, however none have yet reached the desired 70 ps. To reach this level, corrections for the amount of charge sampled by the NINO chips need to be applied. With the use of a dedicated calibration system, a relationship between the width of the digitized signal and the input charge can be determined, allowing for more precise centroiding of the

photon clusters in both space and time. At the time of writing such a system is being developed and is planned to calibrate the NINO test-beam boards, to be used in later versions of the October 2018 analysis. This would provide a significant step in offsetting the issue presented by the naturally larger spread in time seen at positions further from the top of the plate in the October 2018 data. By improving the centroiding of the cluster, the ideal effective granularity of 128 $y'$ pixels should be achieved, hence reducing the time spread per pixel.

### 6.10.2 Photon Counting

The goal for a final TORCH module, fully instrumented with 11 MCP-PMTs, is to detect 30 photons per incident particle. Even taking the largest measured yield per particle from the October 2018 test-beam analysis and scaling for the missing MCP-PMTs, the current expectation would be $\sim 15$ photons. The MCP-PMT quantum efficiencies have a big influence here: MCP-PMT A is very low compared to the MCP-PMTs planned for installation in the final detector, and MCP-PMT B is also lower than desirable. Improvements in the peak quantum efficiencies up to $\sim 25 - 30\%$ (absolute values) should therefore bring the total number of photons close to the final expectation. Further improvements could also be achieved by having a deeper radiator plate or modifying the range of angles accepted by the focusing block, and study of optimised performance versus cost will be the subject of future work.

## 6.11 Conclusions

Three test-beam campaigns have been carried out, two with a small-scale demonstrator and one with a half-sized LHCb module. Single-photon time resolutions better than 100 ps have been achieved, and indications of future work required to reach the target 70 ps have been discussed. The photon-counting efficiency of the half-sized demonstrator is also presented. A deficit of photons compared to the LHCb benchmark is seen, however with MCP-PMTs of improved quantum efficiency, photon yields close to requirement should be achievable.

# 7 Conclusions

## Contents



## 7.1 The Search for $B \to D\mu^+\mu^-$ Decays

Searches for seven decays of the type $B \to D\mu^+\mu^-$ have been made with $8.7\,\text{fb}^{-1}$ of LHCb data, and the branching ratios calculated by comparison of the signal mode to a normalisation channel. The selection efficiencies of the signal and normalisation modes were measured using dedicated Monte Carlo samples and corrections were applied to account for differences between the simulation and real data.

No signals were observed for any of the modes, and so upper limits on the branching ratios have been set at the 95% confidence level, as presented in Table 7.1. All seven upper limits are world-leading.

| Decay | Upper Limit |
|---|---|
| $\mathcal{B}\left(B^0 \to \overline{D}^0\mu^+\mu^-\right)$ | $< 6.1 \times 10^{-8}$ |
| $\mathcal{B}\left(B^0 \to \overline{D}^0 J/\psi\right)$ | $< 1.4 \times 10^{-6}$ |
| $\mathcal{B}\left(B_s^0 \to \overline{D}^0\mu^+\mu^-\right)$ | $< 1.8 \times 10^{-7}$ |
| $\mathcal{B}\left(B_s^0 \to \overline{D}^0 J/\psi\right)$ | $< 1.4 \times 10^{-6}$ |
| $\mathcal{B}\left(B^+ \to D_s^+\mu^+\mu^-\right)$ | $< 3.2 \times 10^{-8}$ |
| $\mathcal{B}\left(B^+ \to D_s^+ J/\psi\right)$ | $< 3.5 \times 10^{-7}$ |
| $\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+\mu^+\mu^-\right)$ | $< 9.6 \times 10^{-8}$ |

**Table 7.1:** The upper limits set for each $B \to D\mu^+\mu^-$ mode at the 95% confidence level. $f_c$ and $f_u$ are the probabilities that a $b$ quark will hadronize into a $B_c^+$ or $B^+$ mesons respectively.

Additionally, the branching ratio of the previously-observed $B_c^+ \to D_s^+ J/\psi$ decay has been measured to be

$$\frac{f_c}{f_u}.\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right) = (1.63 \pm 0.15 \pm 0.13) \times 10^{-5}\,,$$

where the first error is statistical and the second systematic. This value is consistent with the current particle data group average of $(2.00 \pm 0.33) \times 10^{-5}$ [28, 53], determined by combining measurements from ATLAS and LHCb.

The $B_c^+ \to D_s^{*+} J/\psi$ decay is also observed in partially reconstructed form, in a ratio $1.9 \pm 0.2$ with respect to $B_c^+ \to D_s^+ J/\psi$ decays. The $B_c^+ \to D_s^{*+} J/\psi$ decay can occur with differing helicity amplitudes, and the ratio of $\mathcal{A}^{\pm\pm}$ helicities relative to the total is observed to be

$$f_{\mathcal{A}^{\pm\pm}} = \frac{N_{\mathcal{A}^{\pm\pm}}}{N_{\mathcal{A}^{\pm\pm}} + N_{\mathcal{A}^{00}}} = 50 \pm 10\,\%\,. \tag{7.1}$$

The errors quoted on both the above measurements are statistical only.

In summary, six upper limits have been set on previously unstudied decays, the limit on $B^0 \to \overline{D}^0 J/\psi$ has been improved by an order of magnitude, and another measurement of $(f_c/f_u).\mathcal{B}\left(B_c^+ \to D_s^+ J/\psi\right)$ has been made which sits below the current PDG average of $(2.00 \pm 0.33) \times 10^{-5}$, a combination of ATLAS and LHCb measurements.

