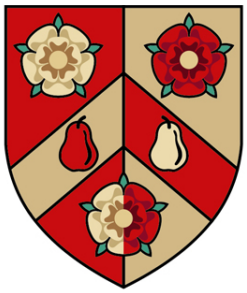
The first full angular analysis of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays





Rizwaan Adeeb Mohammed

Wolfson College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2024



Acknowledgements

I would like to express my gratitude to my supervisor, Malcolm John, for his support and guidance throughout my PhD. His unwavering ambition and belief in my capabilities, even at times when I doubted myself, gave me immense confidence and was instrumental in pushing the analysis forward.

I am also deeply indebted to Donal Hill, without whom this work would not have been envisaged. His mentorship has enabled me to develop in many ways as a physicist. I simply can't thank him enough for his continuous encouragement and advice during my first three years. In addition, I am grateful for his friendship and thought-provoking discussions, which have enriched my PhD journey.

During my time in Switzerland, I had the opportunity to work closely with members of the LHCb VELO team. They welcomed me warmly and were always eager to give me responsibility and help me develop. I am grateful to all of them, particularly Paula Collins, Victor Coco, and Nathan Jurik. I have many fond memories of the countless hours we spent together in the control room, and I'll miss the regular sushi lunches and ice cream breaks.

The LHCb Oxford group have been a source of camaraderie and support for the last four years. I thank all of them for helping to interrupt the constant work with punting, tennis or even just a chat. I feel especially lucky to have shared an office with Jenny and Luke, who have been incredibly entertaining and supportive colleagues.

I must express my appreciation to Asad, Hashi, and Will for their friendship, even when we were spread across three countries. The frankly ridiculous sports debates that we were inescapably drawn into were very enjoyable. I'll remember the adventures on our epic road trip for a long time.

Life in Oxford was greatly enhanced by friendships at Wolfson College. I'm grateful to Joost, Dimitri, Piotr, Christine, and Elizabeth for organising fun pandemic-compliant activities and some unforgettable trips. The lockdown period would have

been a much lonelier time without them.

I could not have reached this point without the continuous support of my family. I am thankful to Irfaan for our regular visits and phone calls, and for being the only person I can talk to about our favourite sports team. I am thankful to my parents for providing me with every opportunity I could ever have asked for. My father's relentless pursuit of self improvement and example of overcoming significant obstacles always inspires me to set myself ambitious goals. I am most grateful to my mum, not just for being my first maths teacher, but for her huge sacrifices that have allowed me to get where I am today.

Abstract

The kinematics of semileptonic $B \to D^* l\nu$ decays are fully described by a twelve-term rate equation parameterised by three decay angles and the hadronic recoil. The twelve terms are each multiplied by an angular coefficient to give the total rate. These angular coefficients are dependent on the fundamental couplings and their measurement may reveal interference between the Standard Model, charged weak interaction and hypothetical New Physics currents. Measurement of the angular distributions of semileptonic decays is compromised by acceptance and reconstruction effects, most notably the lost neutrino. A binned template histogram method is demonstrated, which is model independent and can account for reconstruction biases. This is used to measure the 12 angular coefficients in the 2016-2018 LHCb dataset. The measured values have excellent agreement with the Standard Model predictions. This is the first measurement of these coefficients in the $B \to D^* \mu\nu$ decay mode, as well as the first semileptonic angular analysis performed at LHCb.

Contents

Preface **xiii**

1 Introduction **1**

2 The LHCb experiment **5**
2.1 The Large Hadron Collider 5
2.1.1 LHC data taking periods 7
2.2 The LHCb experiment 7
2.2.1 The Vertex Locator 9
2.2.2 The Dipole magnet 12
2.2.3 Tracking stations 12
2.2.4 Ring-Imaging Cherenkov detectors 14
2.2.5 Calorimeters 15
2.2.6 Muon system 17
2.2.7 Trigger system 19
2.2.8 Track reconstruction 22
2.2.9 Particle identification 23

3 Semileptonic B meson decays **25**
3.1 The Standard Model 25
3.2 Electroweak interaction 26
3.3 Lepton Flavour Universality 27
3.4 $B \to D^* l\nu$ decays 30
3.4.1 $B \to D^* l\nu$ form factors 31
3.4.2 CLN parameterisation 34
3.4.3 BGL parameterisation 34
3.4.4 Summary 35
3.5 Derivation of the angular decay rate 35
3.5.1 Phase-space factor 36
3.5.2 Matrix element 37
3.5.3 Decay rate 40
3.5.4 I_X theoretical predictions 41
3.5.5 Other observables 41

4 Angular analysis methodology 43
4.1 Decay angle definitions . . . 43
4.2 Calculation of the angles . . . 45
4.2.1 B^0 rest frame approximation . . . 45
4.2.2 Decay angles . . . 50
4.3 Angular fit method . . . 53
4.3.1 Extending to four dimensions . . . 59
4.3.2 Adding backgrounds . . . 60
4.3.3 Fit binning and parameters . . . 62
4.3.4 Polynomial corrections . . . 63

5 Angular analysis of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays 65
5.1 Analysis strategy . . . 65
5.2 Datasets and selection . . . 66
5.2.1 Simulation . . . 68
5.3 Background description and reduction . . . 70
5.3.1 Fake D^0 and fake D^* background . . . 71
5.3.2 Combinatorial background . . . 75
5.3.3 Multivariate classifiers . . . 77
5.4 Corrections to simulation . . . 85
5.4.1 Aligning fast simulation with full LHCb simulation . . . 85
5.4.2 Aligning simulation with data . . . 87
5.5 Extracting the generator-level I_X from simulation . . . 94
5.5.1 Introduction . . . 94
5.5.2 Parametric fit . . . 94
5.5.3 Fit validation with toys . . . 97
5.5.4 Method of Moments . . . 97
5.6 Model independence of templates . . . 102
5.6.1 Template shapes under model variation . . . 102
5.6.2 Potential biases from q^2 acceptance effects . . . 105
5.7 Template fit results and validation . . . 112
5.7.1 Validation with simulated data . . . 113
5.7.2 Validation of polynomial corrections . . . 114
5.7.3 Results from template fit to real data . . . 116
5.7.4 Cross-checks . . . 117
5.8 Systematic uncertainties . . . 122
5.8.1 Template statistics . . . 122
5.8.2 Fixed components in angular fit . . . 122
5.8.3 Tracker-only MC corrections . . . 123

5.8.4 Model independence of templates 124
5.8.5 Fake D^0, fake D^* fit 126
5.8.6 D^{**} model uncertainty 127
5.8.7 $B^0 \to D^{*-}\pi^+$ corrections 127
5.8.8 Binning scheme in template fit 128
5.8.9 Residual muon misID background 129
5.8.10 L0 trigger variation 130
5.8.11 Total systematic uncertainties 131

6 VELO online monitoring and tomography 133
6.1 VELO upgrade 133
6.2 VELO closing procedure 136
6.3 Vertex monitoring for VELO closing 137
6.4 RF foil metrology with vertex reconstruction 140
6.4.1 RF foil metrology from first VELO closing 140
6.4.2 VELO aperture measurement 143
6.4.3 RF foil metrology after vacuum incident 145

7 Conclusions 149
7.1 Summary 149
7.2 Looking towards the future 150

Appendices

A Definition of functions used in mass fits 155

B Helicity amplitudes 157
B.1 Leptonic amplitudes 157
B.2 Hadronic amplitudes 159

C Definition of I_X coefficients in terms of helicity amplitudes 161

Bibliography 163

Preface

Any particle physics endevour relies on cooperation and the generosity of colleagues. For the last four years I have worked within the LHCb collaboration and its semileptonic working group. More directly, I have benefited from ideas and code from Donal Hill when starting the analysis that dominates this thesis. All of the work presented in this thesis is my own, except where clearly referenced to others.

In addition to the work shown in this thesis, I undertook several other roles and projects during my PhD:

- **Search for $B_c \to D^{(*)}\mu\nu$ decays (2020-2021)**
  I designed a project that aimed to search for $B_c \to D^0\mu\nu$ and $B_c \to D^{*0}\mu\nu$ decays, and supervised a student who worked on this as part of their MPhys degree. These processes had never been previously measured. The method involved developing a classifier to efficiently enhance the rare signal decay. Then a fit was performed to the B meson corrected mass distribution to identify the $B_c^+ \to D^0\mu\nu$ and $B_c^+ \to D^{*0}\mu\nu$ components.

- **Simulation Liaison (2022-2023)**
  I was responsible for coordinating the simulation sample requests from the LHCb semileptonic working group

- **LHCb UK meetings (2022)**
  I volunteered as an organiser for the LHCb UK student meetings, which are a series of monthly seminars aimed at current PhD students.

- **Vertex Locator (VELO) commissioning (2022)**
  During my time at CERN, I contributed to the commissioning of the new Vertex Locator which was installed in 2022. The performance of the VELO closing monitoring algorithm that I developed is shown in Chapter 6. In addition to this, I undertook several on-call shifts for the VELO. Furthermore, I was involved in reconstructing the first vertices from the LHCb Run 3 data, and assisting with closing the VELO for the first time. Also, I travelled to CERN in April 2023 to assist with the data taking and tomography studies.

1 Introduction

Semileptonic decays, processes which produce a charged lepton and a corresponding neutrino, form significant strands in the rich tapestry of particle physics. Their legacy dates back to 1899 when Ernest Rutherford studied the newly-discovered phenomenon of radioactivity [1]. He identified two distinct components, *alpha* decay which could be absorbed by a thin sheet of aluminium, and *beta* decay which was far more penetrating. The properties of beta decay were further studied in 1900 by Becquerel. By investigating the deflection of beta particles in electric and magnetic fields, he measured their charge to mass ratio [2], demonstrating that they were identical to electrons, which had recently been discovered by JJ Thompson.

In 1911, Meitner and Hahn found that emitted electrons didn't all have the same energy, rather their energies formed a continuous spectrum [3]. This apparently violated the law of conservation of energy, but the problem was resolved in 1930 by Pauli, who postulated that an extremely light neutral particle was also emitted and remained undetected. Then, in 1933, Fermi outlined his famous model for beta decay [4], describing how these *neutrinos* are produced along with electrons in a four-point interaction. Thus, beta decay was the first experimentally observed example of a semileptonic process.

Later, beta decays revealed that weak interactions are not mirror-symmetric, they violate *parity* symmetry. This was first observed in Cobalt-60 decays in

an experiment carried out by Chien-Shiung Wu [5]. This remarkable finding demonstrated that only left-handed quarks and leptons are sensitive to the weak interaction, and that neutrinos have only a left-handed component.

This also laid the foundation for the discovery of *charge-parity* (CP) violation. CP symmetry is the combination of swapping a particle for its corresponding antiparticle (charge conjugation, C) and a spatial inversion (parity, P). It is a vital condition to explain the observed matter-antimatter asymmetry of the universe [6]. Furthermore, semileptonic kaon decays provided the first direct evidence of a process that violates time-reversal symmetry (T) [7]. This result, combined with previous measurements of CP violation in kaons [8], was the first test of the fundamental CPT theorem, which predicts that any physical system should be invariant under the combination of the three symmetries.

Fermi's theory of beta decay was strongly supported by experimental data, however it is invalid at high energy since it predicts that the total cross section is proportional to the square of the energy. It was subsequently understood that Fermi's theory was incomplete, it functioned as an effective theory at low energies, but neglected additional complexity that was measurable at energies greater than approximately 100 GeV. Specifically, Fermi's theory doesn't account for the virtual W boson that is exchanged during the process (see Fig. 1.1).

Moreover, the work of Glashow, Weinberg, and Salam established the unification of the weak and electromagnetic forces [9–11]. The theory was confirmed by the discovery of weak neutral currents in 1973 [12], and then the discovery of the W and Z bosons at their predicted masses in 1983 [13, 14].

The Standard Model (SM), described fully in chapter 3, is our current working theory of fundamental particles and their interactions. Weakly decaying particles are excellent tools for probing the SM as they have long lifetimes, making them sensitive to rare unknown effects. In this regard, semileptonic B meson decays remain an area of active interest at the LHCb experiment. These processes have the highest branching fractions, and have the lowest theoretical uncertainties apart from fully leptonic decays. In addition to searching directly for new phenomena

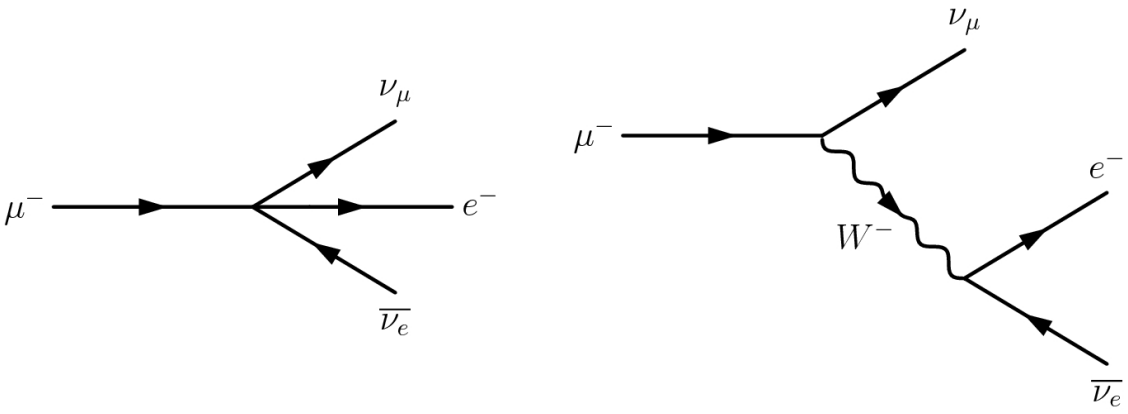


Figure 1.1: Feynman diagrams showing muon decay in Fermi's effective theory (left), and the full electroweak theory (right)

(NP), it is important to make measurements that improve our understanding of SM processes. $B^0 \to D^{*-}\mu^+\nu_\mu$ decays are potentially interesting due to their large event yields made possible by efficient muon detection, and the lesser likelihood of NP affecting the second lepton generation. To access detailed information about the underlying process, the spin-structure of the decay can be exploited by examining the spatial distribution of the final state particles, a technique known as *angular analysis*. This thesis presents the first full angular analysis of the $B^0 \to D^{*-}\mu^+\nu_\mu$ decay[1] using data from the LHCb experiment.

The remainder of the thesis is structured as follows:

- **Chapter 2 – The LHCb experiment**: The Large Hadron Collider is briefly described then each component of the LHCb experiment is outlined in detail.
- **Chapter 3 – Theoretical background**: The Standard Model description of semileptonic decays is presented and the derivation of the angular decay rate for the signal decay is shown.
- **Chapter 4 – Methods**: The calculation of the decay angles is outlined followed by a description of the novel template fit method used to measure the angular coefficients.

[1] charge conjugation is implied throughout except where explicitly stated otherwise.

- **Chapter 5 – Angular analysis**: The full details of the angular coefficient measurement using LHCb data, along with their systematic uncertainties are presented.
- **Chapter 6 – VELO online monitoring**: A brief description of the new VELO and the online monitoring for its closing procedure is given. Also, the results of the VELO tomography are shown.
- **Chapter 7 – Conclusions**: The results and their interpretation are summarised, then possible extensions of this method for future measurements are outlined.

2 The LHCb experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest and most powerful particle accelerator ever built. It is located at CERN, near Geneva, and lies on the border between France and Switzerland. It has a circumference of 26.7 km, and is designed to collide protons or ions accelerated to energies of $6 - 7\,\text{TeV}$ per nucleon.

Protons are fed into the LHC by a system of smaller accelerators, which is illustrated in Fig. 2.1. First, an electric field is used to strip electrons from hydrogen atoms, with the resulting protons fed into Linac2 and accelerated to an energy of 50 MeV. The protons are then injected into the Proton Synchrotron Booster and they are accelerated to an energy of 2 GeV. Next, they are accelerated by the Proton Synchrotron, then by the Super Proton Synchrotron, before finally being injected into the LHC at 450 GeV. The LHC accelerates the protons to their maximum energy (6.5 TeV until 2018, and 6.8 TeV since 2021).

The LHC is made of eight arcs connected by eight straight sections. The protons travel in two beams, contained in two separate beam pipes, in opposite directions and they collide at four interaction points around the ring. The LHC beams are divided into bunches containing around 1.2×10^{11} protons, with a maximum of 2808 bunches per beam. 1232 dipole magnets with a field of 8.3 T are used to

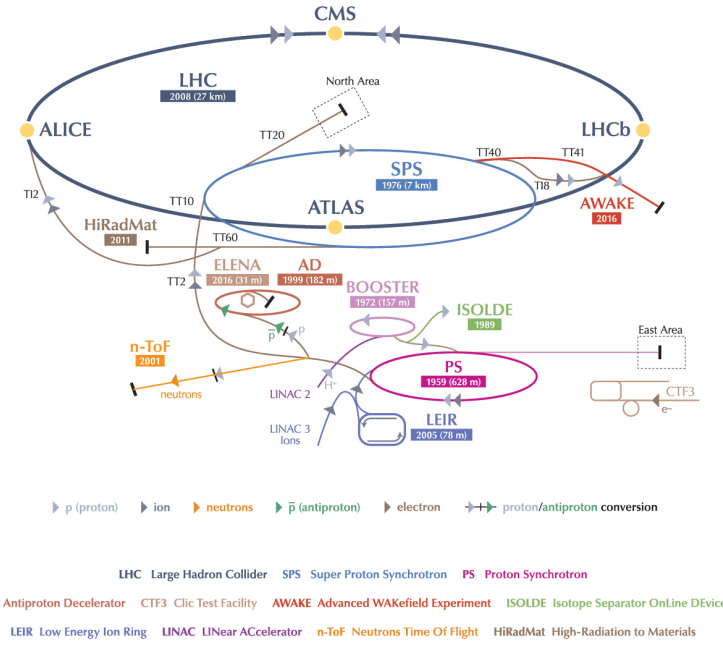


Figure 2.1: The accelerator complex at CERN, along with some of the experiments linked to the accelerators [15].

bend the beams, which are focused by 392 quadrupole magnets. The magnets are superconducting, so they have a cooling system to keep them at their operating temperature of 1.9 K. The protons are accelerated using radiofrequency (RF) cavities. The RF cavities have a voltage of 16 MV, and a frequency of 400 MHz.

Each of the four interaction points on the LHC has a major experiment situated around it. These are ATLAS, CMS, ALICE, and LHCb. ATLAS and CMS are "general-purpose" detectors, predominantly aiming to study high-mass objects like the Higgs boson, the weak bosons, and the top quark, and to perform searches for NP signatures at the TeV scale. ALICE was designed for heavy-ion physics, and studies the physics of the quark-gluon plasma. LHCb is a dedicated flavour physics experiment and will be described in detail in Section 2.2.

2.1.1 LHC data taking periods

The LHC first recorded data from proton-proton collisions in 2009. It entered its first period of data taking for physics in 2010, known as Run 1, which lasted until 2012. The centre of mass energy for the collisions during Run 1 was $\sqrt{s} = 7\,\text{TeV}$ in 2010 and 2011, and it was increased to $\sqrt{s} = 8\,\text{TeV}$ in 2012. The data collected during Run 1 corresponds to an integrated luminosity of approximately $3\,\text{fb}^{-1}$.

Between 2012 and 2015, the LHC was shut down for upgrades. Then, the next period of data taking, Run 2, started in 2015, lasting until 2018. Collisions during Run 2 had a centre of mass energy of $\sqrt{s} = 13\,\text{TeV}$, and a total integrated luminosity of around $6\,\text{fb}^{-1}$. After Run 2, the LHC and the experiments underwent another major upgrade in preparation for Run 3, which started in 2022.

The analysis work presented in this thesis is based on data taken from 2016 to 2018. Accordingly, this chapter describes the LHC and LHCb conditions as they were during Run 2. In addition, Chapter 6 describes work carried out for the new LHCb Vertex Locator for Run 3, which was installed and commissioned in 2022. A description of the Vertex Locator in Run 3 is given in Chapter 6, but the full LHC and LHCb upgrades are otherwise not discussed.

2.2 The LHCb experiment

The LHCb detector is a single-arm spectrometer in the forward region. It has an angular coverage of approximately 10 to 300 mrad in the horizontal (bending) plane, and 250 mrad in the vertical (non-bending) plane[1]. It has been designed to study the behaviour of particles containing b and c quarks produced from the proton-proton collisions at the LHC. As shown in Fig. 2.2, at high energies the production of $b\bar{b}$ pairs is predominantly in the "forward region". This motivates the choice of detector geometry. With this geometry, 24% of $b\bar{b}$ quark pairs are produced inside the LHCb acceptance [16], despite LHCb covering a much smaller fraction of the solid angle.

[1]The coordinate system is shown in Fig. 2.3 and defined at the end of this section.

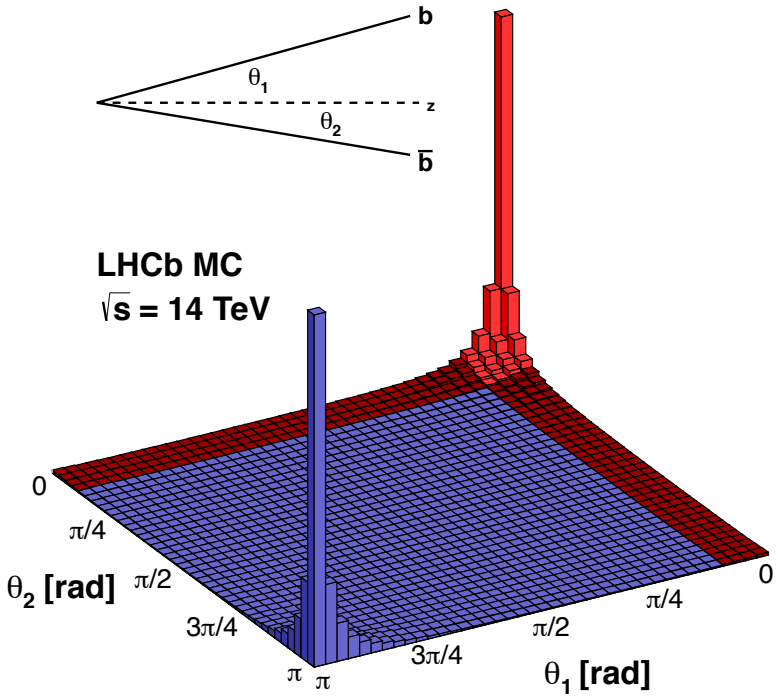


Figure 2.2: Spatial distribution of $b\bar{b}$ pairs produced from pp collisions. The angles θ_1 and θ_2 are defined relative to the z axis, which lies along the beamline. The red region shows the acceptance of the LHCb detector. [16]

The initial physics goals of LHCb were related to precisely probing the heavy flavour sector of the SM. In particular, measurements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements were conceived to provide a rigorous test of the SM by testing the self-consistency of the CKM description of quark interactions with the W boson through measurements of various heavy quark transitions. Furthermore, measurements of processes that are rare in the SM like the $B_s^0 \to \mu\mu$ branching fraction and an angular analysis of $B^0 \to K^{*0}\mu\mu$ decays were envisaged, as these measurements have high sensitivity to NP effects [17].

In addition to its original physics goals, LHCb has proven to be incredibly versatile by making a wide range of further measurements related to beauty and charm hadron decays. These include spectroscopy of exotic hadrons, most notably the discovery of the pentaquark in 2015 [18], many studies in the charm sector, and

measurements from fixed target collisions through the SMOG[2] system. The SMOG system injects a small amount of noble gas into the LHC beam pipe, allowing proton-nucleon collisions to be studied. Most relevant for this thesis however, is the LHCb semileptonic program. Although these decays were well motivated, being theoretically clean and having relatively high branching fractions, it was thought that the presence of unreconstructed neutrinos[3] would severely compromise the measurements, due to the resulting absence of narrow mass peaks.

Despite these challenges, LHCb has made notable contributions by studying semileptonic B decays. For example, measurements of V_{ub} and V_{cb} have been performed [19, 20], and currently have tensions between inclusive and exclusive measurements. Also, there have been several searches for LFV through measurements of $R(D)$ and $R(D^*)$ [21, 22], along with CP asymmetry and production cross-section measurements [23, 24].

The main components of the Run 2 LHCb detector are shown in Fig. 2.3, and will be described in the following sections. The LHCb coordinate system is defined with the z axis along the beamline, the x axis in the horizontal direction pointing away from the centre of the LHC ring, and the y axis in the vertical direction. The origin is defined as the nominal pp collision point.

2.2.1 The Vertex Locator

The Vertex Locator (VELO) is the innermost part of the LHCb detector, placed around the pp collision region and 5.1 mm from the beamline. Its role is to precisely measure the positions of the pp collision points, known as primary vertices (PV), and the positions of the decays of long-lived particles (secondary vertices).

The VELO is comprised of two halves placed either side of the beamline in the x direction. Each half is made of 21 modules, which lie collinear in z and perpendicular to the beamline, as shown in Fig. 2.4. The modules have a semicircular shape and are constructed from silicon strip detectors that provide independent measurements

[2]System for Measuring Overlap with Gas

[3]At LHCb, "semileptonic" is used to denote any decay involving a neutrino and this convention is adopted in this thesis. This is not necessarily the case outside LHCb, e.g. decays such as $B \to K^* \mu\mu$ are sometimes also referred to as "semileptonic".

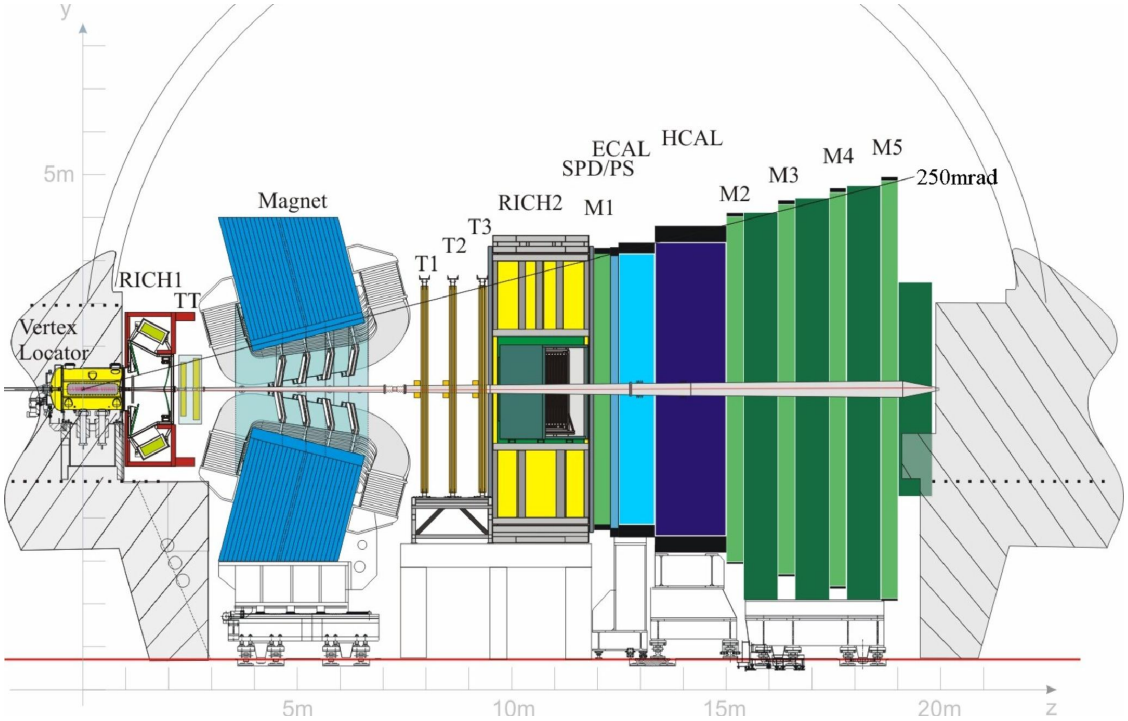


Figure 2.3: Diagram showing the layout of the LHCb detector with the individual subdetectors labelled [25]

of r, the radial distance from z, and ϕ, the azimuthal angle. Fig. 2.5 shows the configuration for a single VELO module.