## 7.2 Test-beam Studies of TORCH Detector Prototypes

A description of TORCH, a novel time-of-flight detector proposed for Upgrades Ib and II of the LHCb experiment, has been presented. To determine whether TORCH can meet the specifications suitable for its inclusion in LHCb, three instrumental test-beam campaigns have been undertaken. A small-scale demonstrator was instrumented in the first two test-beam periods, and a half-sized LHCb module instrumented in the third. Data were recorded with a mixed proton-pion beam with momenta of 3.5, 5 and 8 GeV/$c$; corrections were applied to account for non-linearity and time slew in the digitisation of the Cherenkov photon signals recorded by customised micro-channel photomultiplier detectors. Additionally, a clustering algorithm has been run to optimise the charge-sharing properties of the MCP-PMTs, improving the spatial and temporal resolutions.

The detected photon arrival times have been compared to expectations from an offline reconstruction and the resultant residual distributions constructed. Subtracting contributions from the time reference and the spatial spread of the beam, single-photon time resolutions better than 100 ps have been measured. Plans of future work required to reach the target 70 ps have then been discussed, involving making dedicated measurements of the output-signal width relative to the input charge in the digitisation electronics. This will improve the cluster centroiding, hence reducing the time spread per pixel.

The photon-counting efficiency of the half-sized demonstrator has also been presented and compared to simulation. A deficit of photons compared to expectations at LHCb has been observed, however, with MCP-PMT photon detectors of improved quantum efficiency, photon yields close to requirement are anticipated.

In summary, measurements made with the prototype TORCH detectors continue to approach LHCb requirements, showing significant improvements upon previous campaigns. TORCH should be a valuable subdetector for low-momentum particle identification within LHCb.

# Appendices

# A

# Parameterisations of Physics Distributions

## Contents



## A.1 Introduction

This appendix collects in a common place the various functions which are used to parameterise the physical distributions modelled in this thesis. Table A.1 lists all the functions used, and a brief summary of each is given the subsequent sections.

| Shape | Section | Form |
|---|---|---|
| Gaussian | A.2 | $G(x;\mu,\sigma)$ |
| Crystal-Ball | A.3 | $CB(x;\mu,\sigma,\alpha,n)$ |
| Double Crystal-Ball | A.3 | $DCB(x;\mu,\sigma_1,\sigma_2,\alpha,n,R)$ |
| Hypatia | A.4 | $I(x;\mu,\sigma,\lambda,\zeta,\beta,\alpha,n)$ |
| Two-tailed Hypatia | A.4 | $I_2(x;\mu,\sigma,\lambda,\zeta,\beta,\alpha_l,n_l,\alpha_r,n_r)$ |
| Exponential | A.5 | $E(x;\lambda)$ |
| Chebychev Polynomial * | A.6 | $T_n(x;a_1,a_2,\ldots,a_n)$ |
| Horns/Hill Analytic PDFs | A.7 | $\mathcal{P}_j^p(x;a,b,\xi,\delta,\sigma,R_\sigma,f_G)$ |
| Kernal Density Estimation | A.8 | $\hat{f}(x)$ |

**Table A.1:** Summary of the functions used to parametrise the physical distributions modelled in this thesis. (*) Specifically Chebychev polynomials of the first kind are being referred to.

## A.2 The Gaussian

In cases with a large number of trials, the Gaussian distribution, $G(x;\mu,\sigma)$, can be used to approximate the binomial distribution. This makes the Gaussian well-suited to model distributions in particle physics where there are very large numbers of events. The distribution is given by

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}}.\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \tag{A.1}$$

where $\mu$ is the mean of the distribution and $\sigma$ its standard deviation.

## A.3 The Crystal Ball Function

The Crystal Ball function, $CB(x;\mu,\sigma,\alpha,n)$, consists of a Gaussian with mean $\mu$ and width $\sigma$ which transitions to a power-law tail ($x^{-n}$) with exponent $n$ at transition point $\alpha$ [111]. The power-law tail commonly models losses of energy due to undetected decay products. In this thesis the Crystal Ball function is also used to account for backscattering of primary photoelectrons in MCP-PMTs.

In Chapter 4, a double Crystal Ball shape with shared tail parameters is used. Here, the mean, transition point and power-law exponent are all shared between the shapes, but the core Gaussian components have different widths.

$R$ is the ratio between the amplitudes of the two shapes, giving the final distribution $DCB\left(x;\mu,\sigma_1,\sigma_2,\alpha,n,R\right)$, where $\sigma_1$ and $\sigma_2$ are the widths of the two Gaussian components.

## A.4 The Hypatia Distribution

The Hypatia function is a Crystal-Ball-like function, but with a hyperbolic distribution as the core instead of a Gaussian. The originally-presented distribution, $I\left(x;\mu,\sigma,\lambda,\zeta,\beta,\alpha,n\right)$ [61] , has a single tail, however the Hypatia distribution used in this thesis has a power-law tail on both sides, giving it nine free parameters: $I_2\left(x;\mu,\sigma,\lambda,\zeta,\beta,\alpha_l,n_l,\alpha_r,n_r\right)$, where $x$ is the free variable, $\mu$ is the mean of the hyperbolic distribution, $\sigma$ corresponds the resolution, and $\lambda$ and $\zeta$ are two further parameters of the hyperbolic distribution. $\beta$ introduces a variance-dependence to the mean, with a value of 0 creating a symmetric hyperbolic distribution. $\alpha_{l/r}$ is the transition point to the power law tail for the left/right side, and $n_{l/r}$ is the power-law exponent of the corresponding tail.

## A.5 The Exponential

Exponential distributions, $E\left(x;\lambda\right)$, are commonly used to model backgrounds in particle physics analyses. This is simply an exponential with (in this thesis) a negative exponent, $E\left(x;\lambda\right) = e^{-\lambda x}$.

## A.6 Chebychev Polynomials (of the first kind)

There are two sequences of Chebychev polynomials [113] , so called polynomials of the first and second kind. This thesis makes used of the first kind, $T_n\left(x; a_1, a_2, \ldots, a_n\right)$, defined by the sequence

$$\begin{aligned} T_0\left(x\right) &= 1\,, \\ T_1\left(x\right) &= x\,, \\ T_{n+1}\left(x\right) &= 2xT_n\left(x\right) - T_{n-1}\left(x\right)\,, \end{aligned} \tag{A.2}$$

and where $a_1, a_2, \ldots, a_n$ are multiplicative coefficients applied to $T_1, T_2, \ldots, T_n$. It is assumed the coefficient associated with the $0^{\text{th}}$ polynomial is unity.

## A.7 The Horns and Hill Analytic PDFs

A set of analytic PDFs are used in Chapter 4 to model partially reconstructed $B_c^+ \to D_s^{*+} J/\psi$ decays, where $D_s^{*+} \to D_s^+ \left(\gamma/\pi^0\right)$ and the $\gamma$ or $\pi^0$ is not reconstructed. Three Analytic PDFs are used: Horns, Hill and Little Horns. These shapes are given by $\mathcal{P}_j^p\left(x; a, b, \xi, \delta, \sigma, R_\sigma, f_G\right)$, where the $p$ index represents the missing particle ($\gamma$ or $\pi^0$), and the $j$ index represents the helicity amplitude ($\mathcal{A}^{\pm\pm}$ or $\mathcal{A}^{00}$). The remaining parameters correspond to three sets of components which are convolved to produce the final shape;

- $a$ and $b$, the upper and lower kinematic endpoints respectively, bounding a quadratic which describes the helicity structure;
- $\delta$, $\sigma$, $R_\sigma$, and $f_G$ parametrise a double Gaussian which describes the finite detector resolution, where $\delta$ is the relative shift between the two Gaussians, $\sigma$ is the core width, $R_\sigma$ the width ratio, and $f_G$ the relative fraction of amplitudes between the two; and
- $\xi$ is the slope of a linear function which describes selection effects.

These PDFs are described in further detail in Ref. [114].

## A.8 Kernal Density Estimation

Rather than parameterise a distribution, kernal density estimation [63] allows the probability density function of the distribution to be estimated. This function can then be used as a shape in a fit to data.

Kernal density estimation approximates the distribution as a superposition of $n$ kernals, where n is the number of data points in the distribution. Each contributes $1/n$ to the density function integral. In this thesis, the Gaussian shape is used as the kernal, resulting in

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} G(x - x_i) \ , \tag{A.3}$$

where $\hat{f}$ is the estimation ($f$ is the function being estimated), and $x_i$ the value of the distribution at point $i$.

# B

# Photon Path Reconstruction in TORCH

## Contents



## B.1 Introduction

Included with the main purpose of aiding future TORCH collaborators, this appendix provides a derivation of the path length of a photon generated within TORCH, from the point of emission to the detector. Note: this derivation follows that used in code to generate the reconstruction described in Chapter 6. Simpler derivations are possible.

## B.2 Definition of constants

Constants are written as capital letters or with capital subscripts in the case of angles.

- Let the radiator plate have width $W$, height $H$ and depth $D$.
- Let the angle of the wedge on top of the radiator plate be $\theta_W = 0.628^c = 36°$.
- Let the radius of curvature of the focusing block mirror be $R = 260\,\text{mm}$, the distance from the axis of the cylindrical face to the focusing block be $C = 133.7\,\text{mm}$, and the distance in the $y'$ direction between the top of the radiator plate and perpendicular line which passes through the cylindrical axis $P = 6.691\,\text{mm}$. These parameters can be visualised in Fig. B.3.
- Let the width of the MCP-PMT window be $G$.

## B.3 Definition of initial parameters

The origin is defined as the vertical and horizontal centre of the front face of the radiator plate, with the axes defined the same as in Chapter 6. Relative to this, let $\vec{r_0} = (x_0, y_0, z_0)$ to be the emission point of the photon, with direction $\vec{k} = (k_x, k_y, k_z)$, where $|\vec{k}| = 1$.

Let the detected position of the photon, relative to the radiator plate (as opposed to in the detector plane) be $\vec{r} = (x, y, z)$.

Three angles are also defined, shown in Fig. B.1, and can be specified in terms of the initial parameters:

$$\tan(\theta_x) = \frac{k_y}{k_x}\,, \tag{B.1}$$

$$\tan(\theta_z) = \frac{k_z}{k_y}\,, \tag{B.2}$$

and

$$\theta'_z = \theta_W + \theta_z\,. \tag{B.3}$$

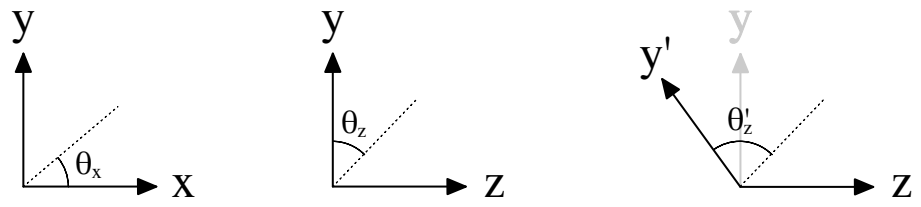



**Figure B.1:** Definition of $\theta_x$, $\theta_z$, and $\theta'_z$. The dotted line in the right-most diagram shows how $y$ relates to $y'$.