The two VELO halves are mounted onto a movable frame. The purpose of this is to move each half to its nominal position for data taking at the start of each fill. Before the start of a fill, the halves are separated in the x direction, with each half held approximately 27 mm away from its final position. This configuration is referred to as the open position. After the injection and when the LHC declares "stable beams" (indicating that the beam positions will remain relatively unchanged), both halves move towards the beam so that the beam is centred inside the VELO, and the aperture formed by the two halves is 5.1 mm. The halves move independently in the x direction, but together in the y direction. Figure 2.6 shows the VELO in the fully closed position for stable beams, and in the fully open position during beam ramp-up.

The VELO is enclosed in a thin aluminium box, known as the RF foil, which separates the secondary VELO vacuum from the LHC primary vacuum and additionally protects the sensitive VELO electronics from beam-induced currents.

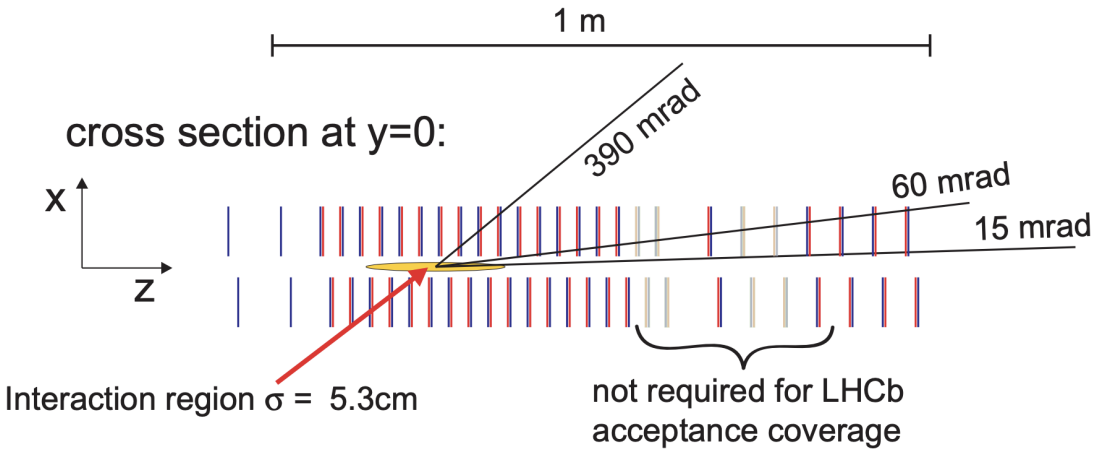


Figure 2.4: Schematic of the VELO in the x, z plane, with the red lines indicating the module positions [26]

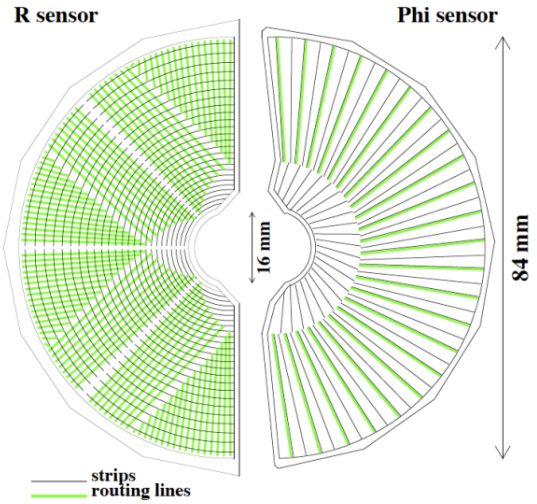


Figure 2.5: Layout of the silicon sensors in the VELO module, showing the separate configurations for the r measurement (left) and the ϕ measurement (right) [27]

Moreover, the silicon sensors are held at a temperature of $(-7 \pm 2)^{\circ}$C using a two-phase CO_2 cooling system. The CO_2 is delivered to each module through a set of pipes which are clamped to the base of a module.

The VELO has a measured tracking efficiency of over 98%. The impact parameter (minimum transverse distance of a track to a primary vertex (PV)), is measured with a resolution of $(15 + 29/p_T)\,\mu$m. The RMS decay time resolution is 50 fs which is crucial for many LHCb physics measurements.

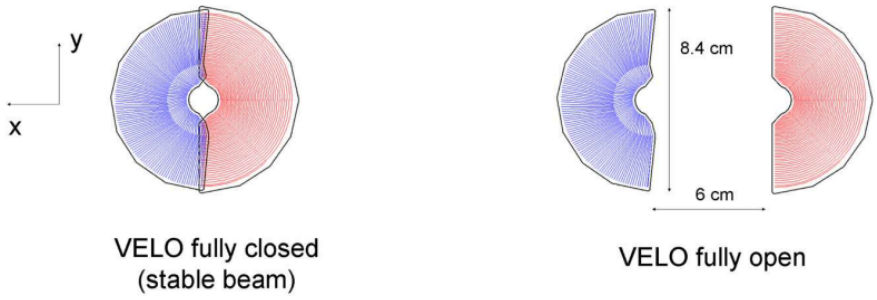


Figure 2.6: Configuration of a VELO module pair in the closed position (left) and open position (right) [28].

2.2.2 The Dipole magnet

A dipole magnet is used to bend the charged particles, enabling the measurement of their momentum. The magnet weighs around 1600 tonnes, is non-superconducting, and has an integrated bending power of around 4 Tm over its length. The polarity of the magnet can be reversed and this is done during a year of data-taking to collect separate data samples of approximately equal size with the "Magnet Up" and "Magnet Down" conditions. This is useful for analysis, since combining these two samples allows the cancellation of systematic uncertainties arising from detector asymmetry effects, which is important for measurements like *CP* violation. The variation of the magnet field strength along the beamline is shown in Fig. 2.7.

2.2.3 Tracking stations

In addition to the VELO, the LHCb detector is equipped with four tracking stations. The Tracker Turicensis [4] (TT) is situated before the magnet, and three tracking stations (T1-3) are situated after the magnet. The TT is made up of four layers. The middle two layers are tilted with respect to the outer layers by $\pm 5°$, which allows two-dimensional resolutions of the particle hits. The layers are built from silicon strip detectors, and each strip has a pitch of around 200 µm, leading to a spatial resolution of around 50 µm. A diagram of two TT layers can be seen in Fig. 2.8.

The T1-3 tracking stations are divided into the Inner Tracker (IT) and the Outer Tracker (OT). The IT has a similar structure to the TT, with four separate

[4]Turicensis refers to the Latin name for Zurich, the institution who built this sub-detector

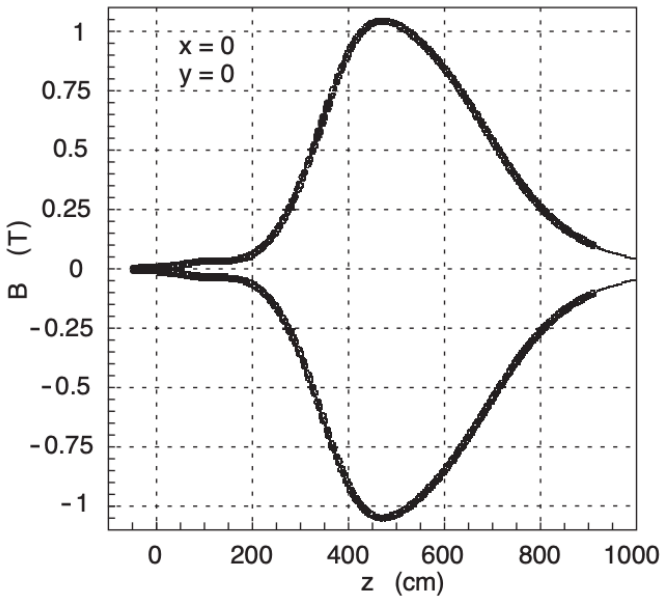


Figure 2.7: Magnetic field map along z axis for the Magnet Up and Magnet Down configurations [28]

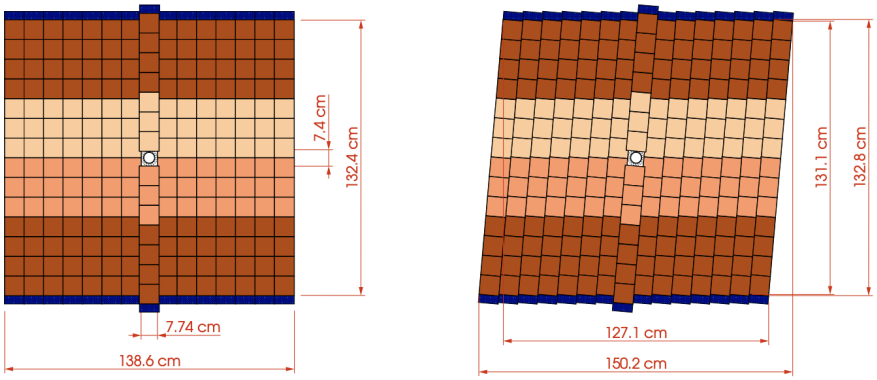


Figure 2.8: First and second layers of the TT. As mentioned in the text, the 2nd layer is rotated by ±5°. The beamline is shown in the centre of both layers [29]

planes that are built from silicon strip detectors. Each plane has seven detector modules. For each station, the IT is made up of four separate boxes, which are arranged around the beampipe. The arrangement is shown in Fig. 2.9. The inner tracker covers an area with a high density of particle hits, so silicon detectors were chosen to maximise the hit resolution. However, the region covered by the OT has a much lower hit density, so a cheaper alternative to the silicon strips is used. The OT uses an array of straw drift tubes, which have inner diameters of 4.9 mm, and are filled with a mixture of Argon and CO_2 gases. This combination of gases

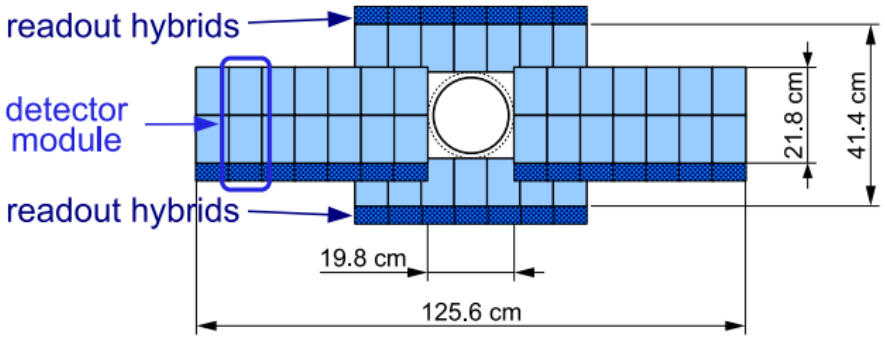


Figure 2.9: Diagram showing the layout of the four modules of the IT, with the beampipe shown in the centre [28]

ensures that the drift time is below 50 ns and the drift-coordinate resolution is below 200 µm. Together, the combination of the silicon and straw-tube tracker provide information on the curving tracks of ionising particles and thus their momentum. The measured momentum has a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV [30].

2.2.4 Ring-Imaging Cherenkov detectors

The purpose of the Ring-Imaging Cherenkov (RICH) detectors is to identify the charged hadrons: π, K, p. The detectors make use of Cherenkov radiation, a phenomenon where charged particles emit radiation when they travel faster than the speed of light in a particular medium. The photons are emitted in a cone around the line of the particle's trajectory, and the angle of the cone (θ_c) depends on the speed of the particle (v) via [31]:

$$\cos\theta_c = \frac{1}{n\beta} \,, \tag{2.1}$$

where β is defined as the ratio v/c, and n is the material's refractive index.

There are two separate RICH detectors. RICH1 is placed between the VELO and TT, whereas RICH2 is after T3. This configuration was chosen to allow excellent separation between particles across a large momentum range - RICH1 is more efficient for low momenta particles (2 - 40 GeV), whereas RICH2 is better for higher momenta (15 - 100 GeV) due to the different refractive indices of the gases used.

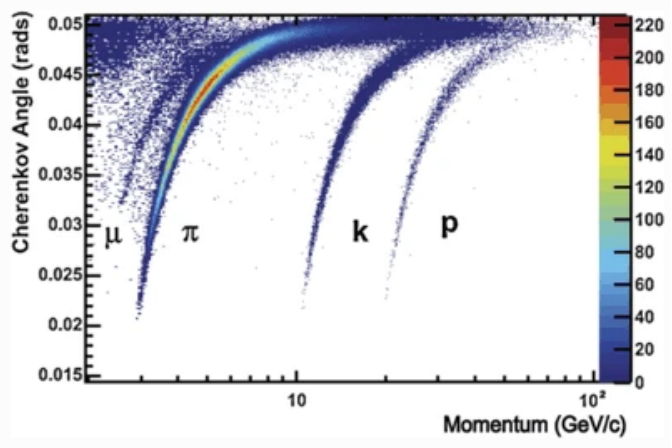


Figure 2.10: Dependence of the Cherenkov angle on particle momentum for protons, kaons, pions, and muons for the C_4F_{10} medium [32]

Both RICH detectors have a similar configuration, using a spherical mirror followed by a plane mirror to reflect and focus the Cherenkov photons into ring images on the photon detector planes, located outside of the LHCb acceptance. The hybrid photon detectors (HPDs) use vacuum tubes with 1024 silicon pixels per tube. The HPDs have focussing electrodes which concentrate the photocathode image onto a small silicon chip. The effective pixel area is $2.5 \times 2.5\,\mathrm{mm}^2$ at the photocathode. The medium chosen to fill the RICH detectors was fluorocarbon gases, specifically C_4F_{10} ($n = 1.0014$) in RICH1 and CF_4 ($n = 1.0005$) in RICH2 [32], because of their low dispersion effects. Fig. 2.10 shows the Cherenkov angle vs momenta for different charged particles. The layouts of RICH1 and RICH2 are shown in Fig. 2.11.

2.2.5 Calorimeters

LHCb uses a system of four calorimeters to measure the transverse energy of hadrons, electrons and photons. These are: a Scintillator Pad Detector (SPD), a PreShower detector (PS), an Electromagnetic CALorimeter (ECAL), and a Hadronic CALorimeter (HCAL). The calorimeters are placed in this order between the first and second muon stations.

The SPD and PS are constructed from scintillating pads. The atoms in the scintillating material are excited when charged particles pass through it, causing them to emit light. This light is then transmitted to a photomultiplier (PMT) via

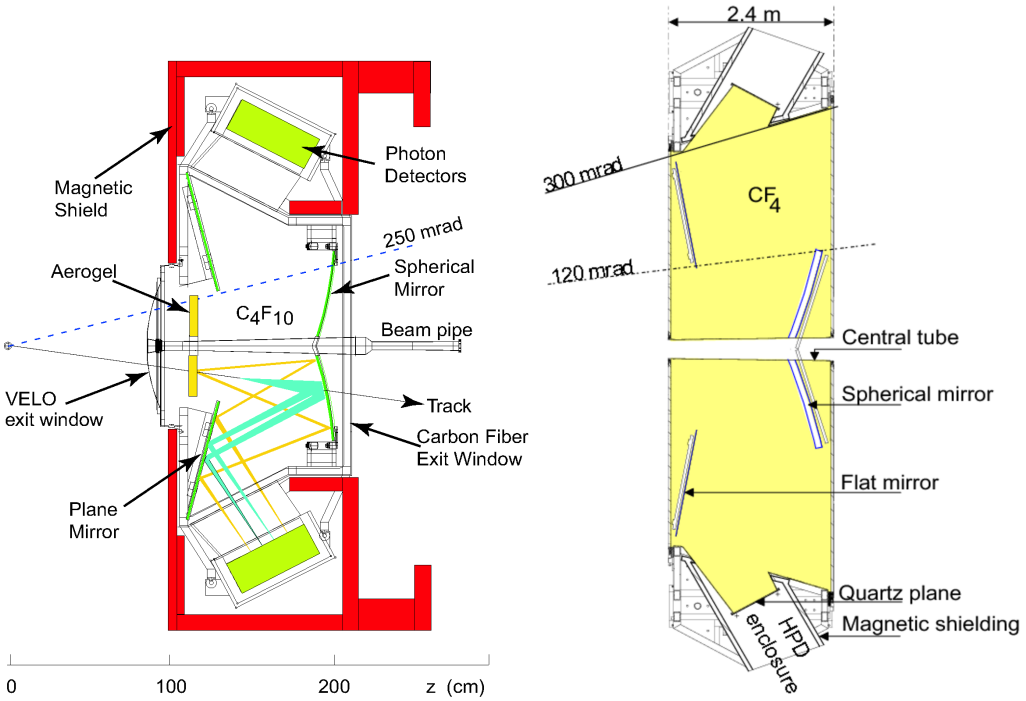


Figure 2.11: Schematics of RICH1 (left) and RICH2 (right) [28]. Note that the aerogel shown in RICH1 was used during Run 1 but removed for Run 2, as it did not give a sufficiently high Cherenkov photon yield.

wavelength-shifting fibres (to convert it to a more easily detectable frequency range). In total, the SPD and PS have 6016 scintillating pads. There is a 15 mm (around 2.5 radiation lengths) layer of lead placed in between the SPD and PS.

The detection of electrons and photons in the calorimeters is primarily performed through their "electromagnetic showers". For an electron entering a material, it loses energy through bremsstrahlung and emits a photon. If this photon has sufficiently high energy, it can undergo pair production, being converted into an e^+e^- pair. These electrons and positrons undergo further material interactions, producing more electrons and photons in a cascade. The energy decreases at each stage until the photons do not have enough energy to pair-produce, and the electrons do not have enough energy to emit bremsstrahlung radiation.

The role of the SPD is to distinguish electrons and photons, by detecting charged particles before they produce a shower. Then, the lead layer causes the electrons and photons to shower, but not charged hadrons or muons. In this way, the PS can distinguish between electrons and pions.

The ECAL modules are arranged in a "shashlik" style, which means alternating layers of lead and scintillator tiles, giving a high level of granularity and excellent energy resolution. In total, for a single ECAL module, there are 66 layers of lead and scintillator, forming a 42 cm stack (equivalent to 25 radiation lengths). The energy resolution of the ECAL is measured to be [33]:

$$\frac{\sigma(E)}{E} = \frac{(9 \pm 0.5)\%}{\sqrt{E}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.003}{E \sin\theta} \tag{2.2}$$

where E is the particle energy in GeV, θ is the angle between the beam axis and a line from the LHCb interaction point to the centre of the ECAL cell, and $\oplus$ signifies addition in quadrature. The first term is the stochastic term, the second is the constant term, and the third is the noise term.

Whereas the ECAL detects electrons and photons, the HCAL absorbs hadrons. Like the ECAL, it is constructed from alternating layers of scintillating tiles and absorbing material, but iron is used instead of lead as the absorber. The HCAL operates using a similar principle to the ECAL, however the hadronic showers are caused by strong interactions with the detector material instead of electromagnetic forces. The HCAL has a length equivalent to only 5.6 interaction lengths, and has an energy resolution of:

$$\frac{\sigma(E)}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\% \,, \tag{2.3}$$

as measured for the data-taking period corresponding to the work presented in this thesis. The two terms are the stochastic and constant terms. For both the ECAL and HCAL, the scintillators are divided into cells, so that the location of energy deposits can be identified. The dimensions of the cells vary across the entire structure, as is shown in Fig. 2.12. Fig. 2.13 shows a diagram of an ECAL and an HCAL cell.

2.2.6 Muon system

The muon system is the last component of the LHCb detector, located the furthest from the beampipe. This reflects the fact that muons are both highly penetrating,

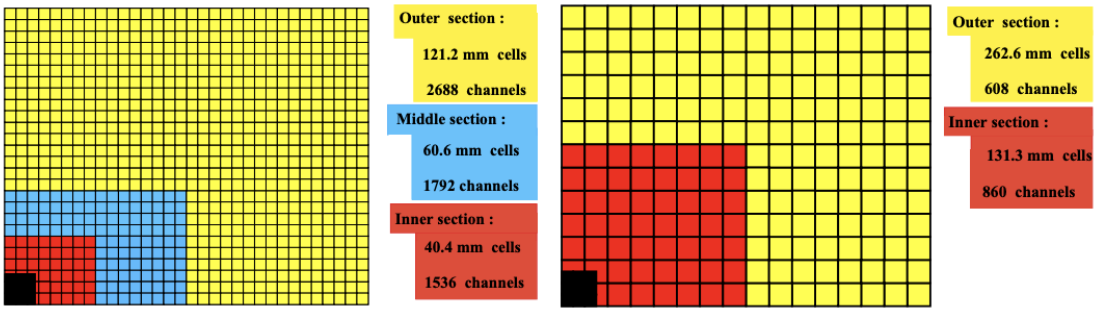


Figure 2.12: Variation in cell size with radial distance from the beampipe for the ECAL (left), and HCAL (right) [28]

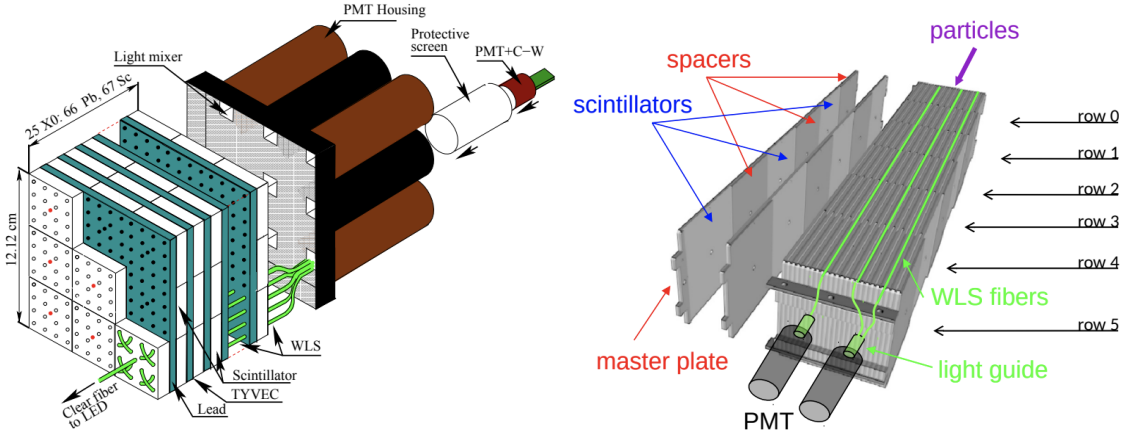


Figure 2.13: ECAL cell (left) and HCAL cell (right) [33]

so the majority of them can travel through the whole calorimeter system, and relatively stable, so only a small fraction decay before reaching the end of the detector. The muon system consists of five separate detectors, labelled M1-5. M1 is placed in front of the calorimeters, and M2-5 are located after them. Between the M2-5 stations are 80 cm thick iron blocks, to absorb any hadrons that have penetrated the calorimeters. Consequently, the higher a muon's momentum, the further through the muon system it reaches.

The inner part of M1 receives the highest level of radiation so uses triple gas-electron-multiplier detectors. The remaining part of M1 and M2-5 are made from multiwire proportional chambers (MWPC). In total, 1368 MWPCs are used, with a gas mixture of $Ar/CO_2/CF_4$ in the ratio (40:55:5). When charged particles pass through the MWPC, the gas molecules are ionised, producing electrons that drift towards the metal wires (which are held under high voltage). The current in

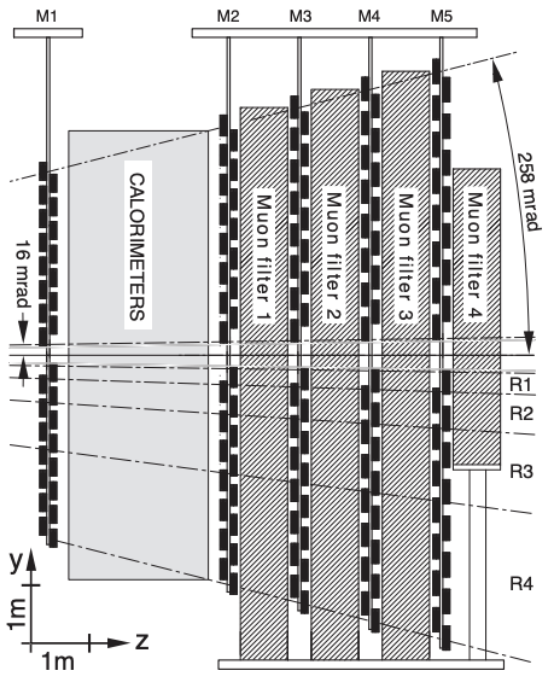


Figure 2.14: Location of the five muon stations along the beamline. The "muon filters" are the iron blocks mentioned in the text. The final muon filter protects M5 from machine-related backgrounds and backsplash from other LHC components.

the wires is detected, and the location of the hit can be identified by combining information from signals produced in a few different wires. A diagram of the muon system is shown in Fig. 2.14.

The efficiency of each muon station is $\geq 99\%$, which gives an overall efficiency of at least 95% for muons with momentum in the range $5 - 40\,\text{GeV}$. Fig. 2.15 shows the variation of efficiency with momentum. The work in this thesis is highly dependent on the efficient performance of the muon system, as this allows very large datasets of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays to be collected.

2.2.7 Trigger system

With the LHC bunch-crossing rate of 40 MHz, the rate of events stored by LHCb must be reduced. This is achieved with a multi-stage system, referred to as the trigger. First, a hardware trigger is used (known as L0). This selects events with high E_T clusters in the calorimeters, and high p_T muons in the muon chambers. Then these selected events are processed by the High Level Trigger (HLT), which is

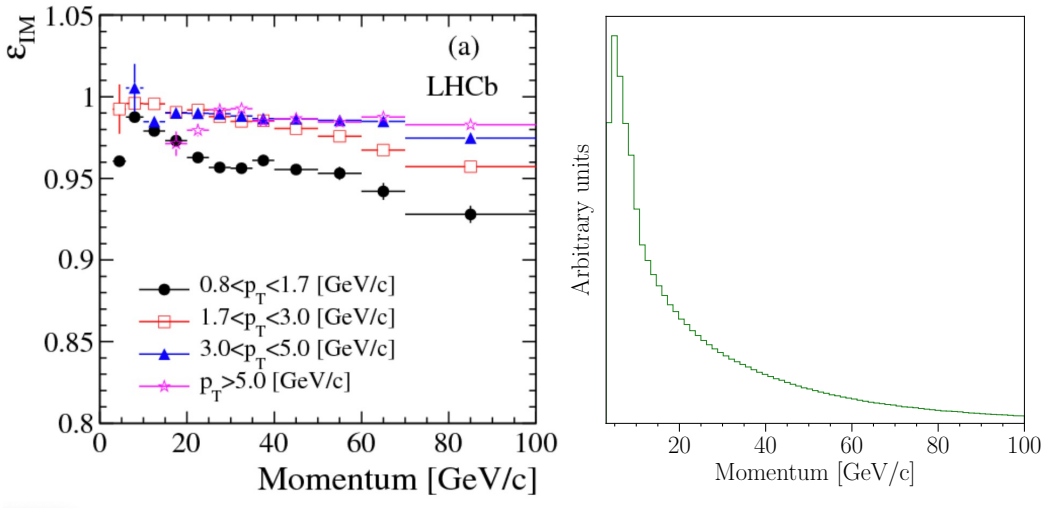


Figure 2.15: Left: muon efficiency as a function of momentum, in bins of p_T. Right: muon momentum spectrum in $B^0 \to D^{*-}\mu^+\nu_\mu$ data sample.

a software-based trigger that runs on a system of CPU cores. It is divided into two stages, HLT1 and HLT2. HLT1 performs a partial reconstruction of track segments, then HLT2 performs a full event reconstruction, before selecting events for storage. An outline of the LHCb trigger system is shown in Fig. 2.16.