## B.4 Path Length in the Radiator Plate

### B.4.1 Main body of the radiator plate

To find the path length within the main body of the radiator plate, $L_p$, consider the $y$-direction only,

$$L_p = \frac{H}{2} - \frac{D}{\tan(\theta_W)} + \begin{cases} -y_0, & k_y > 0 \\ y_0, & k_y < 0 \end{cases}. \tag{B.4}$$

Here, the length from the origin to the bottom of the wedge is given by the first two terms, and then the length adjusted to account for the point of emission through the addition or subtraction of $y_0$. To get the total path length, the distance needs to be scaled by $k_y$, however this factor will be applied at the end when all components are combined into the final path length.

### B.4.2 Wedge at the top of the radiator plate

To determine the length of the path through the wedge at the top of the radiator plate, the position in the $z$-direction where the photon crosses into the wedge, $z_w$, needs to be known, as well as the $z$-direction of travel. This can be calculated by considering the number of reflections the photon makes off the front and back faces, $N_z$, which is in turn derived from $L_p$,

$$N_t = \frac{L_p}{k_y}, \tag{B.5}$$

$$N_z = \frac{N_t k_z + z_0}{D}. \tag{B.6}$$

With the number of reflections known, the sign of the direction, $d$, and crossing point can be found from,

$$d = \begin{cases} -1, & \lfloor N_z \rfloor \text{ is even} \\ +1, & \lfloor N_z \rfloor \text{ is odd} \end{cases}, \tag{B.7}$$

where $\lfloor x \rfloor$ is the floor function, and

$$z_w = \{N_z\} \times D\,, \tag{B.8}$$

where $\{x\}$ is the fractional part defined by $\{x\} = x - \lfloor x \rfloor$.

With the direction and crossing point known, the geometry of the wedge used to derive the path length, $L_w$, as shown in Fig. B.2.

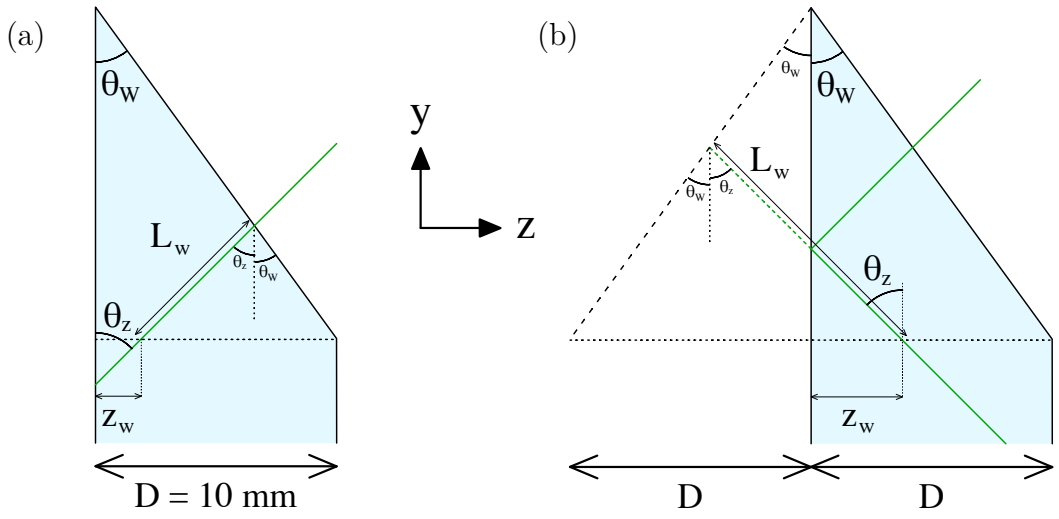



**Figure B.2:** The geometry of the radiator plate wedge and the quantities used to derive the path length of a photon within the wedge for (a) forward and (b) backwards travelling photons.

Beginning with Fig. B.2a, the sine rule can be used to derive $L_w$,

$$\frac{L_w}{\sin\left(\frac{\pi}{2} - \theta_w\right)} = \frac{D - z_w}{\sin\left(\theta_z + \theta_w\right)}\,. \tag{B.9}$$

Considering now Fig. B.2b, it can be seen that the same construction allows $L_w$ to be found, except the horizontal length is given by $D + z_w$. Hence, the horizontal difference is given by $D + dz_w$, with $d$ accounting for the sign change. Converting $\sin\left(\frac{\pi}{2} - \theta_w\right)$ to $\cos\left(\theta_w\right)$ and rearranging gives,

$$L = \frac{\cos\left(\theta_w\right)}{\sin\left(\theta_W + \theta_z\right)}\left(D + dz_w\right). \tag{B.10}$$

Finally, when combining this length with the length measured in the radiator plate, the fact that it is calculated in the $y$-$z$ projection must be accounted for. Similarly as for $L_p$, this will be applied at the end in the total path length calculation.

## B.5 Path of photons in the focusing block to the detector

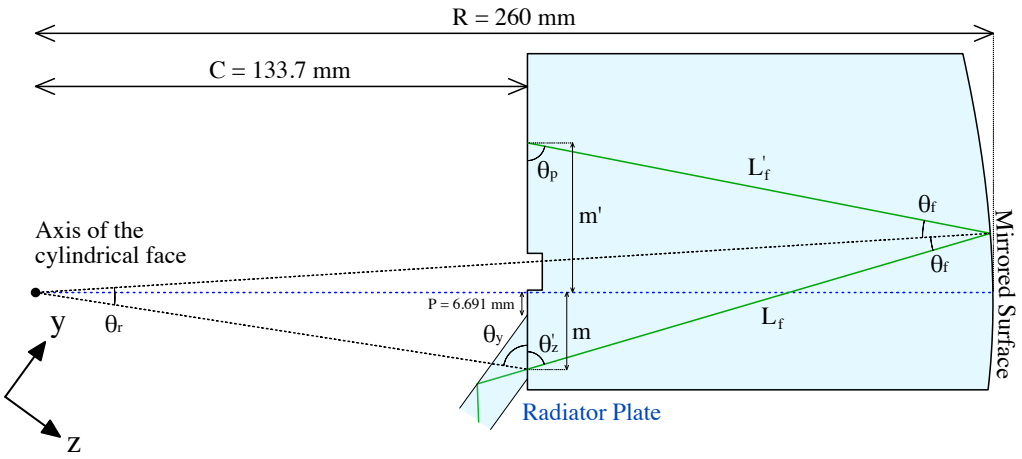



**Figure B.3:** The geometry of the focusing block, with the angles and measurements used in the path-length derivation defined.

$\theta_y$ is determined by considering the right-angled triangle formed by the opposite line of length $C$ and the adjacent line of length $m$, giving

$$\tan\left(\theta_y\right) = \frac{C}{y}\,. \tag{B.11}$$

Knowledge of $\theta_y$ combined with $\theta'_z$ allows the sine rule to be used to find $\theta_f$,

$$\sin\left(\theta_f\right) = \frac{\sqrt{y^2 + C^2}}{R} \times \sin\left(\theta_y + \theta'_z\right). \tag{B.12}$$

$\theta_r$ is then determined by the sum of the internal angles in the triangle formed between the entry point, the axis of the cylindrical face and the point of reflection at the mirror,

$$\theta_r = \pi - \left(\theta_y + \theta'_z + \theta_f\right), \tag{B.13}$$

and similarly for $\theta_p$,

$$\theta_p = \pi - (2\theta_f + \theta'_z)\,. \tag{B.14}$$

With all the relevant angles found, the sine rule can be used to determine $L_f$ and $L'_f$,

$$L_f = \frac{\sin(\theta_R)}{\sin(\theta_y + \theta'_z)} \times R\,, \tag{B.15}$$

and

$$L'_f = \frac{\sin(\theta'_z)}{\sin(\theta_p)} \times L\,. \tag{B.16}$$

Eliminating $\theta_p$ and $\theta_r$ gives,

$$L_f = \frac{\sin(\theta_y + \theta'_z + \theta_f)}{\sin(\theta_y + \theta'_z)} \times R\,, \tag{B.17}$$

and

$$L'_f = \frac{\sin(\theta'_z)}{\sin(2\theta_f + \theta'_z)} \times L\,. \tag{B.18}$$

In addition to the path length in the focusing block, the width of the MCP-PMT window also needs to be accounted for. Using the angle $\theta_p$ this is derived through simple trigonometry,

$$L_{mcp} = \frac{G}{\sin(\theta_p)} = \frac{G}{\sin(2\theta_f + \theta'_z)}\,. \tag{B.19}$$

## B.6 Total Path Length

The total path length can now be found by combining all the previously determined components. As each component was derived in the $y$-direction or $y$-$z$-plane, this must be accounted for by scaling according to the initial photon direction. Thus, the final path length, $L$, is given by,

$$L = \frac{L_p}{k_y} + \frac{L_w + L_f + L'_f + L_{mcp}}{\sqrt{k_y^2 + k_z^2}}\,. \tag{B.20}$$

## B.7 Total $x$-component

Although the path length has been determined, it is often necessary to know the total $x$-component of the photon path, and so it is derived here.

Considering the path in the $x$-$y$ plane only, the $x$-component of the path length is most easily determined by first considering the case that the photon is both emitted and detected with $x = 0$. In this case, the $x$-component is $Nw$, where N is the number of side reflections, as each reflection requires the width of the plate to be traversed, half the width there and half back.

Moving to a case where the photon not emitted in the centre adds or subtracts a factor $x_0$, where the sign is the opposite to that $\theta_x$,

$$-\operatorname{sgn}(\theta_x)\, x_0\,. \tag{B.21}$$

Similarly for the case where the position the photon is detected is displaced from $x = 0$, an additional factor of $x$ must be included, however the sign now matches the direction at detection. This means it depends on the number of side reflections and the sign of $\theta_x$,

$$\operatorname{sgn}(\theta_x)\, x \times \begin{cases} 1, & N \text{ is even} \\ -1, & N \text{ is odd} \end{cases}. \tag{B.22}$$

Combining these cases gives

$$X = Nw - \operatorname{sgn}(\theta_x)\, x_0 + \operatorname{sgn}(\theta_x)\, x \times \begin{cases} 1, & N \text{ is even} \\ -1, & N \text{ is odd} \end{cases}, \tag{B.23}$$

where $X$ is the total $x$-component of the photon path. This can be simplfied to,

$$X_{N\pm} = Nw - \operatorname{sgn}(\theta_x) \begin{cases} (x_0 - x), & N \text{ is even} \\ (x_0 + x), & N \text{ is odd} \end{cases}, \tag{B.24}$$

where $N_\pm$ dictates the number of reflections and sign of $\theta_x$, matching the index of the path as defined in Chapter 6.