L0 trigger

The L0 trigger reduces the event rate down to 1 MHz. It operates by independently selecting:

- The highest E_T hadron, electron and photon clusters in the ECAL and HCAL
- The two highest p_T muons in the muon system

Only these subdetectors are used as they are the only systems that can provide information quickly enough for the L0 trigger to operate at its desired rate. For the calorimeter reconstruction, a zone of 2×2 cells is defined. The total E_T for the zone is calculated, and compared to predefined thresholds, by:

$$E_T = \sum_{i=1}^{4} E_i \sin\theta_i \tag{2.4}$$

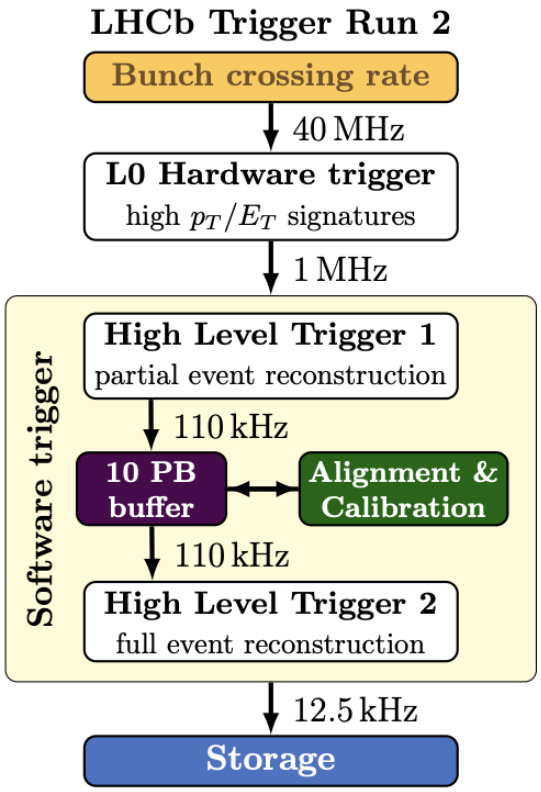


Figure 2.16: LHCb trigger system, showing the event rates at each stage [34]

where θ_i is the angle of a given cell from the nominal collision point. Information from the SPD and PS is used to classify the calorimeter clusters as electron, photon and hadron candidates. In addition, cuts are placed on the number of hits in the SPD cells, to reject events with very high detector occupancy, as these would take too long to process later in the HLT. For this analysis, an upper bound of 600 is placed on the number of hits.

HLT

HLT1 further reduces the event rate from 1 MHz to around 100 kHz for inclusive charm and beauty triggers, and to 40 kHz for inclusive muon triggers. HLT1 does not have sufficient computing power to perform a full event reconstruction, so instead performs a three step pattern recognition procedure:

- Reconstruct VELO track segments

- Extrapolate VELO tracks to TT stations forming *upstream* tracks, then to the T1-3 stations forming *long* tracks
- Fit long tracks using a Kalman filter [35]

Furthermore, the reconstructed VELO tracks are used to determine the primary vertex (PV) position, and muon identification is applied to events with a certain number of hits in the muon stations. These events are then passed to HLT2.

HLT2 has access to the full tracking information, so reconstructs more high quality tracks with better momentum estimates. Also, HLT2 has full particle identification (PID) information, using the calorimeters and the RICH detectors. Secondary vertices are also searched for. HLT2 aims to replicate the offline selection tools used at LHCb as closely as possible.

2.2.8 Track reconstruction

As outlined above, the LHCb detector is equipped with a variety of tracking subdetectors. Information from all of these is combined to determine the trajectory of different final state particles. Fitted tracks can be categorised according to the source of the hits used to reconstruct them, as follows:

- **VELO tracks** - these are tracks formed only from hits in the VELO modules, these candidates leave the LHCb acceptance before reaching the TT.
- **Upstream tracks** - these tracks have hits in the VELO and TT, but are deflected out of LHCb before T1-3
- **T tracks** - they have hits only in the T stations
- **Long tracks** - tracks that have hits in the VELO, TT, and tracking stations. They have the most precise momentum measurement
- **Downstream tracks** - tracks with hits in the TT and T stations, usually used for particles that decay after the VELO, such as the K_S^0

Fig. 2.17 is a schematic showing the different track types with the associated hits.

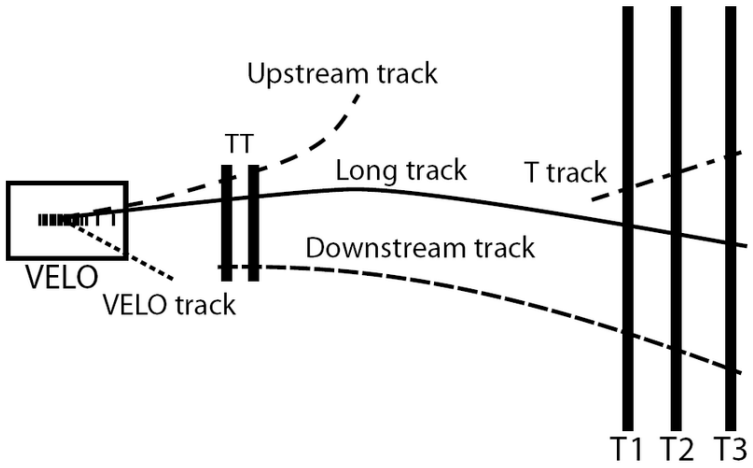


Figure 2.17: Diagram showing the definition of the track types used in LHCb [36]

2.2.9 Particle identification

The LHCb PID system consists of the RICH detectors, calorimeters, and the muon system. Information from each of these sources is combined to produce sets of variables, which are then used by analysts to perform custom selections.

Particularly relevant for the work in this thesis is the muon identification method. A variable `isMuon` is primarily used for this purpose. This places requirements on hits of the muon candidate in the muon stations, which are different depending on the candidate's momentum. These requirements are listed in Tab. 2.1.

p [GeV/c]	Required stations
$p < 3$	*Always false*
$3 < p < 6$	M2 & M3
$6 < p < 10$	M2 & M3 & (M4 || M5)
$p > 10$	M2 & M3 & M4 & M5

Table 2.1: Requirements for a candidate to be flagged `isMuon==True` [37]

For hadrons, the PID information is taken from the RICH detectors. As mentioned earlier, the different masses of the hadrons lead to different Cherenkov cone angles reconstructed in the RICH. To distinguish the particles, a log-likelihood procedure is used. For each track, a hypothesis of the particle being a pion is used, and the other particle hypotheses are evaluated relative to this original one. Then,

the likelihood is maximised, varying the hypotheses. The ΔLL between any particle combination is then calculated using the individual ΔLL between the particles and a pion. For example, for a kaon or proton hypothesis:

$$\Delta LL_{K/p} = \Delta LL_{K/\pi} - \Delta LL_{p/\pi} = \ln \mathcal{L}(K) - \ln \mathcal{L}(p) \ . \tag{2.5}$$

The LHCb PID system provides excellent separation of particle types [38]. Muons can be identified with $\sim 97\%$ efficiency, where the probability of misidentifying a pion is $1 - 3\%$. Also, kaons can be identified with $\sim 95\%$ efficiency, with a $\sim 5\%$ probability of misidentifying a pion.

Identification of electrons and neutrals is not discussed here, as it is not relevant for the work presented in this thesis.

3

Semileptonic B meson decays

3.1 The Standard Model

The Standard Model (SM) of particle physics is a theory describing the fundamental particles that make up our universe and their interactions. The SM is arguably the most important and successful scientific theory to date. Its predictions have been extensively tested by numerous particle physics experiments. It successfully describes three of the four fundamental forces: the weak interaction, the strong interaction, and electromagnetism. The fundamental SM particles are shown in Fig. 3.1, and are divided into fermions and bosons.

Fermions, which have half-integer spin, are matter particles and are divided into quarks and leptons. Quarks can couple to both the strong and electroweak interaction, whereas leptons only interact electoweakly. Both the quarks and leptons are split into three generations. The generations differ by their masses but their electric charge is the same, as is their coupling to the strong interaction. Within the groups of quarks and leptons, the particles are distinguished by *flavour* quantum numbers, i.e. the SM contains six flavours of quarks and six flavours of leptons.

Bosons have integer spin and mediate the three forces through their exchange. The photon mediates the electromagnetic force, the $W^{\pm}$ and Z bosons mediate the weak interaction, and the gluon mediates the strong interaction. The Higgs boson is

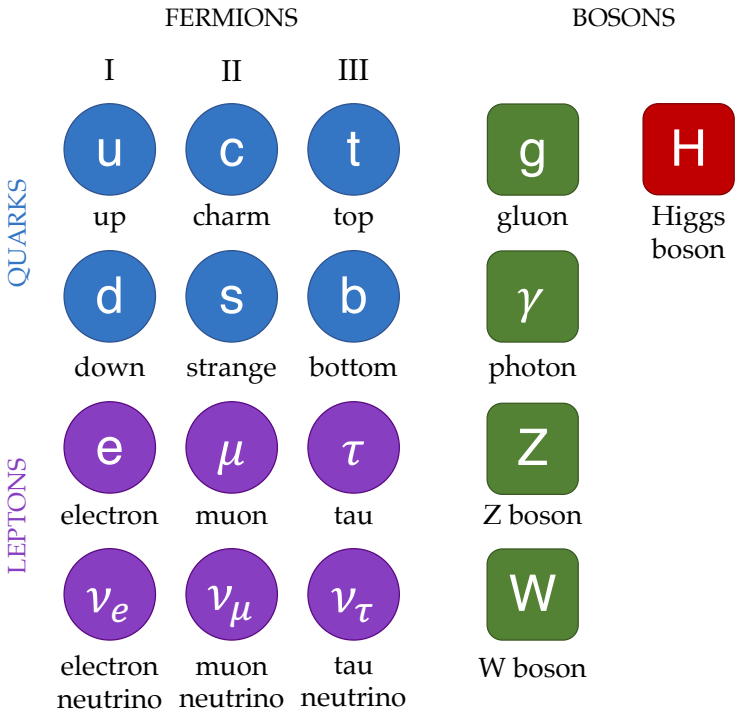


Figure 3.1: Fundamental particles of the Standard Model

required to generate the masses of the $W^{\pm}$ and Z bosons via spontaneous symmetry breaking, and also to generate the fermion masses through a Yukawa coupling.

3.2 Electroweak interaction

The signal decay analysed in this work proceeds via an electroweak process. The electroweak theory accounts for both the electromagnetic and weak interactions, treating them as a single unified interaction. This theory was first proposed by Glashow, Weinberg, and Salam [9–11]. The theory predicts four massless gauge bosons, three of which ($W^{\pm}, Z$) become massive by interacting with the Higgs field, whilst the photon remains massless.

Unlike the strong force, the electroweak interaction can produce quark flavour-changing processes. These are mediated by the W boson, and the phenomenon is described by the CKM mechanism [39]. Experimentally, it is observed that quark transitions within a generation are favoured. Transitions across one generation are

suppressed, and they are further suppressed across two generations.

The Z boson and the photon have similar interactions (neither carries electric charge nor changes quark flavours). However, only the Z boson couples to neutrinos. Only left-handed fermions interact weakly, but in the unified electroweak theory the photon has the same coupling to left and right-handed fermions, while the Z boson distinguishes them. Furthermore, the electroweak interaction leads to P and CP violating effects, as described earlier.

Semileptonic decays, along with the majority of processes studied at LHCb are weak decays of b-hadrons mediated by a W boson.

3.3 Lepton Flavour Universality

Lepton Flavour Universality (LFU) is an accidental symmetry of the SM; the symmetry is not caused by any underlying axiom of the theory. It implies that the three generations of leptons have identical couplings to the electroweak gauge bosons, and consequently the mass differences of the leptons are solely responsible for any differences in their interactions. Many extensions to the SM allow violation of LFU by introducing additional couplings to hypothetical particles. These new particles can be charged Higgs bosons in the minimal supersymmetric extension to the SM, as described in Ref. [40–42]. Alternatively, models involving $W'^{\pm}$ [43–45] and Z' [46, 47] bosons or leptoquarks [48–50] also predict LFU violation. Any experimental evidence of Lepton Flavour Non-Universality would clearly imply the existence of physics beyond the SM [1].

One common way of testing LFU is by measuring branching fraction ratios of the form:

$$\mathcal{R}(H_c) = \frac{\mathcal{B}(H_b \to H_c\, \tau\nu_\tau)}{\mathcal{B}(H_b \to H_c\, l\nu)}\,, \tag{3.1}$$

[1] Lepton flavour violating processes have been observed in the form of neutrino oscillations. Although this phenomenon can be well described by the PMNS model [51], some argue that it isn't a SM process as the origin of neutrino mass is not yet understood. To avoid confusion, all further references to lepton favour violation can be taken to mean *charged* lepton flavour violation.

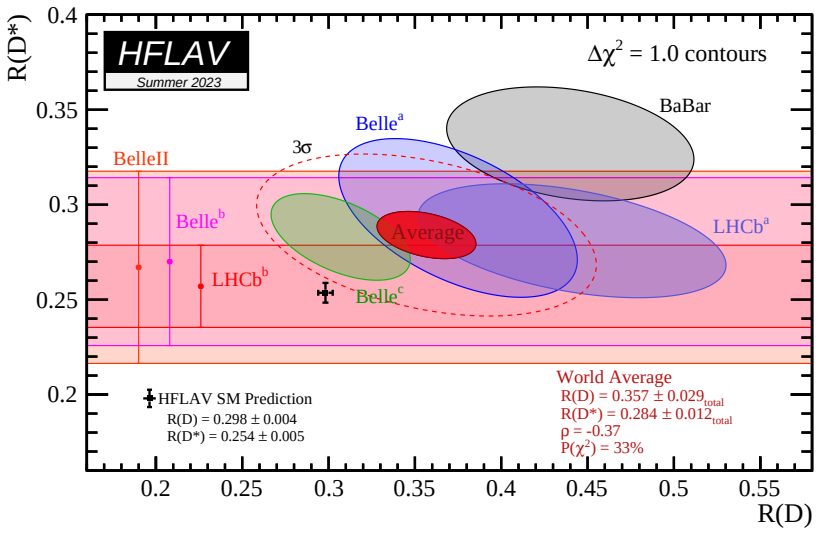


Figure 3.2: Measurements of $R(D)$ and $R(D^*)$ from BaBar, Belle and LHCb, and their average values. The contours correspond to a 68% confidence level for the horizontal bands and 39% for the ellipses. The SM predictions are also shown [57]. The average experimental prediction deviates from the SM value by 3.3σ. The horizontal bands are measurements of $R(D^*)$ [52,58,59]. The ellipses are simultaneous measurements of $R(D^0)$ and $R(D^*)$ [21,53,60].

where H_b is a beauty hadron, H_c is a charm hadron and $l = \{e, \mu\}$. These measurements have been performed at "B factories" – experiments that are designed to study the decays of B mesons, namely BaBar, Belle and Belle II – as well as at LHCb. So far, the measurements have involved $H_c = \{D^{*+}, D^0, \Lambda_c^{(*)+}, J/\psi\}$ with $H_b = \{B^0, B^+, B_c^+, \Lambda_b^0\}$ [21, 52–56]. Figure 3.2 shows the experimental average values for $R(D)$ and $R(D^*)$, along with the SM predicted values. The combined $R(D)$ and $R(D^*)$ averages (including the correlation between the two) differ with the SM predictions by 3.3σ.

Measurements of this type are sensitive tests of the SM, as the ratio allows cancellation of systematic uncertainties from hadronic form factors, as well as efficiency effects. Similar measurements have been made for processes involving $b \to s l^+ l^-$ transitions. In particular, LHCb has measured $\mathcal{R}(K^{(*)})$ [61], defined as

$$\mathcal{R}(K^{(*)}) = \frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}e^+e^-)}\,, \tag{3.2}$$

which was found to be consistent with the SM. Furthermore, an angular analysis of

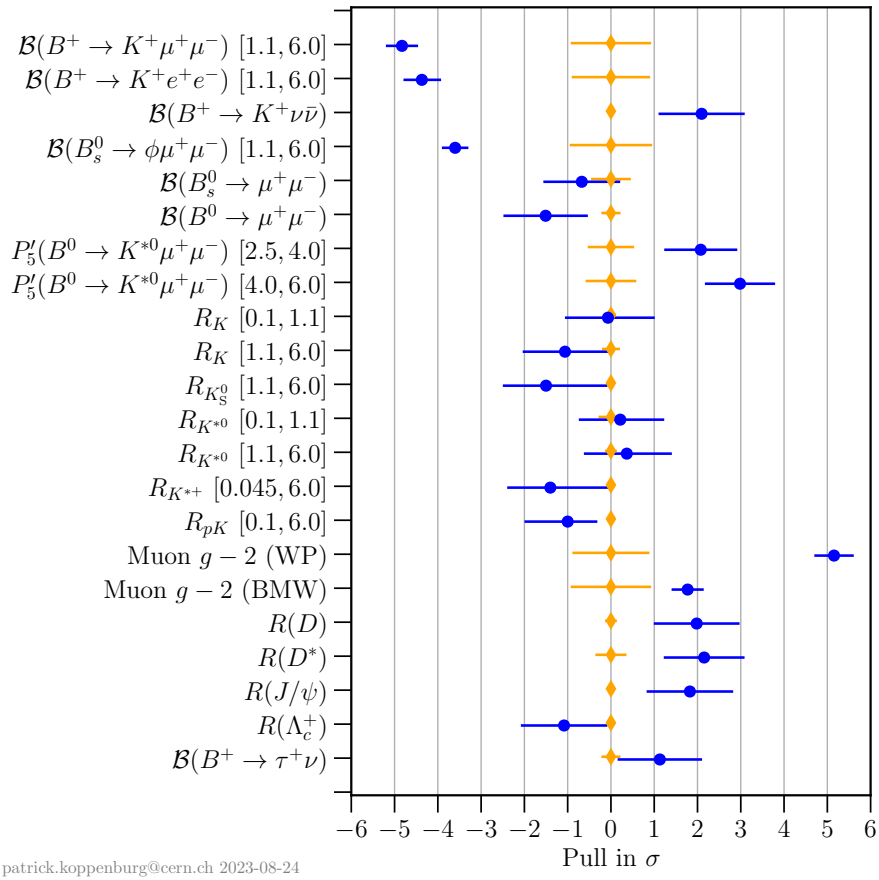


Figure 3.3: The significance of measurements for a selected set of flavour parameters is shown. For each entry, the SM expectation (orange diamond) is set to zero, the experimental and theory uncertainties are summed in quadrature and their sum is normalised to unity. The experimental value is then shifted and scaled accordingly [64].

$B^0 \to K^{*0}\mu^+\mu^-$ decays has been performed by LHCb [62,63], and a 3.7σ discrepancy with the SM was found for one of the angular observables (P_5') in one q^2 bin.

More generally, there has recently been a pattern of experimental signatures for flavour physics processes that indicate possible discrepancies with the SM. Some of these are shown in Fig. 3.3. These experimental hints of deviations from the SM motivate further study of $B \to D^* l\nu$ processes at LHCb. Importantly, the branching fraction ratio measurements alone cannot necessarily distinguish between the possible NP scenarios. Instead, measurements of the angular observables provide more complete information of the decay kinematics, so are a more powerful test of NP. Explicitly, the total decay rate for a $B \to D^* l\nu$ process depends on only four of the 12 angular observables that fully describe the decay. As a consequence, angular

analyses which measure all of the observables can reveal the presence of NP, even if the branching fraction measurements become consistent with the SM in the future.

3.4 $B \to D^* l\nu$ decays

In general, difficulties arise when computing the weak decays of hadrons. The reason for this is that the combination of the strong and electroweak interactions must be considered in order to describe them precisely. The SM amplitude for a semileptonic B meson decay is described by the weak Hamiltonian:

$$H_W = \frac{4G_F}{\sqrt{2}} V_{cb} [\bar{c}\gamma_\mu P_L b][\bar{l}\gamma^\mu P_L \nu_l] \, , \tag{3.3}$$

where G_F is the Fermi constant, V_{cb} is the relevant CKM matrix element, $\bar{c}$ and b are the quark spinors, $\bar{l}$ and ν_l are the lepton spinors and γ^μ represents the four conventional gamma matrices. P_L is the left-handed projection operator:

$$P_L = \frac{1}{2}(1 - \gamma_5) \, . \tag{3.4}$$

The $B^0 \to D^{*-}\mu^+\nu_\mu$ matrix element, $\mathcal{M}$, can be written down from the weak Hamiltonian in Eq. (3.3) as:

$$\mathcal{M} = \langle F|H_W|I\rangle \, , \tag{3.5}$$

where I and F are the initial and final state vectors respectively. To obtain the total decay rate, the matrix element is squared, summed over the final state polarisation states, and integrated over the phase space of the final state particles.

This matrix element can be factorised into the leptonic and hadronic parts (if higher-order electroweak corrections are ignored). The leptonic component can be calculated with a relatively high accuracy, since the electroweak contributions can be treated to lowest order in perturbation theory. However, the hadronic component of the matrix element involves non-perturbative QCD effects, and they must be approximated by *form factors*, as explained in 3.4.1.

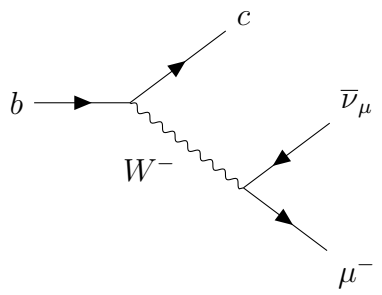


Figure 3.4: Feynman diagram for the quark transition in the $\overline{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu$ decay, mediated by a W boson

The Feynman diagram for the $b \to c$ quark transition via the weak interaction is shown in Fig. 3.4.

3.4.1 $B \to D^* l\nu$ form factors

For composite particles, it is very difficult or sometimes impossible to write down the analytic form for the hadronic part of the matrix element exactly. Instead, it is conventionally described by momentum-dependent functions known as form factors. These functions are Lorentz-invariant, and they approximate the effects of the strong interaction in the $B \to D^*$ transition. The form factors can be written down in several parameterisations which make different assumptions.

For example, dispersive constraints are often used to constrain the form factor shapes at high momentum transfer. The constraints can be based on principles of analyticity (requiring the form factors to be analytic functions in the complex plane), and unitarity (constraining the scattering amplitudes so that their associated probabilities are limited to the interval $[0, 1]$). These approaches are used in the method developed by Caprini, Lellouch and Neubert [65], referred to as the CLN scheme, and in a more general method developed by Boyd, Grinstein and Lebed [66], known as the BGL scheme.

Both the CLN and BGL schemes employ Taylor expansions in the variable z, where z is related to the hadronic recoil, w, defined as:

$$w \equiv \frac{p_B \cdot p_{D^*}}{m_B m_{D^*}} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}} \,, \tag{3.6}$$

where p_B and p_{D^*} are the four-momenta of the B and D^* mesons respectively and m_B, m_{D^*} are their masses. q^2 is the invariant mass of the lepton-neutrino system. z is defined as:

$$z = \frac{\sqrt{w+1} - \sqrt{2N}}{\sqrt{w+1} + \sqrt{2N}} , \tag{3.7}$$

with the value $N = 1$ conventionally chosen since it places the point of zero recoil ($w = 1$) as $z = 0$. This is a convenient conformal transformation, which helps to deal with singularities in the form factor expressions, and thus improves the convergence. z varies smoothly across the q^2 range, and is 0 at the maximum physical value of $q^2 = (p_B - p_{D^*})^2$. The CLN scheme is simpler than the BGL scheme, with fewer free parameters. However, the CLN scheme is model dependent and some have suggested [67] that the extra constraints result in theoretical uncertainties that are no longer negligible compared to current experimental precision.

Furthermore, the QCD light-cone sum rule approach [68] can be applied. This combines information from QCD sum rules with the theory of hard exclusive processes, allowing additional information to be incorporated on the QCD correlation functions.

Another approach is to use lattice QCD, which simulates the behaviour of quarks and gluons on a space-time lattice. This can be used to calculate the hadronic form factors directly [69], or to constrain the form factors by calculating related quantities, such as QCD correlation functions [70]. However, currently the most recent lattice QCD calculations are only available for the high q^2 region, see e.g. Ref. [71]. Thus to obtain the form factors in the whole physical region, the lattice QCD calculations are either extrapolated to low q^2, or they are combined with some of the constraints mentioned above.

The Standard Model process for the $B \to D^*$ transition involves vector and axial vector matrix elements, which are related to the form factor parameters using the Heavy Quark Effective Theory (HQET) [72]:

$$\frac{\langle D^*(v_{D^*}, \epsilon) \left| \bar{c}\gamma^\mu b \right| B(v_B) \rangle}{\sqrt{m_B m_{D^*}}} = h_V(w) \varepsilon^{\mu\nu\rho\sigma} v_{B,\nu} v_{D^*,\rho} \epsilon^*_\sigma , \tag{3.8}$$

$$\frac{\langle D^*\left(v_{D^*},\epsilon\right)|\bar{c}\gamma^\mu\gamma^5 b|\,B(v_B)\rangle}{\sqrt{m_B m_{D^*}}} = ih_{A_1}(w)(1+w)\epsilon^{*\mu} - i\left[h_{A_2}(w)v_B^\mu + h_{A_3}(w)v_{D^*}^\mu\right]\epsilon^*\cdot v_B\ . \tag{3.9}$$

v_B and v_{D^*} are the 4-momenta of the B and D^* mesons respectively, ϵ is the spin vector of the D^*, and $h_V, h_{A_1}, h_{A_2}, h_{A_3}$ are the form-factor parameters.

From this, the decay rate can be written in terms of helicity amplitudes corresponding to the polarisation of the off-shell W boson [73]:

$$H_\pm(w) = (w+1)\left[h_{A_1}(w) \mp \sqrt{\frac{w-1}{w+1}}h_V(w)\right], \tag{3.10}$$

$$H_0(w) = y(w+1)\left\{(w-r)h_{A_1}(w) - (w-1)[rh_{A_2}(w) + h_{A_3}(w)]\right\}, \tag{3.11}$$

$$H_S(w) = y\sqrt{w^2-1}\left[(w+1)h_{A_1}(w) - (1-wr)h_{A_2}(w) - (w-r)h_{A_3}(w)\right], \tag{3.12}$$

where $r = m_{D^*}/m_B$ and $y^2(1-2wr+r^2) = 1$. Using the helicity amplitudes, the $B^0 \to D^{*-}\mu^+\nu_\mu$ decay rate can be written as:

$$\begin{aligned}\frac{d\Gamma}{dw} = |V_{cb}|^2|\eta_{ew}|^2\frac{G_F^2 M_B^5}{16\pi^3}\left(1-\frac{m_l^2}{q^2}\right)^2 r^3\sqrt{w^2-1}\\ \left\{\frac{1}{3y^2}\left(1+\frac{m_l^2}{2q^2}\right)[|H_+(w)|^2 + |H_-(w)|^2 + |H_0(w)|^2] + \frac{m_l^2}{2M_B^2}|H_S(w)|^2\right\},\end{aligned} \tag{3.13}$$

where $\eta_{ew} = 1.0066 \pm 0.0050$ [74], the one loop electroweak correction factor (the leading order correction to the tree-level process). In the above expression, terms involving the lepton mass, m_l^2, are often neglected as they are negligible in the muonic case. In particular, this means the scalar amplitude $H_S(w)$ is often neglected. Then, the decay rate is given by [75]:

$$\frac{d\Gamma}{dw} = \frac{G_F^2 m_B^5}{48\pi^3}\,|V_{cb}|^2\,\sqrt{w^2-1}\,P(w)\,\left(\eta_{\rm ew}\mathcal{F}(w)\right)^2\ , \tag{3.14}$$

where $P(w)$ is a phase-space factor defined as:

$$P(w) = r^3(1-r)^2(w+1)^2\left(1+\frac{4w}{w+1}\frac{1-2rw+r^2}{(1-r)^2}\right), \tag{3.15}$$

and the form-factor $\mathcal{F}(w)$ is defined as:

$$|\mathcal{F}(w)|^2 = \frac{1-2wr+r^2}{w+1} \frac{|H_+(w)|^2 + |H_-(w)|^2 + |H_0(w)|^2}{(5w+1)(1-r)^2 - 8rw(w-1)} \,. \tag{3.16}$$

3.4.2 CLN parameterisation

In the CLN parameterisation, they propose the form-factor function $\mathcal{F}(w)$ can be described by:

$$\mathcal{F}(w) \propto 1 - \rho^2 z + cz^2 - dz^3 \,. \tag{3.17}$$

This has the advantage of a relatively few number of free parameters, however the constraints required make the parameterisation model dependent.