# References


[1] R.-D. Heuer, *CERN and 60 years of science for peace*, AIP Conference Proceedings **1645** (2015) 430.

[2] E. Mobs, *The CERN accelerator complex - August 2018*, `http://cds.cern.ch/record/2636343`, 2018. Accessed: 11-03-2020.

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

[4] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. A **30** (2015) 1530022, `arXiv:1412.6352`.

[5] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`. Accessed: 11-03-2020.

[6] H. Dijkstra, *CP violation prospects at the LHC*, Brazilian Journal of Physics **34** (2004) 1295.

[7] M. Alexander *et al.*, *Mapping the material in the LHCb vertex locator using secondary hadronic interactions*, JINST **13** (2018) P06008, `arXiv:1803.07466`.

[8] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[9] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[10] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. C **73** (2013) 2431, `arXiv:1211.6759`.

[11] A. Papanestis *et al.*, *Performance of the LHCb RICH detectors during the LHC Run II*, Nucl. Inst. & Meth. in Phys. Res. A **876** (2017) 221, `arXiv:1703.08152`.

[12] Y. Guz, *The LHCb Calorimeter system: design, performance and upgrade*, JINST **12** (2017) C07024.

[13] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[14] F. P. Albicocco *et al.*, *Long-term Operation of the Multi-Wire-Proportional-Chambers of the LHCb Muon System*, Tech. Rep. arXiv:1908.02178, CERN, Geneva, 2019.

[15] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, `arXiv:1812.10790`.

[16] LHCb, J. Albrecht, *The LHCb Trigger System: Present and Future*, J. Phys. Conf. Ser. **623** (2015) 012003.

[17] M. Williams *et al.*, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002, CERN, Geneva, 2011.

[18] LHCb collaboration, R. Aaij *et al.*, *Observation of $B_c^+ \to J/\psi\, D_s^+$ and $B_c^+ \to J/\psi\, D_s^{*+}$ decays*, Phys. Rev. **D87** (2013) 112012.

[19] G. Aad *et al.*, *Study of the $B_c^+ \to J/\psi D_s^+$ and $B_c^+ \to J/\psi D_s^{*+}$ decays with the ATLAS detector*, The European Physical Journal C **76** (2016) 4.

[20] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the $B_c^+$ meson mass*, JHEP **7** (2020) 123, arXiv:2004.08163.

[21] P. A. M. Dirac and R. H. Fowler, *The quantum theory of the electron*, Proc. R. Soc. Lond. **117** (1928) 610.

[22] R. P. Feynman and M. Gell-Mann, *Theory of the Fermi Interaction*, Phys. Rev. **109** (1958) 193.

[23] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[24] G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) 1.

[25] S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) 30.

[26] D. Griffiths, *Introduction to Elementary Particles*, Joh Wiley & Sons, Ltd, 2004.

[27] A. D. Sakharov, *Quark-Muonic Currents and Violation of CP Invariance*, JETP Lett. (USSR) **5** (1967), (Engl. Transl.).

[28] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[29] B. R. Martin, *Nuclear and Particle Physics, An Introduction*, Joh Wiley & Sons, Ltd, 2006.

[30] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.

[31] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[32] S. L. Glashow *et al.*, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

[33] B. J. Bjørken and S. L. Glashow, *Elementary particles and SU(4)*, Physics Letters **11** (1964) 255.

[34] G. S. Abrams *et al.*, *Discovery of a Second Narrow Resonance in $e^+e^-$ Annihilation*, Phys. Rev. Lett. **33** (1974) 1453.

[35] J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[36] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.

[37] H. Fritzsch *et al.*, *Advantages of the color octet gluon picture*, Physics Letters B **47** (1973) 365.

[38] G. Altarelli, *Collider Physics within the Standard Model, A Primer*, Springer, 2017.

[39] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books Publishing, L.L.C., 1995.

[40] L. Lellouch *et al.*, *Modern Perspectives in Lattice QCD: Quantum Field Theory and High Performance Computing: Lecture Notes of the Les Houches Summer School*, vol. 93 of *Graduate texts in physics*, Oxford University Press, 2009.

[41] E. Eichten and B. Hill, *An effective field theory for the calculation of matrix elements involving heavy quarks*, Physics Letters B **234** (1990) 511.

[42] N. Isgur and M. B. Wise, *Weak decays of heavy mesons in the static quark approximation*, Physics Letters B **232** (1989) 113.

[43] F. Bernardoni *et al.*, *Decay constants of B-mesons from non-perturbative HQET with two light dynamical quarks*, Physics Letters B **735** (2014) 349.

[44] K. Lingel *et al.*, *Penguin decays of B mesons*, Ann. Rev. Nucl. Part. Sci. **48** (1998) 253.

[45] CDF Collaboration, A. Abulencia *et al.*, *Search for $Z' \to e^+e^-$ Using Dielectron Mass and Angular Distribution*, Phys. Rev. Lett. **96** (2006) 211801.

[46] J. Iizuka, *A Systematics and Phenomenology of Meson Family*, Progress of Theoretical Physics Supplement **37-38** (1966) 21.

[47] S. Okubo, *$\phi$-meson and unitary symmetry model*, Phys. Lett. **5** (1963) 165.

[48] D. H. Evans *et al.*, *Short distance analysis of $\bar{B} \to D^{(*)0}e^+e^-$ and $\bar{B} \to J/\psi e^+e^-$*, Nuclear Physics B **577** (2000) 240.

[49] C. S. Kim *et al.*, *Study of $\bar{B}^0 \to D^0\mu^+\mu^-$ decay in perturbative QCD approach*, Journal of High Energy Physics **2011** (2011) 152.

[50] D. H. Evans *et al.*, *Operator Product Expansion for Exclusive Decays $B^+ \to D_s^+e^+e^-$ and $B^+ \to D_s^{*+}e^+e^-$*, Phys. Rev. Lett. **83** (1999) 4947.

[51] BABAR Collaboration, B. Aubert *et al.*, *Search for $B \to J/\psi D$ decays*, Phys. Rev. D **71** (2005) 091103.

[52] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio $f_s/f_d$ and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[53] LHCb collaboration, R. Aaij *et al.*, *Measurement of $B_c^+$ production in proton-proton collisions at $\sqrt{s}$ =8 TeV*, Phys. Rev. Lett. **114** (2015) 132001, arXiv:1411.2943.

[54] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[55] S. Tolk *et al.*, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039, CERN, Geneva, 2014.

[56] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, p. 79, 2003. arXiv:physics/0308063.

[57] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[58] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13*TeV *pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

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

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

[61] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. A **764** (2014) 150, arXiv:1312.5000.

[62] G. A. Cowan *et al.*, *RapidSim: An application for the fast simulation of heavy-quark hadron decays*, Computer Physics Communications **214** (2017) 239, arXiv:1612.07489.

[63] K. Cranmer, *Kernel estimation in high-energy physics*, Computer Physics Communications **136** (2001) 198, arXiv:hep-ex/0011057.

[64] LHCb collaboration, R. Aaij *et al.*, *First observation of $\overline{B}{}_s^0 \to D_{s2}^{*+} X \mu^- \overline{\nu}$ decays*, Phys. Lett. **B698** (2011) 14, arXiv:1102.0348.

[65] A. L. Read, *Presentation of search results: The $CL_s$ technique*, J. Phys. **G28** (2002) 2693.

[66] L. Moneta *et al.*, *The RooStats project*, PoS(ACAT2010) **093** (2011) 057, arXiv:1009.1003.

[67] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

[68] R. Aaij *et al.*, *Searches for rare decays of the form $B \to D\mu^+\mu^-$*, (2021). (in preparation).

[69] N. H. Brook *et al.*, *Testbeam studies of a TORCH prototype detector*, Nucl. Inst. & Meth. in Phys. Res. A **908** (2018) 256, arXiv:1805.04849.

[70] LHCb collaboration, *PID Plots for Conference*, https://twiki.cern.ch/twiki/bin/view/LHCb/PIDConferencePlots. Accessed: 14-07-2020.

[71] P. A. Cherenkov, *Visible emission of clean liquids by action of $\gamma$ radiation*, C. R. Acad. Sci. USSR **2** (1934) 451.

[72] P. A. Cherenkov, *Influence of magnetic field on the observed luminescence of fluids induced by $\gamma$ rays*, C. R. Acad. Sci. USSR **3** (1936) 413.

[73] W. Allison, *Fundamental Physics for Probing and Imaging*, Oxford University Press, 2006.

[74] I. M. Frank and I. E. Tamm, *Coherent Visible Radiation of Fast Electrons Passing Through Matter*, C. R. Acad. Sci. USSR **14** (1937) 109.

[75] M. J. Charles and R. Forty, *TORCH: Time of flight identification with Cherenkov radiation*, Nucl. Inst. & Meth. in Phys. Res. A **639** (2011) 173, arXiv:1009.3793.

[76] T. Abe *et al.*, *Belle II Technical Design Report*, Tech. Rep. 2010-1, KEK, Tsukuba, 2010. arXiv:1011.0352.

[77] I. Adam *et al.*, *The DIRC Particle Identification System for the BABAR Experiment*, Nucl. Inst. & Meth. in Phys. Res. A **538** (2005) 281.

[78] J. Fast, *The Belle II imaging Time-of-Propagation (iTOP) detector*, Nucl. Inst. & Meth. in Phys. Res. A **876** (2017) 145.

[79] J. L. Wiza, *Microchannel plate detectors*, Nucl. Inst. & Meth. **162** (1979) 587.

[80] G. W. Fraser *et al.*, *Enhanced soft X-ray detection efficiencies for imaging microchannel plate detectors*, Nature **300** (1982) 509.

[81] T. Conneely *et al.*, *The TORCH PMT: a close packing, multi-anode, long life MCP-PMT for Cherenkov applications*, JINST **10** (2015) C05003.

[82] J. S. Milnes *et al.*, *The TORCH PMT, a close packing, multi-anode, long life MCP–PMT for Cherenkov applications*, Nucl. Inst. & Meth. in Phys. Res. A **766** (2014) 183.

[83] T. Gys *et al.*, *Performance and lifetime of micro-channel plate tubes for the TORCH detector*, Nucl. Inst. & Meth. in Phys. Res. A **766** (2014) 171.