3.4.3 BGL parameterisation

The generic form factors, F_i, can be written as [76, 77]:

$$F_i(z) = \frac{1}{P_i(z)\phi_i(z)} \sum_{n=0}^{\infty} a_n z^n \,, \tag{3.18}$$

where the outer functions ($\phi_i(z)$) and the Blaschke factors ($P_i(z)$) are defined in Ref. [76]. The coefficients a_n are constrained by the unitarity bound:

$$\sum_{n=0}^{\infty} |a_n|^2 \leq 1 \,. \tag{3.19}$$

In the physical range $m_l^2 \leq q^2 \leq (m_B - m_{D^*})^2$, z takes the values $0 \leq z \leq 0.056$, so the series expansion can be truncated after a few terms with a very small loss in accuracy[2].

The BGL parameterisation for the $B \to D^*$ transition involves four form factors which are written in the notation of Eq. (3.18):

$$f(z) = \frac{1}{P_{1^+}(z)\phi_f(z)} \sum_{n=0}^{\infty} b_n z^n \,, \tag{3.20}$$

[2]The CLN scheme also involves powers of z, but is not strictly a power series as the number of terms is fixed, so there is no equivalent "truncation".

$$g(z) = \frac{1}{P_{1^-}(z)\phi_g(z)} \sum_{n=0}^{\infty} a_n z^n \ , \tag{3.21}$$

$$\mathcal{F}_1(z) = \frac{1}{P_{1^+}(z)\phi_{\mathcal{F}_1(z)}} \sum_{n=0}^{\infty} c_n z^n \ , \tag{3.22}$$

$$\mathcal{F}_2(z) = \frac{1}{P_{0^-}(z)\phi_{\mathcal{F}_2(z)}} \sum_{n=0}^{\infty} d_n z^n \ . \tag{3.23}$$

The values of a_i, b_i, c_i, and d_i can be obtained from lattice QCD calculations, as reported in Ref. [78] for example.

3.4.4 Summary

In general, many semileptonic analyses have chosen a particular form factor parameterisation and quoted results based on that theoretical approach. However, this has led to tensions between results obtained from different form factor schemes. For example Refs. [79] and [80] report that using Belle data, different $|V_{cb}|$ values are obtained from the BGL and CLN schemes. To avoid this issue, the analysis described in this thesis uses a novel model-independent approach. This ensures that the measured observables are fully interpretable within any form-factor scheme, and thus gives much greater flexibility when testing theoretical models.

3.5 Derivation of the angular decay rate

The derivation of the decay rate governing $B^0 \to D^{*-}\mu^+\nu_\mu$ decays is presented here, mainly following the formalism from Refs. [73, 81, 82].

The calculation of the angular decay rate for the $B^0 \to D^{*-}\mu^+\nu_\mu$ transition proceeds from Fermi's Golden Rule. This gives the first order approximation for a transition from an initial state $|i\rangle$ to a final state $|f\rangle$, and its most general statement is:

$$\Gamma_{fi} = \frac{2\pi}{\hbar} |\langle f|V_{fi}|i\rangle|^2 \rho(E_f) \ , \tag{3.24}$$

where Γ_{fi} is the transition rate, V_{fi} is the Hamiltonian for the transition, and $\rho(E_f)$ is the phase-space factor. If Lorentz invariance is imposed and the states

are normalised to $2E$ particles per unit volume (where E refers to the energy), Eq. (3.24) can be written as:

$$\Gamma_{fi} = (2\pi)^4 \frac{1}{2E_i} \int |\mathcal{M}_{fi}|^2 \prod_{j=1}^{4} \frac{d^3 p_j}{(2\pi)^3\, 2E_j} \delta\left(p_i - \sum_{k=1}^{4} p_k\right) , \tag{3.25}$$

where p_i is the 4-momentum of the initial state, E_i is its energy, $\mathcal{M}_{fi}$ is the matrix element for the transition, and δ represents the Dirac delta function. The squared matrix element is multiplied by a phase-space factor, which is evaluated in the following section. The calculation of the matrix element is then outlined.

3.5.1 Phase-space factor

The above expression relates to the case of four particles in the final state, as the relevant transition here is $B \to D\pi\mu\nu_\mu$. The four-body phase space can be written as the product of two-body phase spaces [82]:

$$\begin{aligned} d\Phi_4 &= (2\pi)^4 \int \prod_{i=1}^{4} \frac{d^3 p_i}{(2\pi)^3\, 2E_i} \delta\left(p - \sum_{j=1}^{4} p_j\right) , \\ &= \frac{dm_{12}^2}{2\pi} \frac{dm_{34}}{2\pi}\, d\Phi_2(m_{12}, m_{34})\, d\Phi_2(p_1, p_2)\, d\Phi_2(p_3, p_4) , \end{aligned} \tag{3.26}$$

where $d^3p = p^2\, dp\, d\cos\theta\, d\phi$ is used to define the two-body phase space $d\Phi_2$ as:

$$d\Phi_2(p_i, p_j) = \frac{1}{16\pi^2} \frac{|\vec{p_i}|}{m_{ij}} d\cos\theta_i\, d\phi_i , \tag{3.27}$$

with $\vec{p_i}$ as the three-momentum in the i, j rest frame, and θ_i, ϕ_i representing the polar and azimuthal angles of particle i respectively. Therefore in this case, $d\Phi_4$ can be evaluated as:

$$d\Phi_4 = \frac{1}{64(2\pi)^8}\, dm_{D\pi}^2\, dq^2 \frac{|\vec{q}|}{m_B}\, d\cos\theta_{D\pi}\, d\phi_{D\pi} \frac{|\vec{p}_D|}{m_{D\pi}}\, d\cos\theta_D\, d\phi_D \frac{|\vec{p}_l|}{\sqrt{q^2}}\, d\cos\theta_l\, d\phi_l . \tag{3.28}$$

Using the 3-momentum definitions:

$$|\vec{p}_D| = \frac{\sqrt{\lambda(m_{D\pi}^2, m_D^2, m_\pi^2)}}{2m_{D\pi}} , \; |\vec{p}_l| = \frac{\sqrt{\lambda(q^2, m_l^2, 0)}}{2\sqrt{q^2}} , \tag{3.29}$$

where the lambda function $\lambda(x,y,z) = x^2 + y^2 + z^2 - 2(xy + xz + yz)$ has been used, and integrating over $(\theta_{D\pi}, \phi_{D\pi}, \phi_D)$:

$$d\Phi_4 = \frac{1}{64(2\pi)^6} \frac{|\vec{q}|}{m_B} \frac{|\vec{p_D}|}{m_{D\pi}} \left(1 - \frac{m_l^2}{q^2}\right) dq^2 \, dm_{D\pi}^2 \, d\cos\theta_D \, d\cos\theta_\ell \, d\chi \ , \tag{3.30}$$

where ϕ_l has been re-written as χ.

3.5.2 Matrix element

To evaluate the matrix element for the $B^0 \to D^{*-}\mu^+\nu_\mu$ decay, the full effective Hamiltonian for the process must first be considered [81]:

$$\begin{aligned}\mathcal{H}_{\text{eff}} = \sqrt{2} G_F V_{cb} \Big[&(1 + g_V)\,(\bar{c}\gamma_\mu b)\left(\bar{\ell}_L \gamma^\mu \nu_L\right) + (-1 + g_A)\,(\bar{c}\gamma_\mu\gamma_5 b)\left(\bar{\ell}_L \gamma^\mu \nu_L\right) \\ &+ g_S (\bar{c} b)\left(\bar{\ell}_R \nu_L\right) + g_P\,(\bar{c}\gamma_5 b)\left(\bar{\ell}_R \nu_L\right) \\ &+ g_T\,(\bar{c}\sigma_{\mu\nu} b)\left(\bar{\ell}_R \sigma^{\mu\nu} \nu_L\right) + g_{T5}\,(\bar{c}\sigma_{\mu\nu}\gamma_5 b)\left(\bar{\ell}_R \sigma^{\mu\nu} \nu_L\right)\Big] + \text{ h.c.}\end{aligned} \tag{3.31}$$

This includes potential NP couplings g_i, where $i \in \{S, P, V, A, T, T5\}$, that are compatible with Lorentz invariance. These represent scalar, pseudo-scalar, vector, axial vector, tensor, and pseudo-tensor contributions. This will treat any NP coupling in the same manner as the vector currents of the SM. If these terms are set to 0, this is equivalent to the SM Hamiltonian in Eq. (3.3). Following this, the matrix element is defined as:

$$\mathcal{M} = \langle D^* l\nu | \mathcal{H}_{\text{eff}} | B^0 \rangle \ . \tag{3.32}$$

This can be split into the leptonic and hadronic contributions:

$$\mathcal{M} = \langle D^* | H_\mu | B^0 \rangle \langle l\nu | L^\mu | 0 \rangle \ , \tag{3.33}$$

where H_μ represents the hadronic current and L^μ is the leptonic current. Including the decay of the D^* resonance:

$$\mathcal{M} = \langle D\pi | D^* \rangle \langle D^* | H_\mu | B^0 \rangle \langle l\nu | L^\mu | 0 \rangle BW_{D^*} \ . \tag{3.34}$$

BW_{D^*} is the Breit-Wigner function that represents the propagation of the resonant state:

$$\begin{aligned} BW_{D^*} &= \frac{1}{m_{D\pi}^2 - m_{D^*}^2 + i m_{D^*}\Gamma_{D^*}} \\ &\approx \frac{\pi}{m_{D^*}\Gamma_{D^*}}\delta(m_{D\pi}^2 - m_{D^*}^2) \ , \end{aligned} \tag{3.35}$$

where the narrow width of the D^* state (Γ_{D^*}) is used to make the approximation above. The matrix element can be re-written in terms of helicity amplitudes using the below relation for the W boson polarisation vectors:

$$\sum_\lambda \eta_\mu^*(\lambda)\eta_\nu(\lambda)\delta_\lambda = g_{\mu\nu}, \quad \delta_{0,\pm} = -\delta_t = -1 \ . \tag{3.36}$$

Then, the vector and axial vector matrix elements become:

$$\begin{aligned} \mathcal{M}_{V(A)}^{\lambda_{D^*},\lambda_\ell} &\propto \left\langle D^*\left(\lambda_{D^*}\right)\left|H^\mu\right|\bar{B}\right\rangle \left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|L_\mu\right|0\right\rangle \\ &= \sum_\lambda \eta_\mu^*(\lambda)\left\langle D^*\left(\lambda_{D^*}\right)\left|H^\mu\right|\bar{B}\right\rangle \eta_\nu(\lambda)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|L^\nu\right|0\right\rangle \delta_\lambda \\ &= \sum_\lambda \delta_\lambda H_{V(A),\lambda}^{\lambda_{D^*}} L_{V-A,\lambda}^{\lambda_\ell} \ . \end{aligned} \tag{3.37}$$

The helicity amplitudes H and L will be defined in the following sections. The corresponding matrix elements for the scalar and tensor currents can be similarly written in terms of the helicity amplitudes.

Leptonic amplitudes

The leptonic components of the matrix element can be written as:

$$\begin{aligned} L_{V-A,\lambda}^{\lambda_\ell}\left(q^2,\chi,\theta_\ell\right) &= \eta_\mu(\lambda)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\gamma^\mu\left(1-\gamma_5\right)\nu\right|0\right\rangle \ , \\ L_{S-P}^{\lambda_\ell}\left(q^2,\chi,\theta_\ell\right) &= \left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\left(1-\gamma_5\right)\nu\right|0\right\rangle \ , \\ L_{T-T5,\lambda\lambda'}^{\lambda_\ell}\left(q^2,\chi,\theta_\ell\right) &= -L_{T-T5,\lambda'\lambda}^{\lambda_\ell} = -i\eta_\mu(\lambda)\eta_\nu\left(\lambda'\right)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\sigma^{\mu\nu}\left(1-\gamma_5\right)\nu\right|0\right\rangle \ , \end{aligned} \tag{3.38}$$

where λ is the W boson helicity in the B reference frame, λ_ℓ is the lepton helicity in the W rest frame, and $\eta_\mu(\lambda)$ is the W polarisation vector. To calculate these expressions, the polarisation vectors for the W boson and the Dirac spinors are required. An example for the calculation of one term is presented.

To calculate the term $L^{+}_{V-A,0}$, the polarisation vector:

$$\eta^{\mu}(0) = \frac{1}{\sqrt{q^2}} \begin{pmatrix} |\boldsymbol{q}| \\ 0 \\ 0 \\ -q_0 \end{pmatrix} , \tag{3.39}$$

where:

$$|\boldsymbol{q}| = \frac{\sqrt{\lambda_{BD^{(*)}}(q^2)}}{2m_B}, \quad q_0 = \frac{m_B^2 - m_{D^{(*)}}^2 + q^2}{2m_B} , \tag{3.40}$$

is used. The spinors are given by:

$$u(\lambda = \pm 1/2) = \begin{pmatrix} \sqrt{E \mp |\boldsymbol{p}|}\xi_{\pm} \\ \sqrt{E \pm |\boldsymbol{p}|}\xi_{\pm} \end{pmatrix}, \quad v(\lambda = \pm 1/2) = \begin{pmatrix} -\sqrt{E \pm |\boldsymbol{p}|}\xi_{\mp} \\ \sqrt{E \mp |\boldsymbol{p}|}\xi_{\mp} \end{pmatrix} , \tag{3.41}$$

where E and $\boldsymbol{p}$ represent the energy and momentum of the particle and:

$$\xi_{+} = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}\, e^{i\chi} \end{pmatrix}, \quad \xi_{-} = \begin{pmatrix} -\sin\frac{\theta}{2}\, e^{-i\chi} \\ \cos\frac{\theta}{2} \end{pmatrix} . \tag{3.42}$$

Using these expressions and the gamma matrices in the Weyl basis gives:

$$L^{+}_{V-A,0}\left(q^2, \chi, \theta_\ell\right) = 2m_\ell \beta_\ell \cos\theta_\ell\, e^{-i\chi} , \tag{3.43}$$

with $\beta_\ell = \sqrt{1 - m_\ell^2/q^2}$. The corresponding expressions for the other terms are listed in Appendix B.

Hadronic amplitudes

Similarly to the leptonic amplitudes, the hadronic components can be written as:

$$\begin{aligned} H^{\lambda_{D^*}}_{V(A),\lambda}\left(q^2\right) &= \left(g_{V(A)} \pm 1\right) \eta^{*}_{\mu}(\lambda) \left\langle D^*\left(\lambda_{D^*}\right) |\bar{c}\gamma^{\mu}\left(\gamma_5\right) b| \, \bar{B} \right\rangle \\ H_{S(P),\lambda}\left(q^2\right) &= g_{S(P)} \left\langle D^*\left(\lambda_{D^*} = 0\right) |\bar{c}\left(\gamma_5\right) b| \, \bar{B} \right\rangle \\ H^{\lambda_{D^*}}_{T(T5),\lambda\lambda'}\left(q^2\right) &= -H^{\lambda_{D^*}}_{T,\lambda'\lambda}\left(q^2\right) = i g_{T(T5)} \eta^{*}_{\mu}(\lambda) \eta^{*}_{\nu}\left(\lambda'\right) \left\langle D^*\left(\lambda_{D^*}\right) |\bar{c}\sigma^{\mu\nu}\left(\gamma_5\right) b| \, \bar{B} \right\rangle , \end{aligned} \tag{3.44}$$

where λ_{D^*} is the helicity of the D^*. The full set of expressions for the amplitudes are shown in Appendix B.

3.5.3 Decay rate

Combining the amplitudes to calculate the matrix element, and adding in the phase-space factor from Eq. (3.30) yields the full angular distribution for $B \to D^*(\to D\pi)\mu\nu_\mu$ in terms of the decay angles $\theta_D, \theta_\ell, \chi$ and q^2:

$$\frac{d^5\Gamma}{dq^2\, dm^2_{D\pi}\, d\Omega} = \frac{\sqrt{\lambda_{BD^*}(q^2)}}{256(2\pi)^6 m_B^3}\left(1 - \frac{m_\ell^2}{q^2}\right)\frac{|\hat{p_D}|}{m_{D\pi}}\left|\mathcal{M}(B \to D^*(\to D\pi)\mu\nu_\mu)\right|^2 , \quad (3.45)$$

where $d\Omega = d\cos\theta_D\, d\cos\theta_\ell\, d\chi$ and $m_{D\pi}$ is the mass of the D,π combination. Integrating over $m^2_{D\pi}$ around the D^{*+} resonance, the angular distribution can be written as:

$$\begin{aligned}\frac{d^4\Gamma}{dq^2 d(\cos\theta_D) d(\cos\theta_L) d\chi} = \frac{9}{32\pi}\Big\{ & I_{1c}(q^2)\cos^2\theta_D + I_{1s}(q^2)\sin^2\theta_D \\ & + \Big[I_{2c}(q^2)\cos^2\theta_D + I_{2s}(q^2)\sin^2\theta_D\Big]\cos 2\theta_L \\ & + \Big[I_{6c}(q^2)\cos^2\theta_D + I_{6s}(q^2)\sin^2\theta_D\Big]\cos\theta_L \\ & + \Big[I_3(q^2)\cos 2\chi + I_9(q^2)\sin 2\chi\Big]\sin^2\theta_L \sin^2\theta_D \\ & + \Big[I_4(q^2)\cos\chi + I_8(q^2)\sin\chi\Big]\sin 2\theta_L \sin 2\theta_D \\ & + \Big[I_5(q^2)\cos\chi + I_7(q^2)\sin\chi\Big]\sin\theta_L \sin 2\theta_D\Big\} .\end{aligned} \quad (3.46)$$

Importantly, this form of the rate equation allows all the form-factor and q^2 dependence to be decoupled from the angular variables, and contained in the $I_X(q^2)$ coefficients. These coefficients are related to the helicity amplitudes introduced earlier, the full definitions are shown in Appendix C.

Note that Γ in Eq. (3.46) represents the average rate for the B^0 and $\overline{B}^0$ decays, and the $I_X(q^2)$ coefficients are the CP-averaged angular observables. This analysis measures the rate-normalised, CP-averaged, q^2-integrated I_X values, which are referred to using a shorthand throughout this thesis [3]:

$$\langle I_X \rangle = \frac{1}{\Gamma(B \to D^*\mu\nu)} \times \int_{m_\ell^2}^{(m_B - m_{D^*})^2} I_X(q^2)\, dq^2 \to I_X \quad (3.47)$$

[3] No further reference to the q^2 variation of the I_X functions will be made, as it is not relavent for this work. As such, all further references to I_X in this thesis relate to the quantity defined in Eq. (3.47).

where:

$$\Gamma(B \to D^* \mu\nu) = \frac{1}{2}\Big(\Gamma(B^0 \to D^{*-}\mu^+\nu_\mu) + \Gamma(\overline{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)\Big) \tag{3.48}$$

3.5.4 I_X theoretical predictions

The SM values for the I_X can be calculated using the EOS software package [83], which allows the estimation of flavour observables and their uncertainties. To perform the calculations, values for the $B \to D^*\mu\nu$ form factors from Ref. [84] were taken as inputs. The calculated values are shown in Tab. 3.1.

Coefficient	Value
I_{1c}	0.54 ± 0.05
I_{1s}	0.34 ± 0.04
I_{2c}	-0.53 ± 0.05
I_{2s}	0.114 ± 0.012
I_{6c}	0.015 ± 0.003
I_{6s}	-0.27 ± 0.05
I_3	-0.180 ± 0.018
I_4	-0.305 ± 0.008
I_5	0.219 ± 0.030
I_7	0
I_8	0
I_9	0

Table 3.1: SM values and uncertainties of I_X calculated with the EOS software. I_7, I_8, and I_9 are defined to be exactly 0 in the SM so they have no associated uncertainty.

3.5.5 Other observables

The differential decay rate is parameterised by 12 observables, denoted I_X above. Alternatively, a different set of observables can be constructed from the I_X. Although this analysis measures the I_X directly, these alternative observables are relevant as some of them have direct physical interpretations. They are [81]:

- **Decay rate**

$$\Gamma = \frac{1}{4}(3I_{1c} + 6I_{1s} - I_{2c} - 2I_{2s}) \ . \tag{3.49}$$

- **Forward-backward asymmetry**

$$\mathcal{A}_{FB} = \frac{3}{8}\frac{(I_{6c} + 2I_{6s})}{d\Gamma/dq^2} \ . \tag{3.50}$$

- **D^* polarisation fraction**

$$F_L^{D^*} = \frac{3I_{1c} - I_{2c}}{3I_{1c} - I_{2c} + 6I_{1s} - 2I_{2s}} \ . \tag{3.51}$$

- **$R_{A,B}$**

$$R_{A,B} = \frac{\Gamma_A}{\Gamma_B} \ , \tag{3.52}$$

where

$$\begin{aligned} \Gamma_A &= \frac{1}{4}(3I_{1c} + 2I_{6s} - 3I_{2c} - 6I_{2s}) \ , \\ \Gamma_B &= \frac{1}{2}(3I_{1c} + 2I_{6s} + I_{2c} + 2I_{2s}) \ . \end{aligned} \tag{3.53}$$

- **$A_3 - A_9$**

$$A_3 = \frac{1}{2\pi}\frac{I_3}{d\Gamma/dq^2} \ , \tag{3.54}$$

$$A_4 = -\frac{2}{\pi}\frac{I_4}{d\Gamma/dq^2} \ , \tag{3.55}$$

$$A_5 = -\frac{3}{4}\frac{I_5}{d\Gamma/dq^2} \ , \tag{3.56}$$

$$A_{6s} = -\frac{27}{8}\frac{I_{6s}}{d\Gamma/dq^2} \ , \tag{3.57}$$

$$A_7 = -\frac{3}{4}\frac{I_7}{d\Gamma/dq^2} \ , \tag{3.58}$$

$$A_8 = \frac{2}{\pi}\frac{I_8}{d\Gamma/dq^2} \ , \tag{3.59}$$

$$A_9 = \frac{1}{2\pi}\frac{I_9}{d\Gamma/dq^2} \ . \tag{3.60}$$

4 Angular analysis methodology

The first step in performing the angular analysis is to calculate the relevant kinematic variables describing the signal decay mode, hereafter referred to as the *decay angles*. These angles are defined by the momenta of the final-state particles in a particular rest frame.

In this chapter, the definitions of the angles and the procedure used to calculate them in data and simulation are outlined. Then, a full description of the novel template fit methodology used to measure the angular coefficients is given.

4.1 Decay angle definitions

The $B^0 \to D^{*-}\mu^+\nu_\mu$ decay mode can be described by three decay angles $\cos\theta_D$, $\cos\theta_L$ and χ (as shown in Fig. 4.1), and the invariant mass of the muon-neutrino system, q^2.

The three decay angles are defined below for the B^0 decay, using the following convention. $\vec{p_a}^b$ is a vector representing the momentum of particle a in the rest frame of particle b. The unit normal vector to the plane defined by the momentum vectors of a and b in the rest frame of c can be defined as:

$$\hat{\eta}^c_{ab} = \frac{\vec{p}^c_a \times \vec{p}^c_b}{|\vec{p}^c_a \times \vec{p}^c_b|} \tag{4.1}$$

In this basis, the $\overline{B}^0$ decay is defined as a CP transformation of the B^0 decay. This is an equivalent formalism to the one described in Refs. [85, 86].

The angle θ_D

θ_D is the angle between the direction of the D meson in the D^* rest frame and the direction of the D^{*-} in the B rest frame

$$\cos\theta_D = \frac{\vec{p}_{D^0}^{D^{*-}} \cdot \vec{p}_{D^{*-}}^{B^0}}{|\vec{p}_{D^0}^{D^{*-}}||\vec{p}_{D^{*-}}^{B^0}|} = -\frac{\vec{p}_{D^0}^{D^{*-}} \cdot \vec{p}_{B^0}^{D^{*-}}}{|\vec{p}_{D^0}^{D^{*-}}||\vec{p}_{B^0}^{D^{*-}}|} \tag{4.2}$$

The angle θ_L

θ_L is the angle between the direction of the muon in the virtual W rest frame and the direction of the W in the B rest frame

$$\cos\theta_L = \frac{\vec{p}_{\mu}^{W} \cdot \vec{p}_{W}^{B^0}}{|\vec{p}_{\mu}^{W}||\vec{p}_{W}^{B^0}|} = -\frac{\vec{p}_{\mu}^{W} \cdot \vec{p}_{B^0}^{W}}{|\vec{p}_{\mu}^{W}||\vec{p}_{B^0}^{W}|} \tag{4.3}$$

The angle χ

χ is the azimuthal angle between the plane formed by the D^* and the plane formed by the W decay.

$$\cos\chi = \hat{\eta}_{\mu\nu_\mu}^{B^0} \cdot \hat{\eta}_{D^0\pi^-}^{B^0} \tag{4.4}$$

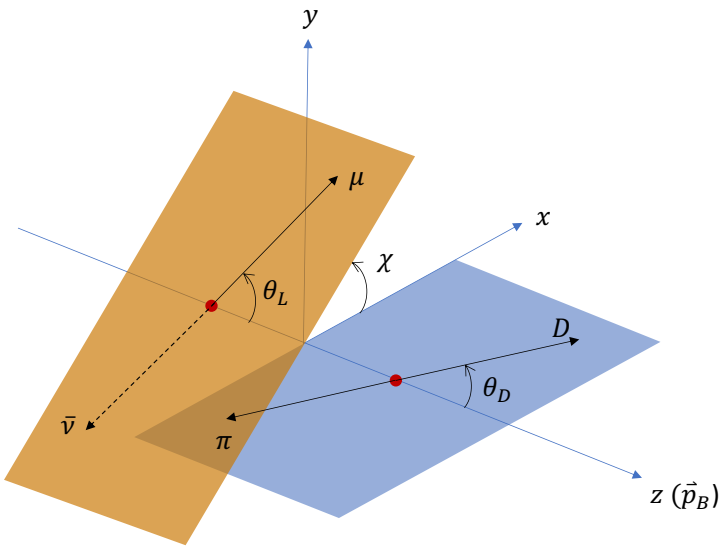


Figure 4.1: Schematic of the $\overline{B}^0 \to D^*(\to D\pi)l\bar{\nu}$ decay, defining the decay angles.

4.2 Calculation of the angles

4.2.1 B^0 rest frame approximation

The signal decay produces a neutrino, which cannot be detected at LHCb. Additionally, at a hadron collider, it is not possible to reconstruct the rest frame of the initial collision. This is in contrast to e^+e^- colliders where knowledge of the initial rest frame can be exploited to calculate the "missing" energy and momentum carried by the neutrino. Instead, for this analysis, the B^0 rest frame must be approximated in order to calculate the decay angles.