[84] N. Harnew *et al.*, *Status of the TORCH time-of-flight project*, Nucl. Inst. & Meth. in Phys. Res. A **952** (2018) 161692, arXiv:1812.09773.

[85] T. M. Conneely *et al.*, *Extended lifetime MCP-PMTs: Characterisation and lifetime measurements of ALD coated microchannel plates, in a sealed photomultiplier tube*, Nucl. Inst. & Meth. in Phys. Res. A **732** (2013) 388.

[86] S. M. George, *Atomic Layer Deposition: An Overview*, Chem. Rev. **110** (2010) 111.

[87] R. Gao *et al.*, *Development of precision Time-Of-Flight electronics for LHCb TORCH*, JINST **9** (2014) C02025.

[88] R. Gao *et al.*, *Development of scalable electronics for the TORCH time-of-flight detector*, JINST **10** (2015) C02028.

[89] R. Gao *et al.*, *Development of TORCH readout electronics for customised MCPs*, JINST **11** (2016) C04012.

[90] F. Anghinolfi *et al.*, *NINO: an ultra-fast and low-power front-end amplifier/discriminator ASIC designed for the multigap resistive plate chamber*, Nucl. Inst. & Meth. in Phys. Res. A **533** (2004) 183.

[91] F. Anghinolfi *et al.*, *NINO, an ultra-fast, low-power, front-end amplifier discriminator for the Time-Of-Flight detector in ALICE experiment*, in *Nuclear Science Symposium Conference Record, 2003 IEEE*, vol. 1, p. 375, 2003.

[92] M. Mota *et al.*, *A flexible multi-channel high-resolution time-to-digital converter ASIC*, in *Nuclear Science Symposium Conference Record, 2000 IEEE*, vol. 2, p. 9/155, 2000.

[93] A. V. Akindinov *et al.*, *Design aspects and prototype test of a very precise TDC system implemented for the Multigap RPC of the ALICE-TOF*, Nucl. Inst. & Meth. in Phys. Res. A **533** (2004) 178.

[94] The LHCb Collaboration, LHCb Experiment, *Physics case for an LHCb Upgrade II*, Tech. Rep. LHCb-PUB-2018-009, CERN, Geneva, 2018.

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

[96] R. Forty, *RICH pattern recognition for LHCb*, Nucl. Inst. & Meth. in Phys. Res. A **433** (1999) 257.

[97] N. Harnew *et al.*, *Status of the TORCH time-of-flight project*, Nucl. Inst. & Meth. in Phys. Res. A **952** (2020) 161692, `arXiv:1812.09773`.

[98] S. Bhasin *et al.*, *Test-beam studies of a small-scale TORCH time-of-flight demonstrator*, Nucl. Inst. & Meth. in Phys. Res. A **961** (2020) 163671, `arXiv:2002.07632`.

[99] T. H. Hancock *et al.*, *Beam tests of a large-scale TORCH time-of-flight demonstrator*, Nucl. Inst. & Meth. in Phys. Res. A **958** (2020) 162060, `arXiv:1904.11235`.

[100] M. van Dijk, *Design of the TORCH detector: A Cherenkov based Time-of-Flight system for particle identification*, PhD thesis, University of Bristol, 2016.

[101] L. Durieu *et al.*, *The CERN PS east area in the LHC era*, IEEE Transactions on Nuclear Science **1** (1998) 228.

[102] L. Durieu and O. Ferrando, *Design of T9 (ATLAS/CMS) for EHNL*, Tech. Rep. PS/PA/Note 96-39, CERN, Geneva, 1996.

[103] I. Rubinskiy, *An EUDET/AIDA Pixel Beam Telescope for Detector Development*, Physics Procedia **37** (2012) 923.

[104] C. Hu-Guo *et al.*, *First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope*, Nucl. Inst. & Meth. in Phys. Res. A **623** (2010) 480.

[105] P. Baesso *et al.*, *The AIDA-2020 TLU: a flexible trigger logic unit for test beam facilities*, JINST **14** (2019) P09019, `arXiv:2005.00310`.

[106] Y. Liu *et al.*, *EUDAQ2 – A Flexible Data Acquisition Software Framework for Common Test Beams*, JINST **14** (2019) P10033, `arXiv:1907.10600`.

[107] J. Allison *et al.*, *Recent developments in Geant4*, Nucl. Inst. & Meth. in Phys. Res. A **835** (2016) 186.

[108] S. Liu *et al.*, *LUT-based non-linearity compensation for BES III TOF's time measurement*, Nuclear Science and Techniques **21** (2010) 49.

[109] L. Castillo García, *Study of a prototype module of a precision time-of-flight detector for particle identification at low momentum*, PhD thesis, École polytechnique fédérale de Lausanne (EPFL), 2016.

[110] E. D. Palik, *Handbook of Optical Constants of Solids*, Academic Press, 1998.

[111] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986.

[112] E. P. M. Gabriel, *Observation of new $\Xi_c^0$ states decaying to $\Lambda_c^+ K^-$, and development of the TORCH time-of-flight detector*, PhD thesis, University of Edinburgh, 2020.

[113] P. L. Chebyshev, *Théorie des mécanismes connus sous le nom de parallélogrammes*, Mémoires des Savants étrangers présentés à l'Académie de Saint-Pétersbourg **7** (1854) 539, (in French).

[114] F. Cheung, *Measurements of the CKM angle $\gamma$ at the LHCb experiment*, PhD thesis, University of Oxford, UK, 2016.