If the flight direction of the B^0 meson is known, it is possible to estimate its momentum up to a two-fold ambiguity [87]. The total momentum can be split into the reconstructed, or *visible* component, $\vec{P}_{\text{vis}}$, and the *missing* component, $\vec{P}_{\text{miss}}$. In this case, $\vec{P}_{\text{vis}}$ corresponds to the $D^*\mu$ system, which is fully reconstructed, and $\vec{P}_{\text{miss}}$ is the neutrino momentum. The visible component is decomposed into transverse and parallel components, with respect to the B^0 flight vector, $\vec{F}$. $\vec{F}$ is calculated

from the difference between the B^0 vertex and the primary vertex positions:

$$P_{\rm vis}^{\perp} = \left| \vec{P}_{\rm vis} \times \frac{\vec{F}}{|\vec{F}|} \right| , \tag{4.5}$$

$$P_{\rm vis}^{\parallel} = \vec{P}_{\rm vis} \cdot \frac{\vec{F}}{|\vec{F}|} . \tag{4.6}$$

The transverse component of the missing momentum, $P_{\rm miss}^{\perp}$ must be equal to $P_{\rm vis}^{\perp}$. Since the B^0 mass (m_B) is well known [88], the parallel component of the missing momentum can be found from a quadratic equation:

$$P_{\rm miss}^{\parallel} = -a \pm \sqrt{r} , \tag{4.7}$$

where

$$a = \frac{P_{\rm vis}^{\parallel} \left(m_B^2 - M_{\rm vis}^2 - 2 \left(P_{\rm vis}^{\perp} \right)^2 \right)}{2 \left(\left(P_{\rm vis}^{\parallel} \right)^2 - E_{\rm vis}^2 \right)} \tag{4.8}$$

and

$$r = \frac{E_{\rm vis}^2 \left(m_B^2 - M_{\rm vis}^2 - 2 \left(P_{\rm vis}^{\perp} \right)^2 \right)^2}{4 \left(\left(P_{\rm vis}^{\parallel} \right)^2 - E_{\rm vis}^2 \right)^2} + \frac{\left(E_{\rm vis} P_{\rm vis}^{\perp} \right)^2}{\left(P_{\rm vis}^{\parallel} \right)^2 - E_{\rm vis}^2} . \tag{4.9}$$

The terms $E_{\rm vis}$ and $M_{\rm vis}$ are the visible energy and mass, respectively. From this, two solutions can be obtained for the B^0 momentum:

$$P_{\pm} = P_{\rm vis}^{\parallel} - a \pm \sqrt{r} . \tag{4.10}$$

Note that due to resolution effects, some events have negative values of r, which is non-physical. These are subsequently removed from the analysis.

Two methods are considered for resolving the two-fold ambiguity in Eq. (4.10). The accuracy of the two methods can be evaluated using the signal MC. For every kinematic variable, the simulation contains both the *reconstructed* quantity (which emulates what is measured in real data), and the *true* quantity (which has detector resolution effects removed). It follows that for each method the reconstructed angular shape can be compared with the corresponding truth-level shape, to assess which method yields a closer approximation.

Firstly, the ambiguity can be resolved by requiring $r = 0$. It can be shown [89] that this is equivalent to requiring the opening angle between the B^0 flight vector and the $D^{*-}\mu^+$ momentum vector to be at a maximum:

$$\theta^{max}_{B^0,Y} = \arcsin\left(\frac{m^2_{B^0} - m^2_Y}{2m_{B^0}\,|\vec{p}_Y|}\right) \tag{4.11}$$

where Y represents the $D^{*-}\mu^+$ system and $\theta_{B^0,Y}$ is the angle between the B^0 flight vector and the momentum vector of Y. Conceptually, this can be seen as a way of "averaging" the two solutions for the B^0 momentum. This avoids the problem of dealing with events with negative r values, but makes no attempt to choose which of P_+ and P_- is the correct solution.

A more sophisticated way of dealing with the B^0 momentum ambiguity is to use an independent estimate to select the correct solution. This method was demonstrated in Ref. [90]. Figure 4.2 shows the two dimensional distribution of values for the estimated B^0 momentum, according to Eq. (4.10), for a MC sample of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays. On the plot, a vertical and horizontal band are clearly visible, which correspond to events where P_- and P_+ are the true solutions, respectively. The bands are relatively narrow, which means there is a strong motivation to be able to choose between P_+ and P_-.

This can be achieved by training a linear regression algorithm. Two variables are used in the training: the B^0 flight length $|\vec{F}|$ and $1/\sin\theta_{\text{flight}}$. $|\vec{F}|$ is related to the B^0 momentum (P_B) by:

$$P_B = \frac{m_B|\vec{F}|}{t} \tag{4.12}$$

where m_B is the B^0 mass and t is its decay time. θ_{flight} is the polar angle of the flight vector. Both $|\vec{F}|$ and $1/\sin\theta_{\text{flight}}$ are observed to be weakly correlated with the true B^0 momentum, P_{true}. Therefore a regression algorithm trained on these two variables is expected to produce an estimate of the B^0 momentum that is also correlated with P_{true}. This estimated momentum can then be used to pick between P_+ and P_-, by choosing whichever of these two solutions is closer to the estimate.

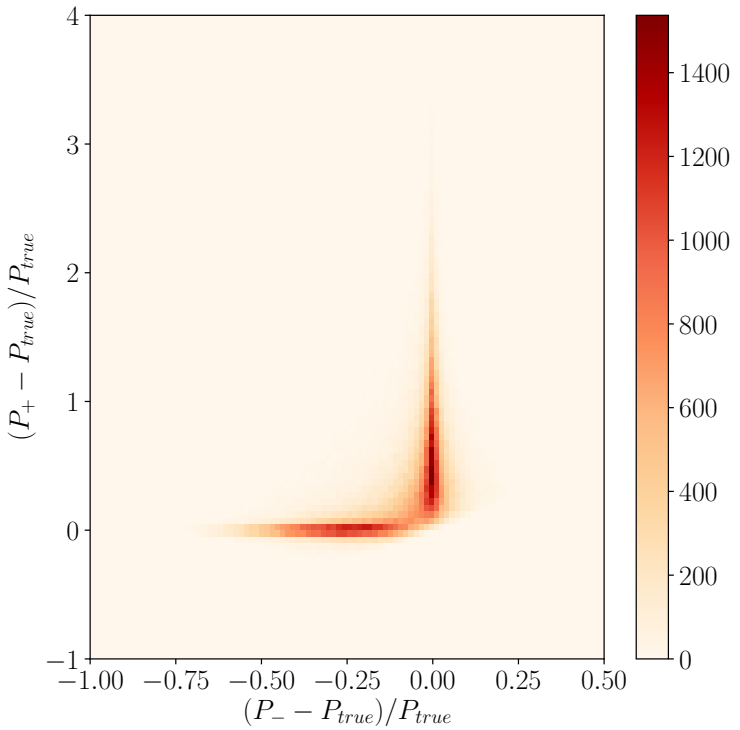
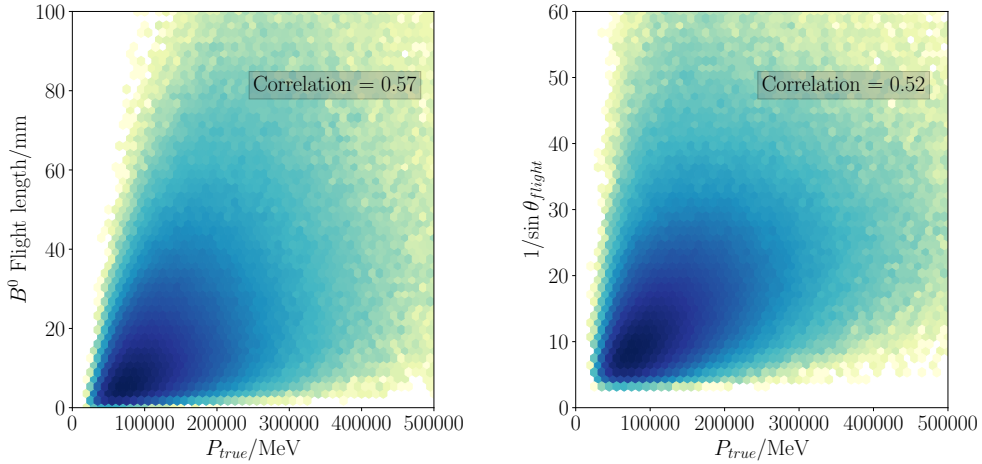
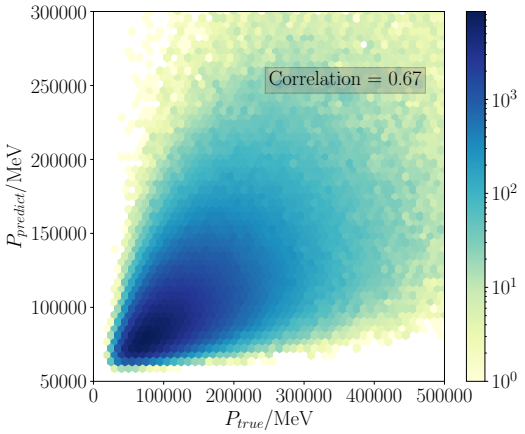


Figure 4.2: 2D density plot showing the solutions for the B^0 momentum, using $B^0 \to D^{*-}\mu^+\nu_\mu$ fullsim MC, according to Eq. (4.10)

Table 4.1 shows the performance of the algorithm. Although the two training variables are only weakly correlated with P_{true}, the linear regression algorithm uses this information to create an estimate for the B^0 momentum that has a stronger correlation with P_{true}. The correlations are also plotted as two-dimensional histograms in Fig. 4.3. This correlation indicates that the regressor estimate will pick the correct solution to Eq. (4.10) more than 50% of the time, which should give a better estimate than the previous method of setting r to zero. In addition to the linear regression algorithm, higher order models were tested. Figure 4.4 shows the effect of increasing the order of the regression model. As the order increases, the predicted B momentum moves slightly closer to the truth-level B momentum, but this shift is almost negligible. Since the training time increased considerably with the higher order models, no sufficient benefit was gained from changing the order of the model.



(a) Flight length **(b)** Flight angle



(c) Predicted momentum

Figure 4.3: (a) and (b) show correlations between the true B^0 momentum and the variables chosen to train the regression algorithm. (c) shows the correlation between the true B^0 momentum and the estimated momentum from the algorithm.

Variables	Correlation coefficient
P_{true} vs $1/\sin\theta_{\text{flight}}$	0.52
P_{true} vs $\lvert\vec{F}\rvert$	0.57
P_{true} vs $P_{predict}$	0.67

Table 4.1: Table showing the correlation coefficients between the true B^0 momentum and the regressor training variables, as well as the regressor estimate for the B^0 momentum ($P_{predict}$)

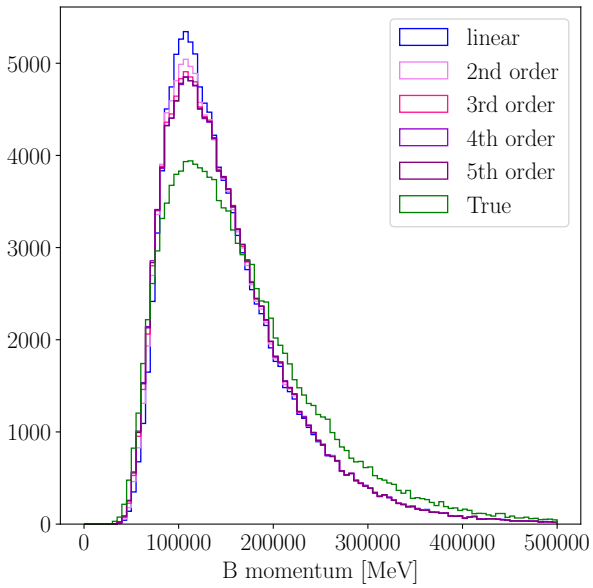
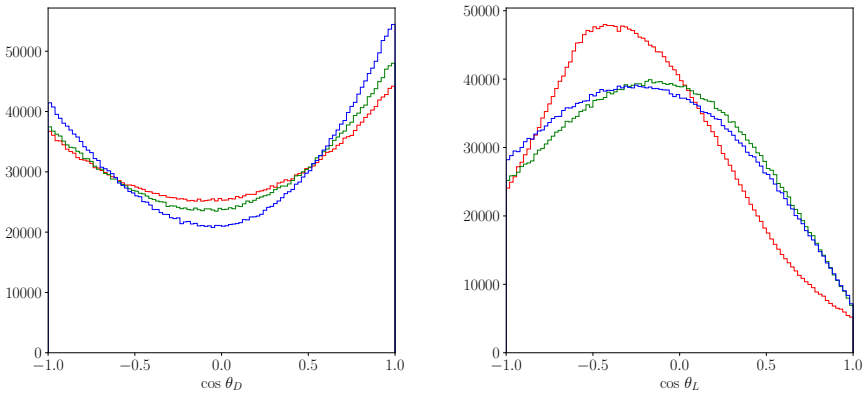


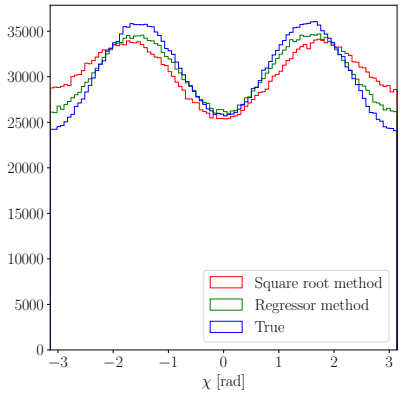
Figure 4.4: Predicted B^0 momentum for different models compared to the true B^0 momentum

4.2.2 Decay angles

With the estimate for the B^0 momentum, the three decay angles can be calculated. Using an MC sample of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays, both methods are used to calculate the angles, and the resulting distributions can be compared to the truth-level distributions, as shown in Figure 4.5. It is clear that for all three angles, the linear regression estimate provides a better approximation of P_{true} than the square-root method. By far the biggest improvement is seen in $\cos\theta_L$, which has a noticeably different shape for the two calculation methods. The improvement with the linear regression method can be seen more clearly in Figure 4.6. These plots show the resolutions for each angle, where the resolution is calculated by subtracting the estimated value from the true value for each event. The widths of the resolution functions are shown, and all the angles have an improved resolution with the regression method; the $\cos\theta_L$ resolution is more than twice as good. Figure 4.7 shows two dimensional histograms of the decay angle resolutions, calculated using the regressor method.

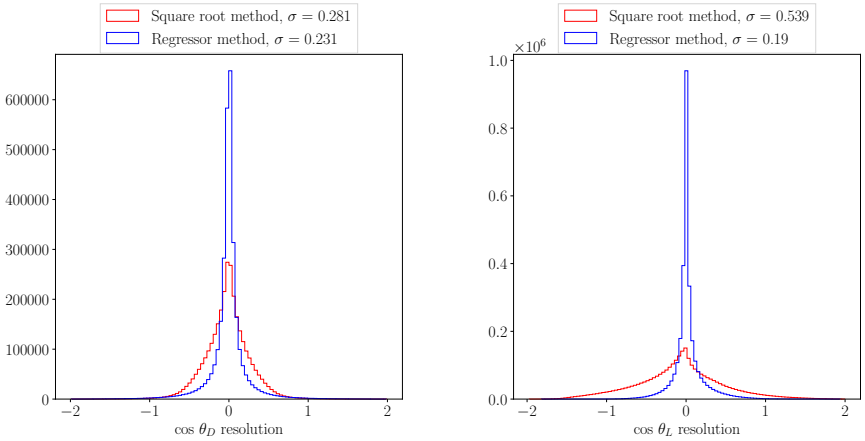


(a) $\cos\theta_D$ **(b)** $\cos\theta_L$

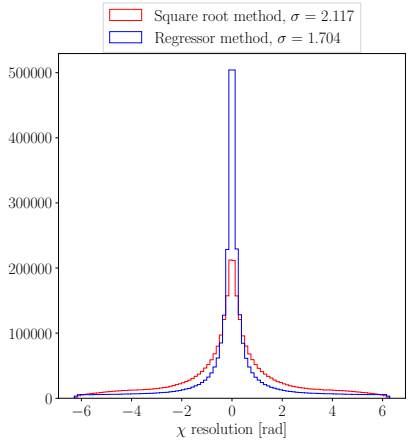


(c) χ

Figure 4.5: Plots showing the decay angle calculation, comparing the two methods described with the true angular distributions



(a) $\cos\theta_D$ **(b)** $\cos\theta_L$



(c) χ

Figure 4.6: Angular resolutions for the two calculation methods; the resolution is defined in the text

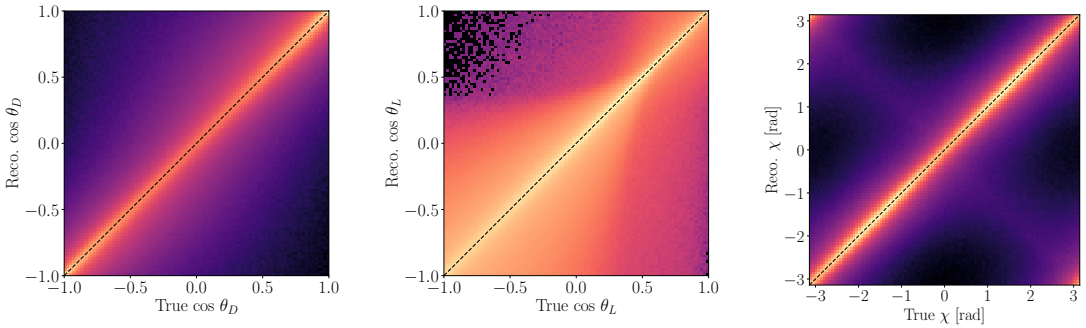


Figure 4.7: 2D histograms showing the distributions of reconstructed angular variables versus their true values in $B^0 \to D^{*-}\mu^+\nu_\mu$ MC. Lighter colours indicate a higher density of events

4.3 Angular fit method

The measurement presented in this thesis employs a novel binned multidimensional template fitting procedure to extract the angular observables. The key goal of the method is to create templates from simulation that have fully general shapes, and are not dependent on the particular decay model used to generate the simulation. In this way, the resulting I_X coefficients are decoupled from any form factor scheme and thus can be interpreted in a multitude of ways.

The fit uses the reconstructed variables $(\cos\theta_D, \cos\theta_L, \chi)$ as a three-dimensional basis, and the 12 angular functions in Eq. (3.46) are each represented by three-dimensional histograms. Each template histogram is normalised, multiplied by its corresponding I_X coefficient, and summed to create the *probability density function* (PDF). The I_X coefficients effectively act as scaling factors that define the contribution of each histogram to the total PDF. The fit then varies the contribution from each of the templates, thus measuring the 12 I_X coefficients.

Histogram templates created from the pure angular functions will not model the angular shapes in the data accurately, since the data is sculpted by reconstruction biases, especially from the missing neutrino, as well as other detector resolution effects. In contrast, a template fit in the reconstructed angular space accounts for these effects, by ensuring they are included directly in the PDF.

Template fits require that the MC reliably describes the event migration caused by reconstruction and acceptance effects, so that the fit can accurately map between the "true" and "reconstructed" distributions. These effects can be studied by creating acceptance functions in the fit variables, by making ratios of the "true" and "reconstructed" distributions in the MC. The acceptance functions show the form of the event migration derived from the MC. However, for this decay mode, the shape of the acceptance functions depends strongly on the decay model used to generate the MC.

This can be seen in Fig. 4.8, which shows the acceptance functions derived from MC with different decay models. To create these samples, in the generation, the form factors describing the decay were varied from their default values by different

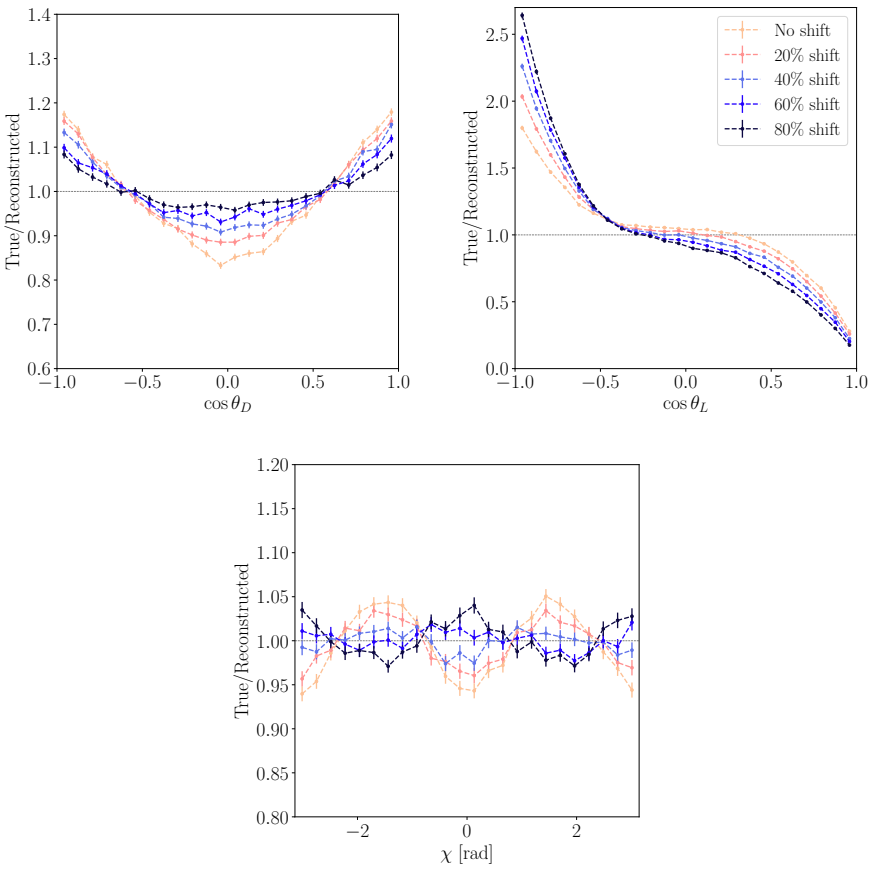


Figure 4.8: Acceptance functions from MC generated with different EvtGen form factors. The acceptance function is defined as the ratio of the truth-level and reconstructed value for each variable

amounts. The plots show that the shape of the acceptance functions vary between the different models, and that the bigger the change to the form factors, the bigger the shift in the shape of the acceptance function. The effects are seen in Fig. 4.8 because the "reconstructed" variables in MC are biased in two separate ways:

- Reconstruction and detector acceptance effects
- Biases from the choice of model used to generate the MC

The combination of these two effects mean that the "reconstructed" angular variables cannot be taken as a reliable model for the reconstruction and detector acceptance effects. This is because if these variables are used to create the fit templates without any corrections, the fit results will be model-dependent, as templates taken from

MC generated with a different decay model would have different shapes. Clearly, templates created from MC generated with a particular decay model will not accurately model the event migration in MC generated with a different model. To avoid this, when creating the 12 template histograms, per-event weights are assigned, in order to cancel the model-dependence in the templates, leaving a set of general shapes which can successfully describe the reconstruction and detector acceptance effects. In this analysis, we therefore assume that the MC, after all corrections are applied, can fully model all reconstruction and acceptance effects present in the data.

The weights are calculated using the following equation:

$$w_{I_X} = \frac{f_X}{\sum\limits_x I_x^{true} f_x} \tag{4.13}$$

where $X \in \{1c, 1s, 2c, 2s, 3, 4, 5, 6c, 6s, 7, 8, 9\}$, and f_X are the angular functions that are multiplied by the I_X coefficients in Eq. (3.46). Explicitly, to calculate the weight for a particular event in a particular template, the angular function associated with that template is divided by the total model function. The total model here is defined as the normalised decay rate shown in Eq. (3.46). As an example, the weight for an event in the histogram associated with the I_{6c} term is given by:

$$w_{I_{6c}} = \frac{\cos^2\theta_D^{true}\cos\theta_L^{true}}{\sum\limits_x I_x^{true} f_x}. \tag{4.14}$$

The I_X^{true} coefficients are taken from a parametric fit to a generator level sample of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays, which will be described in section 5.5.2. Since this truth-level sample uses the same model as the LHCb MC, the denominator of Eq. (4.13) represents the total model used to generate the signal MC. In this way, a set of 12 weights (w_{I_X}) can be calculated for each MC event by using its true values for the angles and applying the equivalent of Eq. (4.13).

Note that the weight calculation in Eq. (4.13) differs from the method shown in Ref. [89], which binned the angular functions before calculating the weights. It was found that the binning led to small biases when large samples were required, as

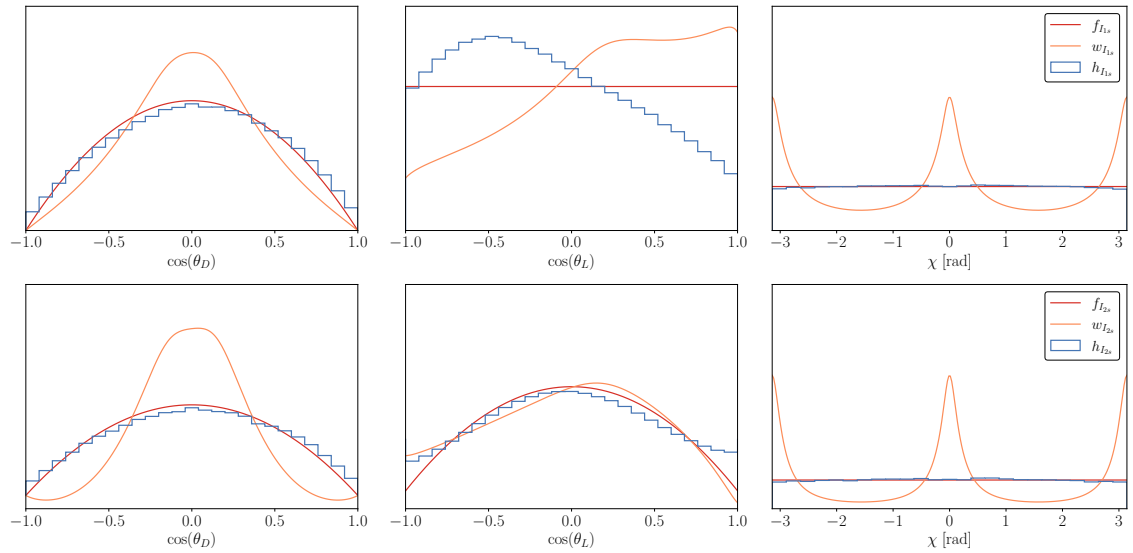


Figure 4.9: Each plot shows the shape of the angular function (red), the angular per-event weights (orange), and 1d template projection (blue) for the terms associated with I_{1s} (top) and I_{2s} (bottom)

the function variation across a bin was too large to be neglected. For this reason, continuous functions were used instead, which yields greater accuracy in the fit.

Next, histogram templates (h_{I_X}) are created, where each event is multiplied by the appropriate weight. In this way, a signal PDF can be created using the reconstructed angular variables, so any biases arising from the reconstruction are included, assuming the data is well modelled by simulation. The angular shapes of two of the 12 templates are shown in Fig. 4.9. Figure 4.10 shows the truth-level model used in the denominator of the weight calculation.

The per-event weights encode the angular shapes of each of the 12 templates. Since a division by the total model function (determined by the truth-level parametric fit) is performed in the weight calculation, the model dependence of the histogram templates is removed. In other words, the histogram templates can be created using MC generated with an arbitrary decay model, as long as the corresponding generator-level sample used to calculate the denominator of the w_{I_X} terms uses the same model. There is no requirement to select a realistic form factor scheme, as this method creates general fit templates, which are statistically identical regardless of the model used in the MC sample initially. This is shown in Section 5.6.1.

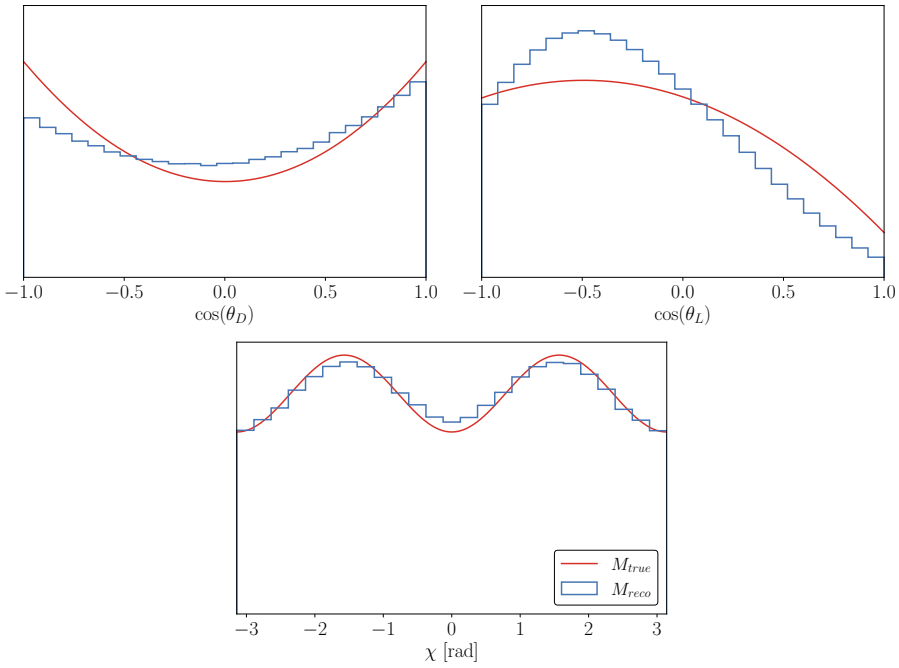


Figure 4.10: One dimensional projections for the total truth-level model (red). This was created by plotting Eq. (3.46) with the truth-level I_X values measured in the parametric fit. Also shown is the sum of the 12 signal histograms scaled by their truth-level I_X value (blue)

Note that the template weights are necessary for two main purposes:

- Encoding the angular shape for each of the 12 templates
- Removing the model dependence from the templates

For that reason, this method is more powerful than simply using acceptance functions taken from MC to correct the templates. This can be demonstrated by creating templates using acceptance functions instead of the template weights described above. The acceptance functions can be created by making ratios of true and reconstructed angles in MC. These functions can successfully correct reconstruction biases in the MC that was used to create the function themselves. However, when applying the acceptance functions to MC generated with an alternative decay model, the resulting template shapes will not model the simulation well. This is because, as shown in Fig. 4.8, the acceptance functions are model dependent. Therefore, the weight procedure first described in Ref. [89] is preferred, to preserve the model independence of the angular fit templates.

In this method, all reconstruction and detector acceptance effects are encoded directly in the PDF, in a model-independent way. The measured I_X coefficients are equivalent to the theoretical q^2-integrated I_X, so there is no requirement to perform any post-fit "unfolding".

The total signal PDF is:

$$PDF = \frac{1}{3}(4 - 6I_{1s} + I_{2c} + 2I_{2s})\ h_{I_{1c}} + \sum_X I_X\ h_{I_X} \tag{4.15}$$

where $X \in \{1s, 2c, 2s, 3, 4, 5, 6c, 6s, 7, 8, 9\}$. This is equivalent to Eq. (3.46) with the angular functions replaced by h_{I_X}, and the value of I_{1c} constrained by a normalisation condition:

$$\frac{1}{4}(3I_{1c} + 6I_{1s} - I_{2c} - 2I_{2s}) = 1. \tag{4.16}$$

This can be derived by integrating Eq. (3.46) over the three angles, to obtain the total decay rate, and requiring that to be equal to 1, as is conventionally done in angular analyses.

Figure 4.11 shows the angular function for each of the 12 terms. Additionally, it shows the 12 templates summed in equal proportions, and the templates summed according to the theoretical I_X (according to Tab. 3.1). This demonstrates how the overall angular shapes in the MC can be recovered by summing the 12 templates in a particular proportion.

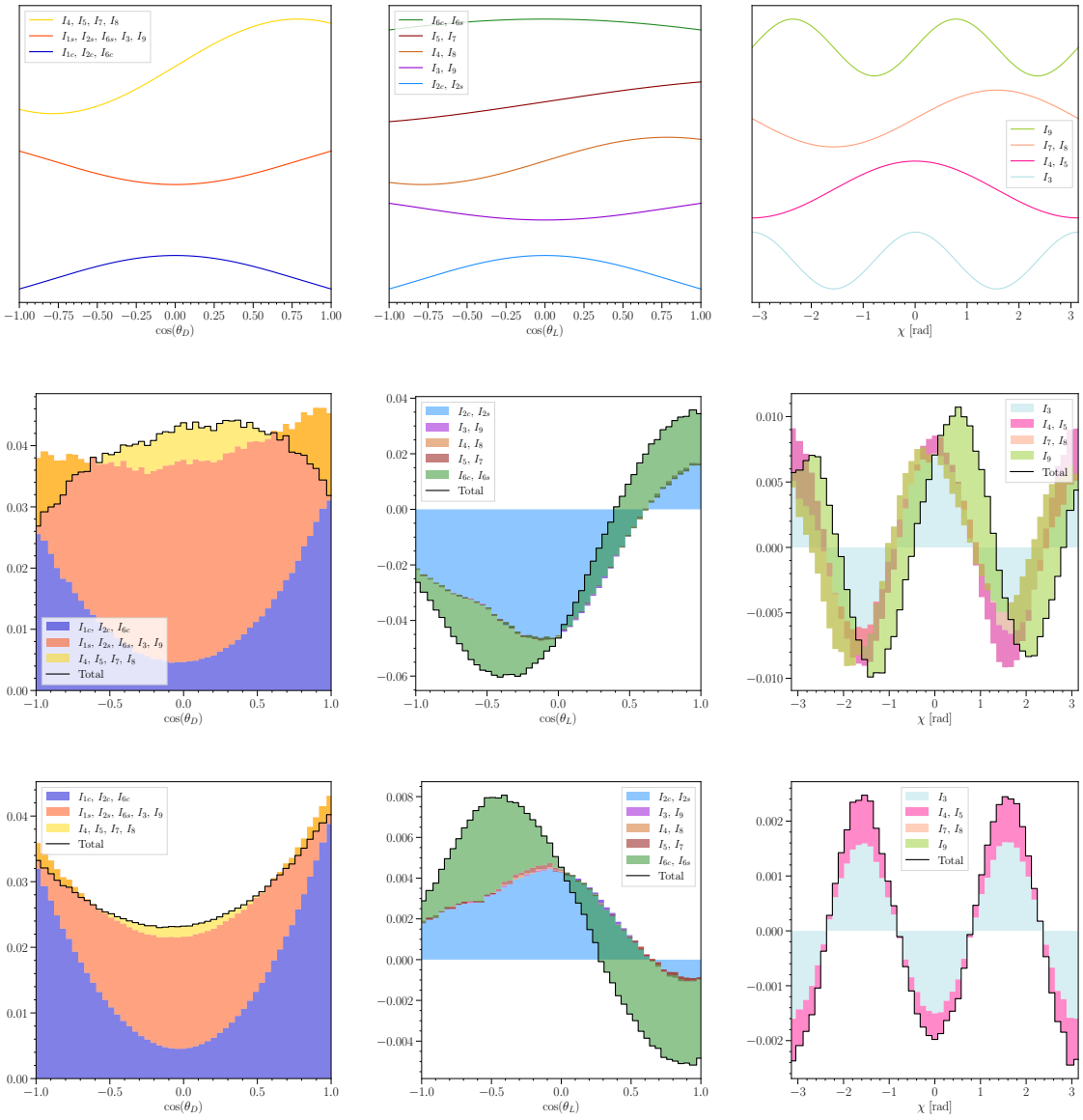


Figure 4.11: Top: One-dimensional angular functions associated with each angular term. The terms with the same function for a particular angle have been grouped. Middle: One-dimensional projections showing the 12 templates summed in equal proportions. Bottom: One-dimensional projections showing the 12 templates summed using the predicted I_X values.

4.3.1 Extending to four dimensions

The signal PDF can be modelled perfectly in three dimensions. However, when fitting to a data sample including backgrounds, it is useful to include an additional variable to improve background discrimination. It would be difficult to reliably fit the data sample in only three dimensions, since some of the background modes have similar

angular distributions to the signal. Therefore, using an additional 4th fit variable (m_{corr}) helps to measure the background fractions accurately. m_{corr} is defined as [91]:

$$m_{corr} \equiv |p_\perp| + \sqrt{|p_\perp|^2 + m^2_{D^*\mu}} \,, \tag{4.17}$$

where $|p_\perp|$ is the component of the reconstructed momentum of the $D^*\mu$ system that is transverse to the B^0 flight direction, and $m_{D^*\mu}$ is the mass of the $D^*\mu$ system.

Importantly, the signal template is made in the same way in the four-dimensional fit as in the three-dimensional fit described earlier. In particular, the m_{corr} shape is also described by a sum of 12 templates. This means that as the I_X coefficient values are varied in the fit, the m_{corr} signal shape is adjusted along with the angular shapes. Therefore, the signal fit templates are still model independent in four dimensions.

The effect of varying the I_X coefficients on the model shapes is shown in Fig. 4.12. The shape with the default truth-level I_X values is plotted, along with a shaded band representing the shape variation when the I_X values are randomly fluctuated by 10%.

4.3.2 Adding backgrounds

The background components are included in the PDF by forming a four-dimensional background histogram, h_{bkg}, for each component and multiplying it by the relative fraction for the component in the data, f_{bkg}. The signal PDF in Eq. (4.15) is multiplied by a signal fraction, f_{sig}, and the background terms are added. These background fractions can be floating or fixed in the angular fit, and one fraction is constrained so that all fractions sum to unity. The full PDF including backgrounds is:

$$PDF = \left(\frac{1}{3}(4 - 6I_{1s} + I_{2c} + 2I_{2s})\; h_{I_{1c}} + \sum_X I_X \; h_{I_X}\right) f_{sig} + \sum_{bkg} f_{bkg} \; h_{bkg} \,, \tag{4.18}$$

where the h_x terms now represent four-dimensional templates for both the signal and background. For each data-taking year, separate templates are created for the signal and background modes. Then these are added together, and scaled by the appropriate luminosity ratios, shown in Tab. 4.2.

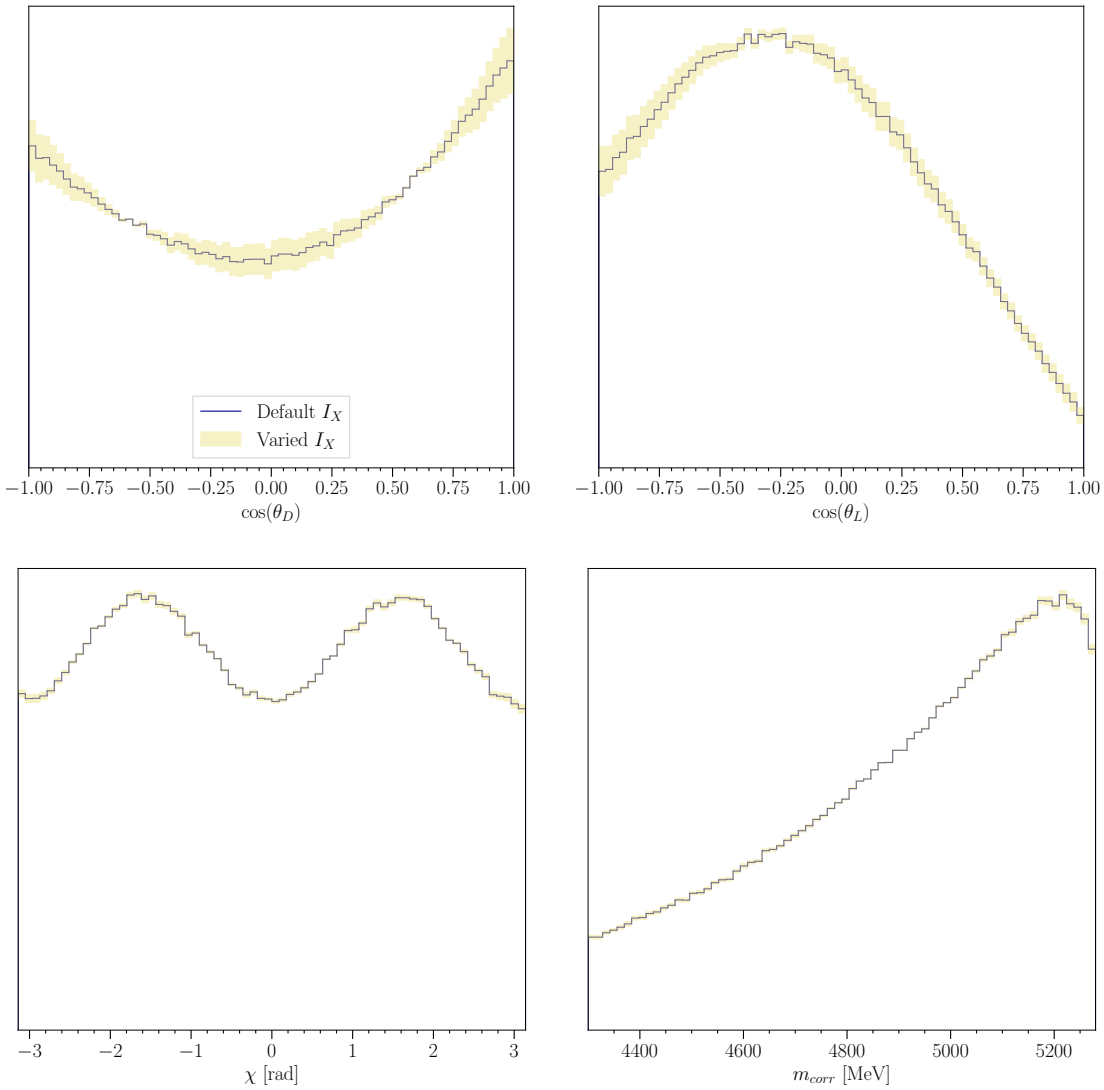


Figure 4.12: Effect of varying the underlying I_X values on the shapes of the fit variables. The blue line represents the shape with the default (truth-level) I_X. The yellow band is the $\pm 1\sigma$ region for random variations of the I_X within 10% of their default values.

Condition	2016	2017	2018	Total
MagDown	0.824	0.721	0.964	2.509
MagUp	0.787	0.735	1.043	2.566
Both	1.612	1.456	2.007	5.075

Table 4.2: Integrated luminosity split by year of data taking and magnet polarity, in fb^{-1}

4.3.3 Fit binning and parameters

Table 4.3 shows a list of the fixed and floating components in the angular fit. The

Parameter	Status
I_{1s}	floating
I_{2c}	floating
I_{2s}	floating
I_{6c}	floating
I_{6s}	floating
I_3	floating
I_4	floating
I_5	floating
I_7	floating
I_8	floating
I_9	floating
I_{1c}	fixed by normalisation
Signal fraction	floating
$B \to (D^*\pi)\mu\nu$ fraction	floating
$B \to (D^*\pi\pi)\mu\nu$ fraction	floating
$B_s \to (D^{**})\mu\nu$ fraction	floating
$B \to D^*\tau\nu$ fraction	floating
Combinatorial fraction	fixed
$B \to D^*DX$ fraction	fixed by normalisation

Table 4.3: Parameters in 4d angular fit

binning scheme is shown in Table 4.4. The bins have variable widths, which were chosen to ensure approximately equal signal yield in each bin. This avoided having regions of the fit with a very low density, which led to increased fit stability. The fit uses a binned, negative log-likelihood minimisation.

	$\cos(\theta_D)$	$\cos(\theta_L)$	χ [rad]	m_{corr} [MeV]
Bin 1	[-1.0, -0.843]	[-1.0, -0.902]	[-3.142, -2.555]	[4500.006, 4776.203]
Bin 2	[-0.843, -0.674]	[-0.902, -0.813]	[-2.555, -2.034]	[4776.203, 4859.147]
Bin 3	[-0.674, -0.495]	[-0.813, -0.729]	[-2.034, -1.559]	[4859.147, 4925.13]
Bin 4	[-0.495, -0.308]	[-0.729, -0.646]	[-1.559, -1.086]	[4925.13, 4980.899]
Bin 5	[-0.308, -0.118]	[-0.646, -0.561]	[-1.086, -0.573]	[4980.899, 5029.563]
Bin 6	[-0.118, 0.07]	[-0.561, -0.471]	[-0.573, -0.003]	[5029.563, 5072.929]
Bin 7	[0.07, 0.251]	[-0.471, -0.373]	[-0.003, 0.57]	[5072.929, 5112.358]
Bin 8	[0.251, 0.424]	[-0.373, -0.26]	[0.57, 1.085]	[5112.358, 5148.676]
Bin 9	[0.424, 0.584]	[-0.26, -0.124]	[1.085, 1.558]	[5148.676, 5182.552]
Bin 10	[0.584, 0.734]	[-0.124, 0.051]	[1.558, 2.034]	[5182.552, 5214.864]
Bin 11	[0.734, 0.872]	[0.051, 0.301]	[2.034, 2.555]	[5214.864, 5246.562]
Bin 12	[0.872, 1.0]	[0.301, 1.0]	[2.555, 3.142]	[5246.562, 5279.63]

Table 4.4: Binning scheme in 4d angular fit. Each variable has 12 bins so there are $12^4 = 20,736$ bins in total.

4.3.4 Polynomial corrections

To correct residual data/MC discrepancies in $\cos\theta_L$ and χ, a procedure is used to adjust the template shapes within the angular fit itself. This is performed by firstly creating two one-dimensional polynomial functions, one for $\cos\theta_L$ and one for χ. These functions are created by summing a set of Chebyshev polynomials, where the fraction of each polynomial is a floating parameter. For example:

$$f(\chi) = z_0 T_0(\chi) + z_1 T_1(\chi) + z_2 T_2(\chi) + z_3 T_3(\chi) + z_4 T_4(\chi) \tag{4.19}$$

$$= (z_0 - z_2 + z_4) + (z_1 - 3z_3)\chi + (2z_2 - 8z_4)\chi^2 + 4z_3\chi^3 + 8z_4\chi^4 \,, \tag{4.20}$$

where z_i are the floating polynomial parameters. An equivalent function is defined for $\cos\theta_L$, where the corresponding polynomial parameters are labelled y_i.

Then, the polynomial functions are binned using the same binning scheme as the fit templates. Two uniform binned functions are also created for $\cos\theta_D$ and m_{corr}. The outer product of the four functions is calculated, yielding a single four-dimensional binned correction histogram. All of the four-dimensional fit templates that are derived from simulation (all apart from the combinatorial

template) are multiplied by the correction histogram, which results in a set of corrected 4d templates.

By allowing the polynomial parameters to float in the fit, the shape of the correction histogram is adjusted at each iteration of the fit, at the same time as the other parameters are fluctuated. Also, the binned polynomial functions are normalised at each iteration, such that their average value is equal to 1. This fixes the value of the 0th order term in the polynomial function (z_0 in eq. (4.20)). This ensures that multiplying the fit templates by the correction histogram does not affect the component yields. Since each of the 12 signal templates is multiplied by the same correction histogram, there is no resulting bias of the I_X.

Two alternative parameterisations for the polynomial correction functions were tested. One method was to define a generic 5th order polynomial and vary the coefficients directly. Another way was to define a polynomial by summing powers of sine and cosine functions. Both of these methods yielded polynomial shapes that were consistent with those obtained from the Chebyshev method.

5 Angular analysis of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays

5.1 Analysis strategy

An angular analysis of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays[1] was performed using the technique described in Chapter 4. The LHCb data sample used corresponds to 5.1 fb^{-1} and was taken between 2016 and 2018. D^0 meson candidates were selected and combined with a "slow" pion to make a D^* meson. The D^* was then combined with a muon of opposite charge to create a B^0 meson.

Selection cuts were placed to reduce contributions from background processes. These are outlined in Section 5.2. Furthermore, two dedicated boosted decision tree (BDT) classifiers were trained to remove specific backgrounds, as described in Section 5.3.3. The angular fitting method requires the simulation-derived templates to fully describe the biases from detector and selection effects on the data. For this reason, correction weights for the simulation were carefully calculated in order to account for small discrepancies between the simulation and data. The correction procedure is shown in Section 5.4.

[1] Charge conjugation is implied, so the equivalent process was carried out for the $\overline{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu$ decays

It is necessary to obtain the truth-level I_X values that describe the simulation, as they are required to cancel out the model dependence when constructing the signal templates. These values are calculated from a parametric fit to a generator-level simulation sample, described in Section 5.5.2. The obtained values are then validated using the Method of Moments, shown in Section 5.5.4.

As the method used in this work is novel, the central claim of model independence must be validated. Studies to this effect are presented in Section 5.6.

The four-dimensional angular fit is then performed on the data sample to measure the I_X coefficients. The fit was first applied to pseudodata which verified that the procedure was unbiased. All results are shown in Section 5.7. The main sources of systematic uncertainty were evaluated, and are summarised in Section 5.8. The analysis was performed blind, i.e. the values of the I_X coefficients from the data fit were hidden until all cross-checks were completed.

5.2 Datasets and selection

Pre-selection cuts[2] were applied to the data and simulation to increase signal purity and to reduce the size of the samples. The cuts are listed in Tab. 5.1. For the final-state particles, a minimum threshold was placed on the χ^2 of the impact parameter with respect to its origin vertex ($IP\chi^2$), as well as the momentum (p) and momentum transverse to the beamline (p_T). The composite particles have a requirement on the χ^2 of the fit used to calculate the vertex position (Vertex χ^2_{fit}). For the B and D^0 mesons, the direction angle (DIRA) was cut on. This is the angle between the reconstructed momentum of the particle and a line drawn between the decay vertex and the PV. The D^0 meson is reconstructed using the decay $D^0 \to K^-\pi^+$.

At the hardware trigger level, candidates are required to have a hadron with high transverse energy in the calorimeters. Alternatively, one of the other particles in the event, not reconstructed as part of the signal candidate, must fulfil any trigger requirement. Following this, further offline selection requirements were added, which are summarised in Tab. 5.2. ΔM is defined as $m(D^*) - m(D^0)$. m_{corr}

[2]These cuts are referred to as "stripping selection" at LHCb

Event-Level Cuts	**Variable**	**Selection**
Global Event Cut	nSPDhits	< 600
PV cut	nPV	> 0
Particle Cuts	**Variable**	**Selection**
μ	isMuon	True
	$IP\chi^2$	> 16
	Ghost track probability	< 0.5
	PIDμ	> -200
	p	> 3 GeV/c
K, π	p_T	> 300 MeV/c
	p	> 2 GeV/c
	$IP\chi^2$	> 9
	isMuon	False
	$DLLK$	$> 4(K), < 2(\pi)$
	Ghost track probability	< 0.5
D^0	$\lvert p_T(\pi)\rvert + \lvert p_T(K)\rvert$	> 2.5 GeV/c
	Vertex χ^2_{fit}/dof	< 4
	$DIRA$	> 0.999
	$FD\chi^2$	> 25
	$m - \langle m \rangle$	< 80 MeV/c^2
D^*	$\lvert p_T(D^0)\rvert + \lvert p_T(\pi_s)\rvert$	> 1.4 GeV/c
	$m - \langle m \rangle$	< 85 MeV/c^2
	Vertex χ^2_{fit}/dof	< 8
B^0	m	$\in [2.2$ GeV/c^2, 10 GeV/$c^2]$
	Vertex χ^2_{fit}/dof	< 6
	$DIRA$	> 0.999

Table 5.1: Pre-selection cuts applied to data and simulation. $\langle m \rangle$ refers to the PDG value for the particle's mass. $FD\chi^2$ is the χ^2 distance of the particle decay vertex from the related PV.

Variable	Cut
$m(D^*)$	$m(D^*)_{PDG} \pm 30$ MeV
$m(D^0)$	$m(D^0)_{PDG} \pm 30$ MeV
ΔM	143 - 148 MeV
m_{corr}	4000 - 5350 MeV
D^* vertex χ^2	< 7.5
π_s track χ^2/dof	< 2
τ BDT	> 0.898
iso BDT	> 0.566
D^0 FD z_{sig}	> 2
μ ProbNNmu	> 0.4
$p_T(\mu)$	> 1500 MeV

Table 5.2: Further offline selection cuts applied to data and simulation

is defined in Eq. (4.17). The τ BDT and iso (isolation) BDTs will be outlined in Sec. 5.3.3. The D^0 flight significance is defined as:

$$D^0 \text{ FD } z_{sig} = \frac{D^0 \text{ vtx}(z) - B^0 \text{ vtx}(z)}{\sqrt{D^0 \text{ vtx}(z_{err})^2 + B^0 \text{ vtx}(z_{err})^2}} , \tag{5.1}$$

where vtx(z) is the z coordinate of the particle's decay vertex, and vtx(z_{err}) is the error on that measurement. ProbNNmu is a variable that is created using machine learning techniques, combining PID and tracking information to estimate the probability that a given particle is a true muon. Plots of the variables used for this part of the selection are shown in Fig. 5.1.

5.2.1 Simulation

LHCb simulation is used to create the templates which model the data in the angular fit. Since the data samples are relatively large, corresponding large simulation samples are required to describe the data reliably and to allow fine-grained binning.

Simulating the full LHCb detector is computationally expensive and the resulting samples require considerable storage space [92]. As a result, generating large simulation samples efficiently is done using *tracker-only* simulation. With this method, optical photons are not generated in the RICH system and the calorimeter

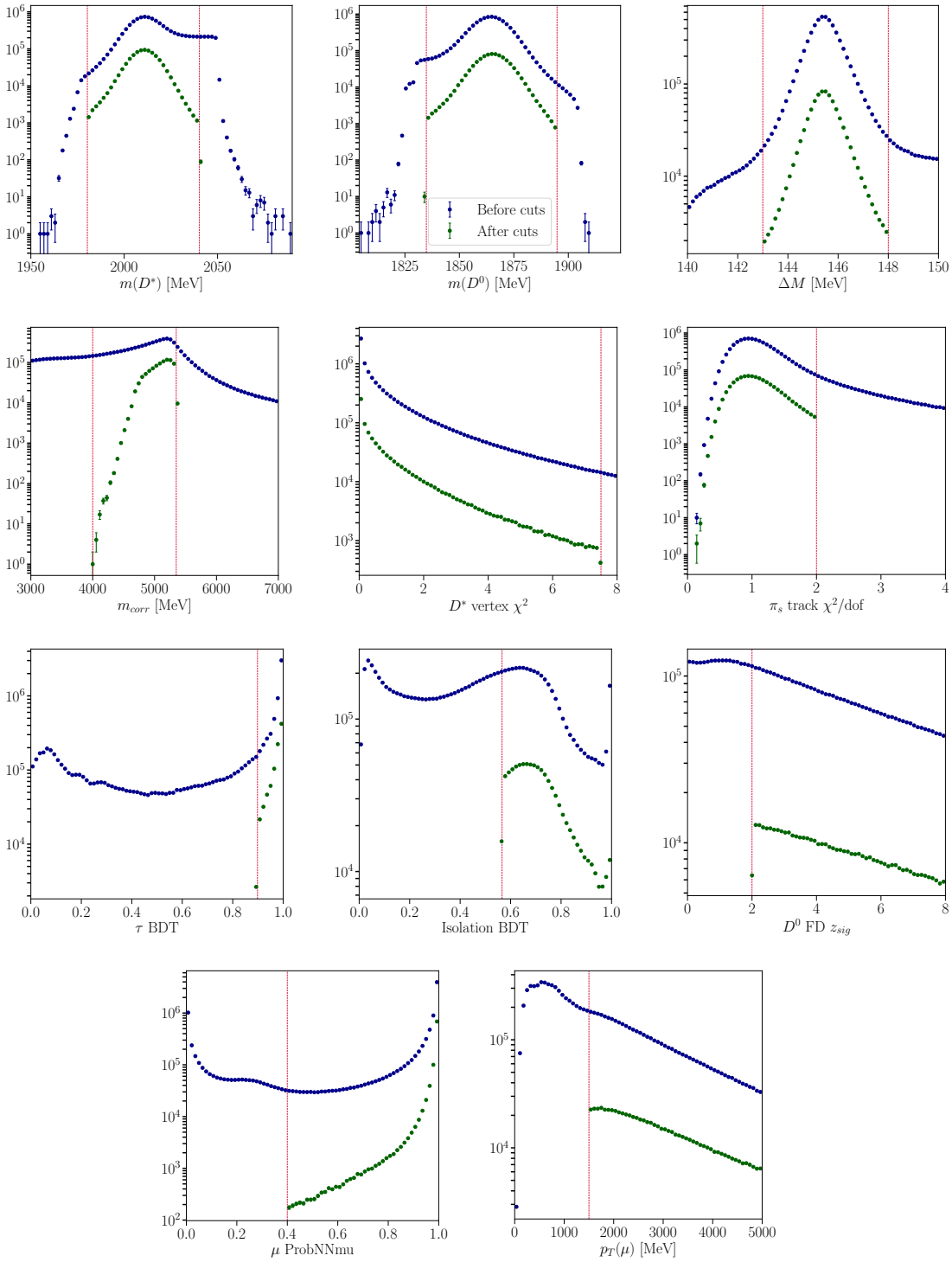


Figure 5.1: Distributions of variables used for the additional offline selection, plotted in data. The blue points show the sample before the additional offline selection, and the sample after all offline selection is applied is plotted in green. The dashed red lines show the values of the cuts.

and muon chambers are removed entirely. Consequently, the tracker-only simulation is around 10 times faster and uses 25% less disk space [93]. However, this means that the resulting samples contain less information than the full simulation. In particular, the trigger system cannot be emulated so this effect must subsequently be accounted for, as described in Sec. 5.4.1. This process involves comparing the large tracker-only MC samples to corresponding smaller validation samples generated with the full LHCb simulation.

5.3 Background description and reduction

Due to the high reconstruction efficiency of muons at LHCb, the dataset in this analysis has a high signal purity. The dominant background contributions considered are described below.

- $\boldsymbol{B^0 \to D^{*-}\tau^+\nu}$ **decays**, with $\tau \to \mu\nu\bar{\nu}$. These processes have a similar angular structure to the signal and are reduced by a BDT described in Section 5.3.3.

- $\boldsymbol{B \to D^{**}\mu\nu}$ **decays**. These are processes that involve higher excited D^* states (referred to as D^{**} throughout) that decay to the $D^{*-}(2010)$ state, which is present in the signal decay. The excited charm resonances modelled in this analysis are shown in Tab. 5.3. These include contributions from B^0, B^+, and B_s decays. This background is the most dominant background remaining after the full selection is applied, and it is reduced with an isolation BDT described in Section 5.3.3.

- $\boldsymbol{B \to D^*DX}$ **decays**. These are B^0 or B^+ decays to a charm hadron which then decays semileptonically to a muon. They are also reduced by the BDT described in Section 5.3.3.

- **Combinatorial background**. This refers to random combinations of a D^* and a muon which originate from different primary vertices. These processes

State	Mass (MeV)	Width (MeV)	Decay
$D_0^*(2300)^{0,\pm}$	2343 ± 10	229 ± 16	$D^{*+}\pi^-\pi^0$, $D^{*+}\pi^+\pi^-$, $D^{*+}\pi^0\pi^0$
$D_1(2420)^{0,\pm}$	2422.1 ± 0.8	31.3 ± 1.9	$D^{*+}\pi^-$, $D^{*+}\pi^0$
$D_1(2430)^{0,\pm}$	2412 ± 9	314 ± 29	$D^{*+}\pi^-$, $D^{*+}\pi^0$
$D_2^*(2460)^{0,\pm}$	2461.1 ± 0.7	47.3 ± 0.8	$D^{*+}\pi^-$, $D^{*+}\pi^0$, $D^{*+}\pi^+\pi^-$, $D^{*+}\pi^0\pi^0$
$D_1^*(2600)^{0,\pm}$	2627 ± 10	141 ± 23	$D^{*+}\pi^+\pi^-$ $D^{*+}\pi^0\pi^-$
$D^*(2640)^{0,\pm}$	2637 ± 6	< 15	$D^{*+}\pi^+\pi^-$ $D^{*+}\pi^0\pi^-$
$D_3^*(2750)^{0,\pm}$	2763.1 ± 3.2	66 ± 5	$D^{*+}\pi^+\pi^-$ $D^{*+}\pi^0\pi^-$
$D(3000)^{0,\pm}$	3214 ± 57	186 ± 81	$D^{*+}\pi^+\pi^-$ $D^{*+}\pi^0\pi^-$
$D_{s1}(2536)^+$	2535.21 ± 0.28	0.92 ± 0.05	$D^{*+}K_S^0$
$D_{s2}(2573)^+$	2569.1 ± 0.8	16.9 ± 0.7	$D^{*+}K_S^0$

Table 5.3: Excited charm states and their decay modes modelled in this analysis. The masses and widths displayed are taken from the PDG averages [88], apart from the $D(3000)$ which uses the measurement in Ref. [94].

are modelled by the "wrong-sign" sample that is created by selecting $D^{*+}\mu^+$ candidates in data. The rate of these processes is estimated in Section 5.3.2.

- **Fake D^0 and fake D^* background**. These are random $K\pi$ and $D^0\pi$ combinations respectively. This background is removed from the data by fitting $m(D^0)$ and ΔM and extracting sWeights, as described in Section 5.3.1.

- **Misidentified muon background**. This refers to $B \to D^*h$ decays, where h is a charged hadron that is misidentified as a muon. It is reduced by a PID cut on the muon and is small enough to be neglected in the final angular fit, although a systematic uncertainty is assigned for any residual background.

5.3.1 Fake D^0 and fake D^* background

One way of dealing with backgrounds from fake D^0 and fake D^* processes is to use a data-derived template, and include this in the angular fit, as was done in Ref. [95]. However, an alternative approach is preferred, using the `sPlot` technique [96] to statistically subtract these backgrounds from the data sample, thus removing the need for a dedicated template in the fit.

Using the `RooFit` package [97], a two-dimensional fit is performed to $m(D^0)$ and ΔM. The fit consists of three components, which are modelled as shown in Tab. 5.4. Some of the signal shape parameters were gaussian constrained by separate fits to simulation. The $B^0 \to D^{*-}\mu^+\nu_\mu$ MC sample was used for these fits. The signal and fake D^* components have the same shape in $m(D^0)$, and both background components have the same shape in ΔM.

From this fit, the sPlot technique allows a set of per-event sWeights to be calculated. The variation of the sWeights with the mass variables is shown in Fig. 5.2. These sWeights are then applied when making the four-dimensional data histograms for the angular fit, and the effect of this is to statistically subtract the fake D^0 and fake D^* backgrounds from the data. The fits were carried out independently for the 2016, 2017 and 2018 data samples, and separately for MagUp and MagDown. As an example, the fit projections for one sample are shown in Fig. 5.3.

Component	$m(D^0)$ shape	ΔM shape
Signal	2 sided Crystal Ball	2 sided Crystal Ball + Gaussian
Fake D^*	2 sided Crystal Ball	RooDstD0BG
Fake D^0	Exponential	RooDstD0BG

Table 5.4: Table showing the component shapes for the fit described in Section 5.3.1. The functions are defined in Appendix A.

Correlation between mass variables and angular fit variables

The `sPlot` technique requires that the discriminating variables used in the fit are uncorrelated with the control variable, which the sWeights are applied to. Accordingly, the correlations between $\{m(D^0), \Delta M\}$ and $\{\cos\theta_D, \cos\theta_L, \chi, m_{corr}\}$ were checked, and are summarised in Tab. 5.5. Additionally, the correlations are depicted in Fig. 5.4. There are no significant correlations, so the sWeights can be used in the template fit to reliably subtract the fake D^0 and fake D^* background.

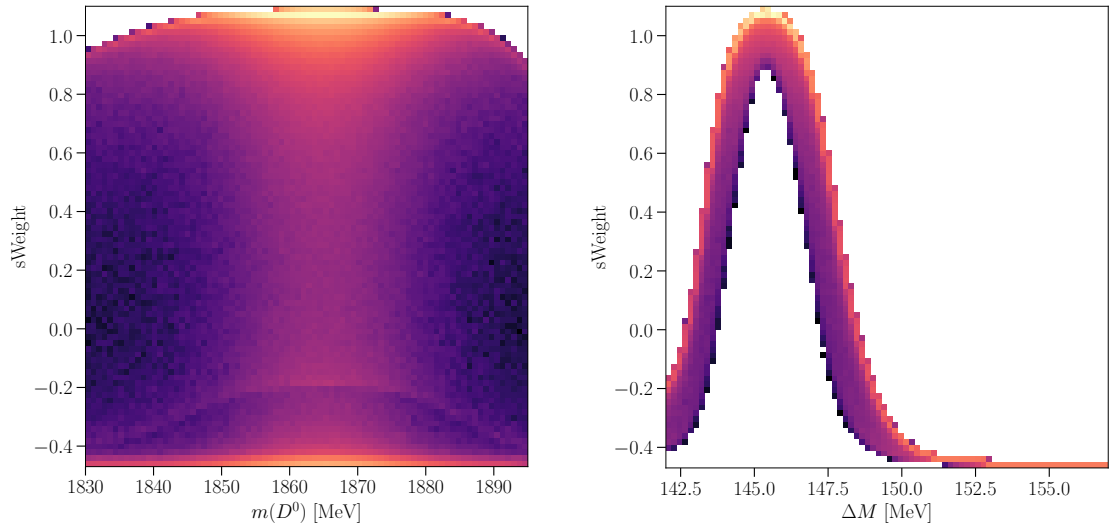


Figure 5.2: 2D histograms showing the variation of sWeights with the fit variables. Lighter colours indicate higher density of events.

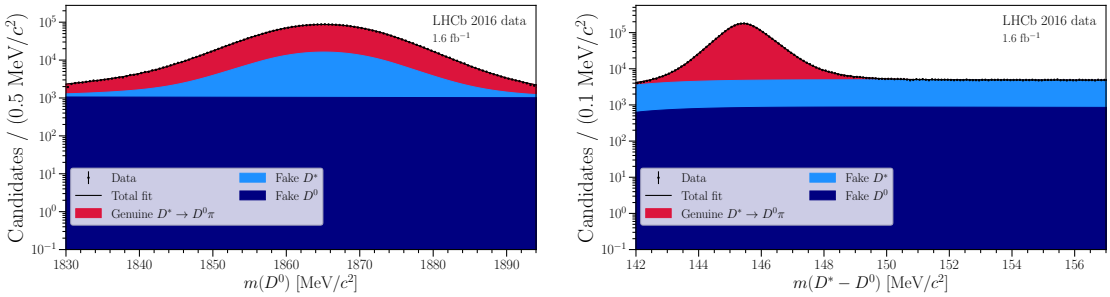


Figure 5.3: Fit projections for $m(D^0)$ (left) and ΔM (right), for the 2016 MagUp sample.

	$\cos\theta_D$	$\cos\theta_L$	χ	m_{corr}
$m(D^0)$	0.0059	-0.0033	0.0002	0.0131
ΔM	-0.0389	0.0055	0.0004	-0.0035

Table 5.5: Correlation coefficients between variables in mass fit and those in angular fit

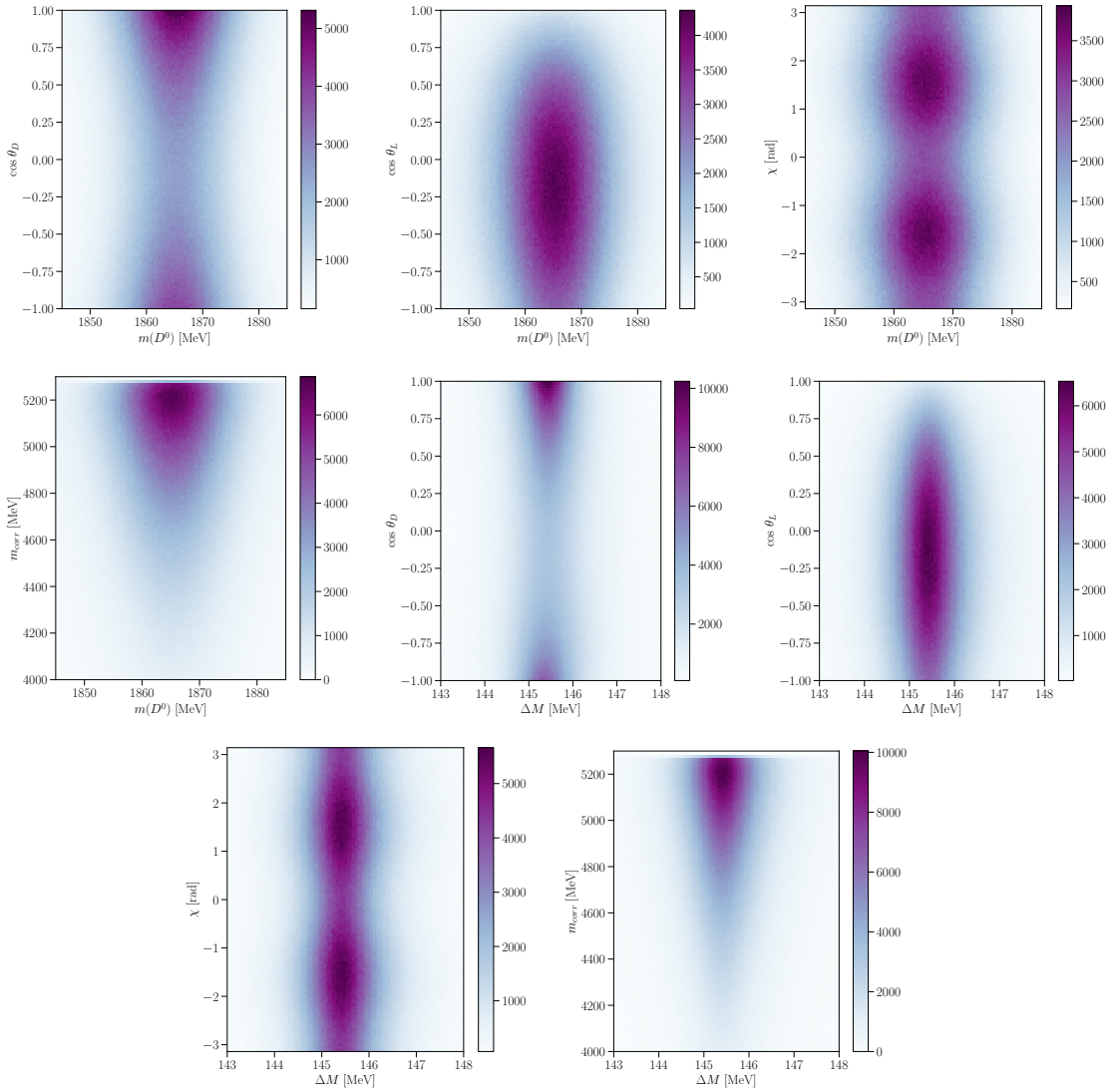


Figure 5.4: 2D histograms of mass variables and angular fit variables, showing they are uncorrelated

5.3.2 Combinatorial background

As mentioned previously, the combinatorial background is modelled by a sample of $D^{*+}\mu^+$ (and charge conjugate) candidates[3] selected in the same way as for the $D^{*+}\mu^-$ data sample. Applying the same selection to these events will give a sample that should accurately model the shape and yield of the combinatorial background in the data.

This assumption can be checked by comparing the wrong-sign sample to the data sample in a high region of m_{corr}, referred to as the sideband region. The region $m_{corr} \in [10000, 13000]$ MeV is chosen for this comparison as all other dominant backgrounds are negligible here. This can be seen by looking at m_{corr} in the Monte Carlo samples for the various background modes, as shown in Fig. 5.5. In this region, the shapes of the wrong-sign and right-sign data samples are consistent, which is shown in plots of the normalised yields in Fig. 5.6 and the ratios in Fig. 5.7. However, the yield of the wrong-sign sample was significantly higher than that of the right-sign sample. In other words, it appears as though random $D^{*+}\mu^+$ combinations are not selected at the same rate as random $D^{*-}\mu^+$ combinations. To account for this, the ratio between the wrong-sign and right-sign yields, averaged across the sideband region, is taken as a correction factor, and this is applied to the fixed combinatorial background yield in the angular fit. The value of the combinatorial background yield is calculated by multiplying the ratio of wrong-sign to right-sign events in the whole sample by the correction factor from the m_{corr} sideband.

This comparison was performed separately for each data year, so a different correction factor was applied to each of the 2016, 2017 and 2018 samples. The correction factor for each year, along with the average, is shown in Tab. 5.6. The consistency of the right-sign and wrong-sign shapes in this region of m_{corr} was checked by calculating a χ^2/dof; this is also shown in Tab. 5.6.

[3]These combinations are sometimes referred to as "same-sign" events but in this chapter the term "wrong-sign" is used instead.

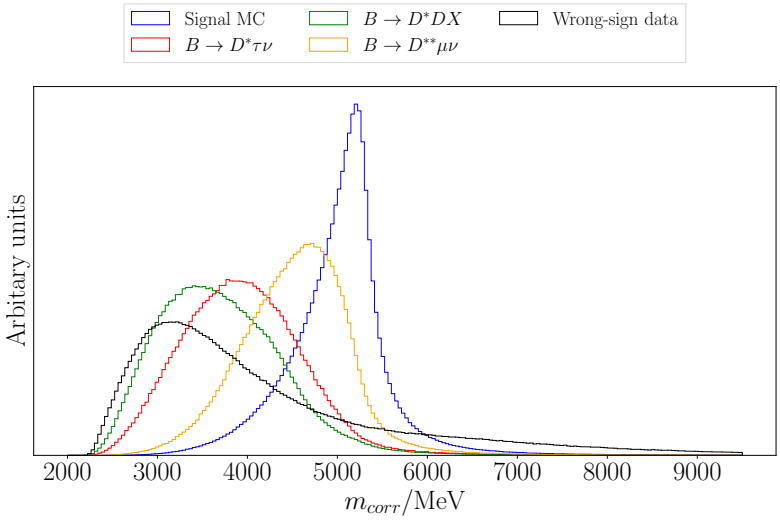
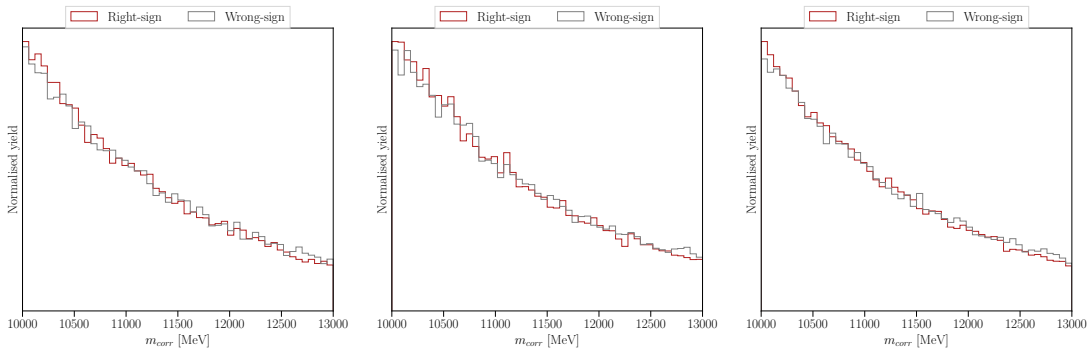


Figure 5.5: m_{corr} distribution for signal and background modes

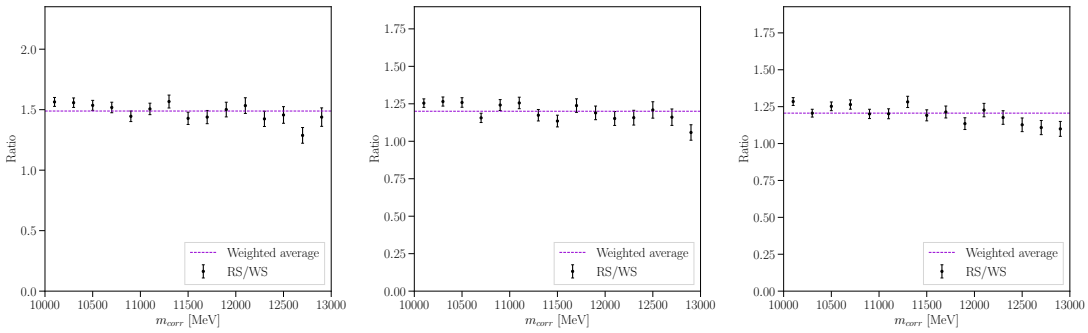
Data year	Combinatorial correction	χ^2/dof
2016	1.489 ± 0.072	1.09
2017	1.200 ± 0.059	1.16
2018	1.206 ± 0.058	1.13
Total	1.29 ± 0.036	1.13

Table 5.6: Table showing the correction factors applied to the combinatorial background templates for each data year. The χ^2/dof represents the deviation between the right-sign and wrong-sign shapes. The "Total" value shows the average correction across all years, weighted by luminosity



(a) 2016 data **(b)** 2017 data **(c)** 2018 data

Figure 5.6: Normalised yields of right-sign and wrong-sign data samples in the high m_{corr} sideband, showing that the two samples have consistent shapes across this region of m_{corr}



(a) 2016 data **(b)** 2017 data **(c)** 2018 data

Figure 5.7: Ratio of right-sign and wrong-sign yields in the high m_{corr} sideband

5.3.3 Multivariate classifiers

The use of machine learning techniques is well established in high energy physics as they are extremely efficient at solving classification problems involving multiple overlapping variables. The most popular class of algorithms are boosted decision trees (BDTs) [98], which can be applied to particle identification problems [99] and to aid background discrimination [100].

This work uses two separate BDTs to remove backgrounds from $B^0 \to D^{*-}\tau^+\nu$ and $B \to D^{**}\mu\nu$ decays; they are outlined below.

τ BDT

To reduce the contributions from $B^0 \to D^{*-}\tau^+\nu_\tau$, a boosted decision tree (BDT) classifier is trained with the `XGBoost` package [101]. Although only $B^0 \to D^{*-}\tau^+\nu_\tau$ MC was used as the background sample in the BDT training, the BDT is also very efficient at removing $B \to D^*DX$ events.

In both of these decays, the muon originates from a different vertex to the D^*, unlike in the signal mode. Therefore, when the D^* meson is reconstructed, the muon flight direction will not necessarily align with the D^* vertex. This is shown in Fig. 5.8. Consequently, the signal mode is expected to have a better B^0 and D^* vertex χ^2 than the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B \to D^*DX$ modes, so this is one of the variables used to train the BDT. The muon momentum and transverse momentum are also used as inputs to the BDT. This is because, since an additional decay is

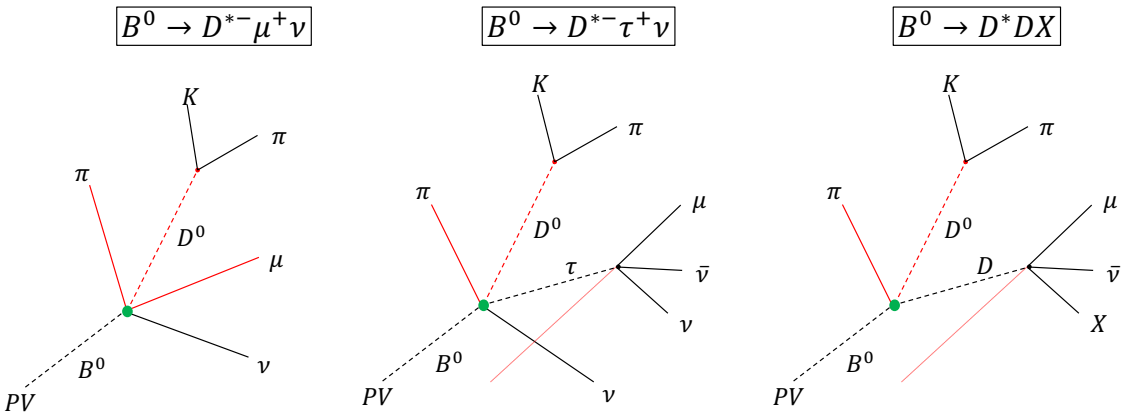


Figure 5.8: Diagram comparing the kinematics of the signal mode with the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B \to D^*DX$ background modes. The muons in the background modes do not point back to the D^* vertex (shown in green), so the vertex has a larger resolution. The D^{*-} meson decays very quickly to $D^0\pi^-$, so its flight vector is not shown.

required to produce the muon in these background modes compared to the signal mode, these variables are expected to have lower values on average in the background decays. The muon IP χ^2 and B^0 FD χ^2 were found to be shifted to lower values in the $B^0 \to D^{*-}\tau^+\nu_\tau$ mode compared to the signal mode. Finally, the corrected mass, m_{corr}, is used, as defined in Eq. (4.17). This variable gives a good separation for these backgrounds, as the τ and the D meson have a much higher mass than the muon. The distributions for all the training variables are shown in Fig. 5.10.

The BDT was trained using signal MC and $B^0 \to D^{*-}\tau^+\nu_\tau$ MC, and was then applied to all the MC samples[4] as well as the data. Sample weights were assigned to the signal and background samples to balance the class sizes. The performance of the BDT is shown in Fig. 5.9. The area under the Receiver Operating Characteristic (ROC) curve [102] is 0.91.

The distributions of the BDT variables in signal and background MC can be seen in Fig. 5.10. The BDT working point was chosen to be 0.898. With this cut applied, the signal efficiency is 70.0%, and the efficiencies on the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B \to D^*DX$ backgrounds are 9.4% and 4.7% respectively.

[4]The MC samples used to train the BDT were split in half to avoid applying the BDT to data that was used to train it.

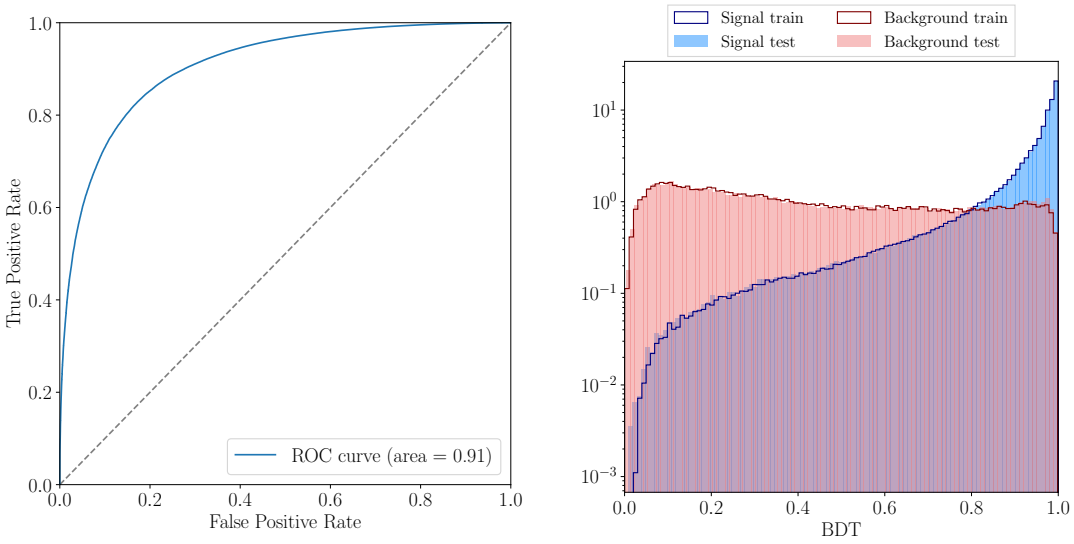


Figure 5.9: ROC curve and BDT distributions for signal and background

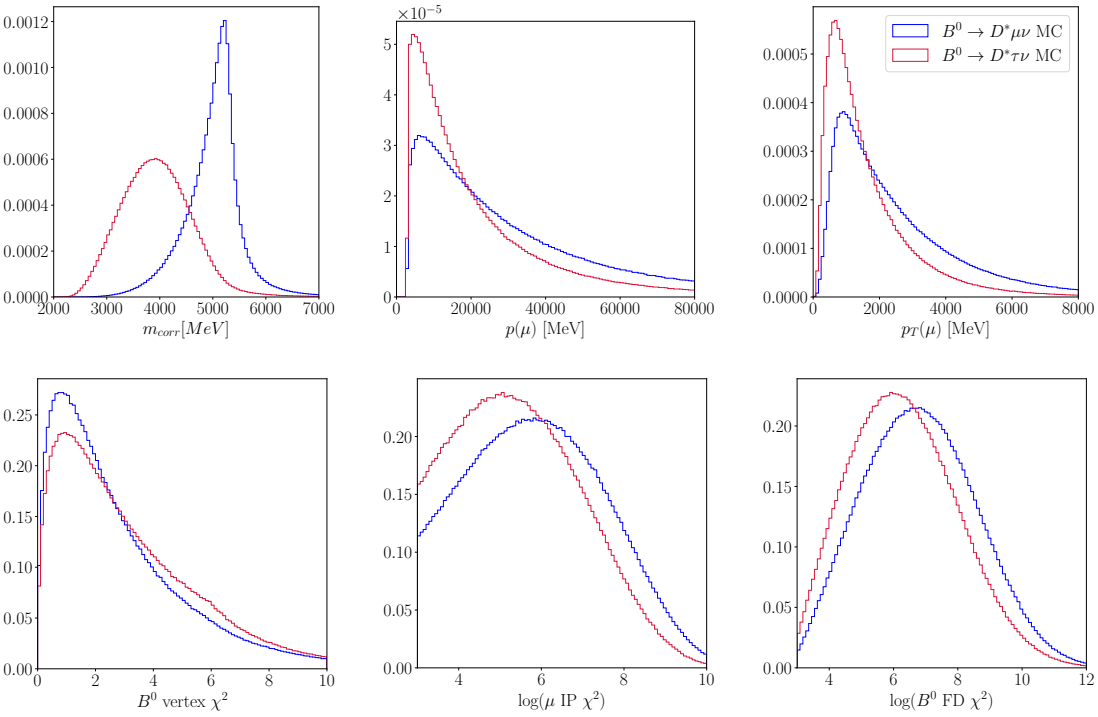


Figure 5.10: Signal (blue) and background (red) distributions of variables used in the BDT training

Isolation BDT

In addition, another BDT is used to control the rate of $B \to D^{**}\mu\nu$ processes, which form the most dominant background in this work. These backgrounds are very "signal-like" and thus are difficult to reduce. They also have some angular structure, which has not been well studied, meaning there is some uncertainty in the angular shapes present in the simulation. Therefore, it is extremely important to reliably measure the rate of these processes.

The input variables for the classifier are predominantly "isolation" variables. These tools are a way of isolating the signal decay from other tracks, which helps to separate the signal from background with partially reconstructed decays. Since the $B \to D^{**}\mu\nu$ decays will have one or two unreconstructed pions (or kaon from a B_s decay), these samples can be expected to have a different distribution to the signal in the isolation BDT variable.

Like the $B^0 \to D^*\tau\nu$ BDT, the isolation BDT was trained with `XGBoost`. Sample weights were used to balance the sizes of the signal and background samples. If the signal and background classes have unequal sizes, this can result in a model with insufficient information about the smaller class. This leads to higher levels of mis-classifcation for the smaller class. In this case, the BDT was trained with and without balanced sample sizes, and the balanced case gave a wider BDT distribution, with better separation between signal and background. The area under the ROC curve is 0.82, and the BDT performance in both cases can be seen in Figure 5.12. The BDT was checked for overtraining, and no significant issues were found, which can be seen by comparing the train and test samples in Figure 5.12.

The background training sample for this BDT was a combined MC sample, composed of the $B \to D^{**}\mu\nu$ decay modes used in the angular fit, in their expected proportions. The resulting BDT distribution has been plotted separately for each of the background modes in Fig. 5.11. The modes with two missing pions have a lower BDT distribution than those with 1 track, indicating that the extra missing particle allows better discrimination.

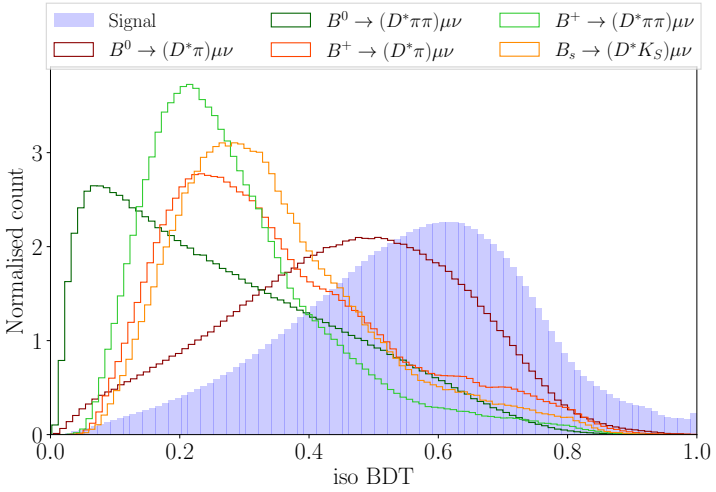


Figure 5.11: Isolation BDT distribution in signal (shaded) and background (line) modes

In total, 38 isolation variables were considered for the BDT. A KS test [103] was used to identify which variables gave the best separation between the signal and background modes. The variables were ranked using the calculated score from the KS test, and the 10 variables with the best score were used in the BDT training. Four additional non-isolation variables were also used, as this led to increased performance of the BDT compared with only training on isolation variables. These additional variables are: $p_\perp$, log(1-DIRA(B)), B vertex χ^2 and $m^2_{miss} = (p_B - p_{D^*} - p_\mu)^2$, where p_B, p_{D^*} and p_μ are the 4-momenta of the B, D^* and muon respectively. In addition, the BDT was trained several times with a varying number of isolation variables, and the performance was analysed by measuring the area under the ROC curve, and the BDT training time. In general, increasing the number of isolation variables increases the area under the ROC curve, but it was observed that using more than 10 variables yielded negligible gains in performance, whilst increasing the training time.

The isolation variables are defined below. In every case, the isolation is performed relative to the B candidate tracks. The "ptasy" and "cc IT" are cone isolation variables [104], which are constructed by drawing a cone around the candidate tracks in $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ where ϕ is the azimuthal angle and η is the pseudorapidity, and checking for tracks and energy deposits in the calorimeter within this cone. The algorithm returns variables determining the activity around the candidate

tracks. The numbers in each variable relate to the size of the cone. The other two variables listed below are created by taking a vertexed particle and sequentially adding other tracks from the event to build a new vertex.

- **B ptasy**: The p_T asymmetry, defined as:

$$\text{ptasy} = \frac{|\vec{p_T}(track) - \vec{p_T}(cone)|}{|\vec{p_T}(track) + \vec{p_T}(cone)|}\,, \tag{5.2}$$

  where *track* refers to the B candidate and $p_T(cone)$ is the vector sum of the p_T of all charged particles found within a cone of a certain size, excluding the signal tracks.

- **B cc IT**: Transverse isolation of the cone with respect to the candidate track. This is defined as:

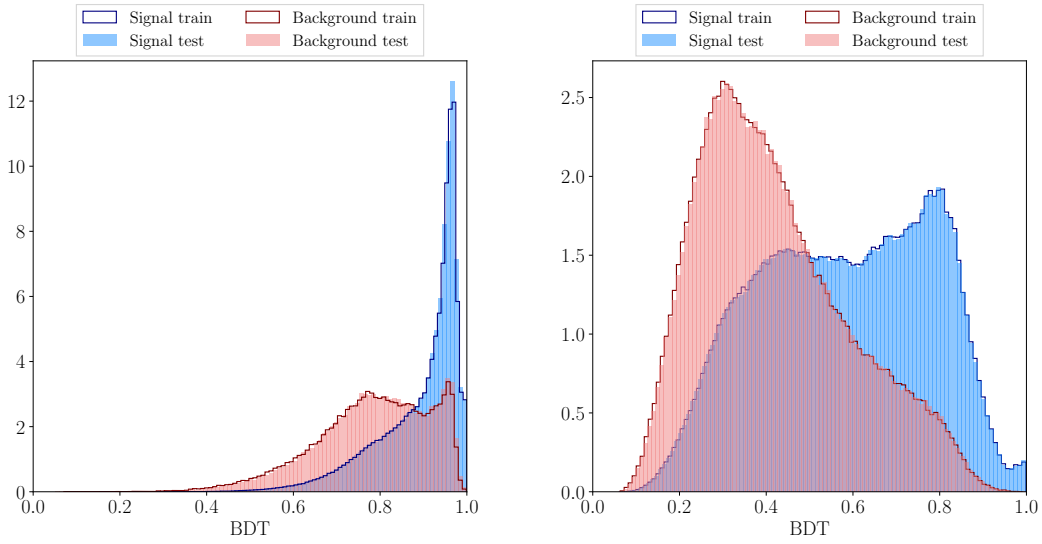
$$\text{IT} = \frac{p_T(track)}{p_T(track + cone)} \tag{5.3}$$

- **B smallest $\Delta\chi^2$ mass 1 track**: mass of the candidate with the smallest $\Delta\chi^2$

- **B smallest $\Delta\chi^2$ mass 2 tracks**: mass of the candidate with the smallest $\Delta\chi^2$ when adding one track to the combination that has the smallest $\Delta\chi^2$ when adding one track

A tight cut was placed on the isolation BDT to reduce the contribution of $B \to D^{**}\mu\nu$ decays (which is the most dominant background) in the data to a low level. An alternative approach was considered, where the isolation BDT is used instead of m_{corr} as the 4th fit variable. However, this would lead to a higher fraction of $B \to D^{**}\mu\nu$ in the data sample, which leads to increased systematic uncertainties. Therefore, it was preferred to use m_{corr} in the fit, and cut tightly on the BDT.

The distributions of the BDT variables in signal and background MC can be seen in Fig. 5.14. The BDT working point was chosen to be 0.566. With this cut, the signal efficiency is 59.4% and the background efficiency is 24.7%.



(a) No sample weights **(b)** With sample weights

Figure 5.12: Signal and background distributions in the isolation BDT, showing effect of sample weights

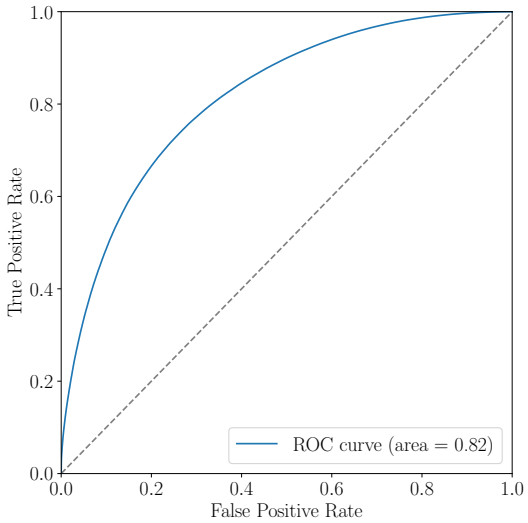


Figure 5.13: ROC curve for isolation BDT

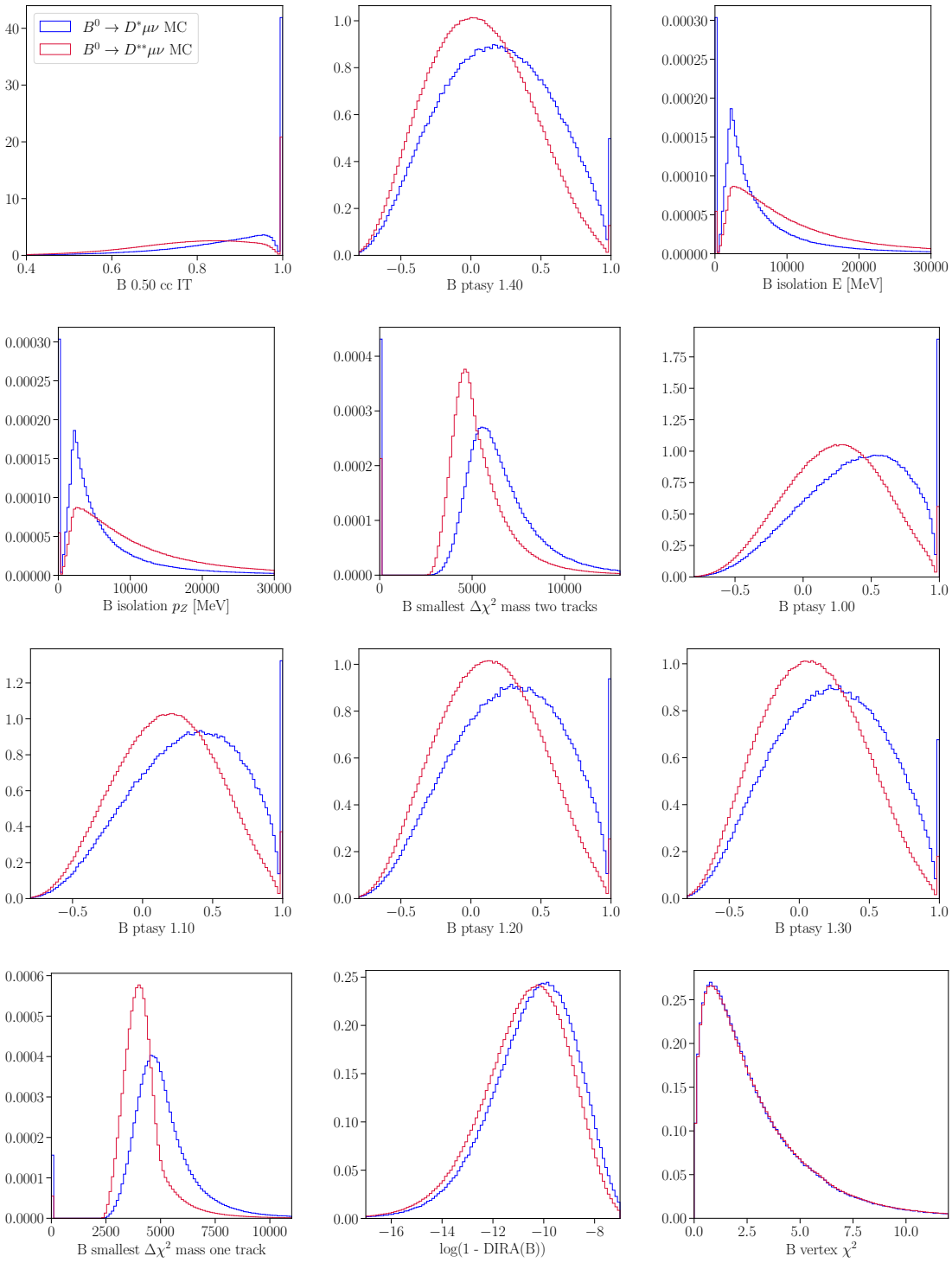

$B^0 \to D^* \mu\nu$ MC
$B^0 \to D^{**} \mu\nu$ MC
B 0.50 cc IT
B ptasy 1.40
B isolation E [MeV]
B isolation p_Z [MeV]
B smallest $\Delta\chi^2$ mass two tracks
B ptasy 1.00
B ptasy 1.10
B ptasy 1.20
B ptasy 1.30
B smallest $\Delta\chi^2$ mass one track
log(1 - DIRA(B))
B vertex χ^2

Figure 5.14

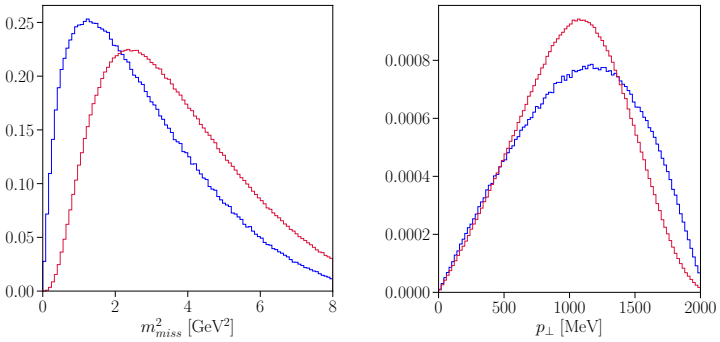


Figure 5.14: Signal and background distributions of variables used in the BDT training

5.4 Corrections to simulation

The method of describing the data with simulation-derived templates to extract the angular observables strictly requires the templates to reliably describe all biases in the data from acceptance and reconstruction effects. Therefore, deficiencies in the simulation must be corrected. Two methods of applying corrections are outlined. Firstly, the MC samples generated with a simplified, fast simulation are corrected to align them with the more realistic full LHCb simulation. Then, a data-driven study using $B^0 \to D^{*-}\pi^+$ decays is used to account for data/MC discrepancies. Residual data/MC discrepancies are handled using polynomial corrections within the angular fit itself, as described in Sec. 4.3.4.

5.4.1 Aligning fast simulation with full LHCb simulation

As mentioned earlier, tracker-only simulation is used to create the fit templates. Tracker-only simulation cannot emulate the hardware trigger. Therefore, to account for the trigger selection on the data, a set of correction weights is derived and applied to the tracker-only simulation when creating the fit templates.

To carry out the reweighting, the fullsim validation MC samples are used, since they can emulate the effect of the trigger on the data sample. The trigger selection was applied to the fullsim MC, fullsim/tracker MC ratio histograms were made in the fit variables. Cubic splines were fitted to the histograms, using the same process as the sequential reweighting described in the previous section. A single set of

weights was obtained. These weights are applied to the tracker-only simulation when making the angular fit templates and they correct all the fit variables simultaneously, shifting the MC shapes to account for the trigger selection on the data sample, and other small residual differences. Since the weights from one step are used in the ratio histogram for the next step, the final correction weights could depend on the order chosen for the variables. This was checked by changing the order and recalculating the corrections, but no significant difference was observed. The weights are applied to all the fit templates that are derived from tracker-only simulation, i.e. all apart from the combinatorial template. The tracker-only samples used to calculate these corrections are separate from those used to make the fit templates.

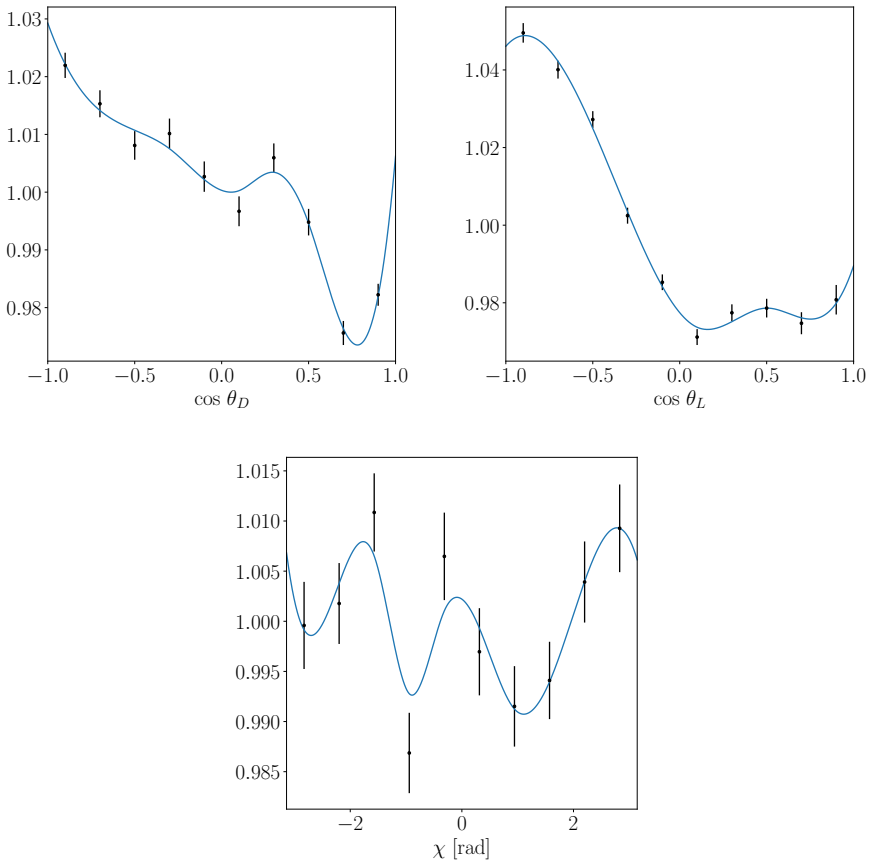


Figure 5.15: Ratio histograms for the three decay angles, fitted with cubic spline functions

Figure 5.15 shows the correction spline functions for the three decay angles, and Fig. 5.16 shows a histogram of the correction weights obtained at the end

of the reweighting procedure.

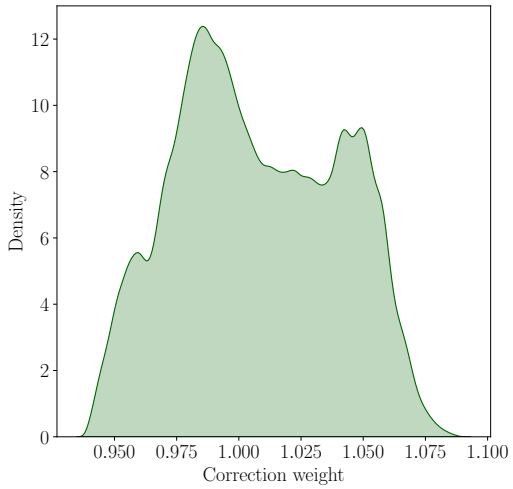


Figure 5.16: Distribution of final correction weights applied to MC samples

5.4.2 Aligning simulation with data

When comparing the data and simulation samples, some small MC mismodelling effects were seen in $\cos\theta_D$. Specifically, the simulation was observed to be slightly higher than the data at negative $\cos\theta_D$, and slightly lower at positive $\cos\theta_D$. One method of correcting these discrepancies is to compare data and simulation in a control mode. $B^0 \to D^{*-}\pi^+$ was chosen for this control study, since it is a Scalar $\to$ Vector + Scalar decay. It is a fully polarised decay and thus has only one decay angle describing the spin-1 D^* system, which is the equivalent of $\cos\theta_D$. Another advantage is that it is a fully reconstructed decay, meaning backgrounds can be more easily isolated and removed.

When processing the $B^0 \to D^{*-}\pi^+$ events, it is important to align the selection as closely as possible to the $B^0 \to D^{*-}\mu^+\nu_\mu$ data to avoid introducing additional biases. Therefore the same preselection was used, except that the `isMuon` requirement on the muon was removed.

Before comparing the data and simulation for the control sample, the dominant backgrounds must be removed. To achieve this, a fit was carried out to $m(D^*\pi) -$

$m(D^*) + m(D^*)_{PDG}$, where $m(D^*)_{PDG}$ is the D^* mass reported by the Particle Data Group (PDG) [75]. This fit yielded a set of sWeights, which was applied to the data sample to remove the most dominant backgrounds. Additionally, fake D^0 and fake D^* backgrounds were reduced by cuts to $m(D^0)$ and ΔM. These cuts are the same as those used for the signal selection, as shown in Tab. 5.2.

Fit to $m(D^*\pi)$

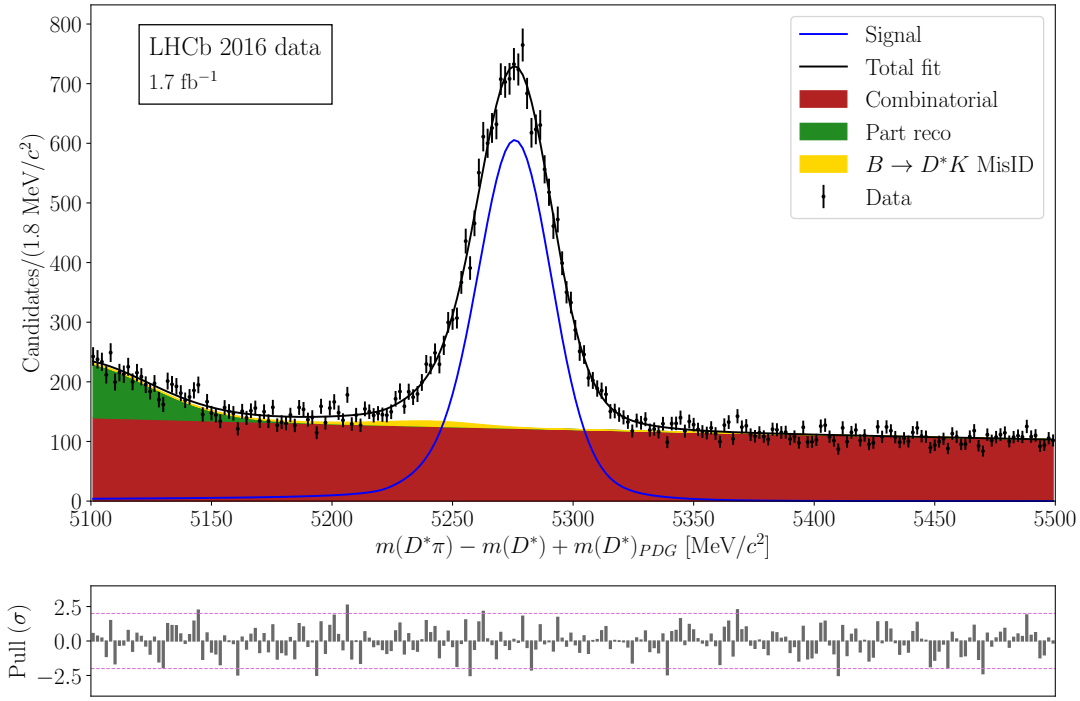


Figure 5.17: Fit to $m(D^*\pi)$ using 2016 data. The components are described in the text. The fit has a $\chi^2/\text{dof} = 1.35$.

The $m(D^*\pi)$ fit is shown in Figure 5.17. The signal is modelled by the sum of a Crystal Ball PDF [105] and a Gaussian PDF, with shape parameters (all parameters apart from the normalisation factor) constrained from a separate fit to MC. The combinatorial background is modelled by an exponential function. There is a relatively high level of combinatorial background in this sample, due to many random pions passing the selection, since a dedicated $B^0 \to D^{*-}\pi^+$ preselection is not used. The partially reconstructed background (decays such as $B \to D^*\pi\pi$) is modelled by a Gaussian PDF. Finally, there is a small component of $B^0 \to D^{*-}K^+$

decays, where the kaon is misidentified as a pion. The fit was performed using the `zfit` package [106], and has $\chi^2/dof = 1.35$.

Initial reweighting for multiplicity and muon p_T

The first reweighting step aims to align the normalised `nTracks` distribution of the $B^0 \to D^{*-}\pi^+$ sWeighted data to the corresponding distribution from the truth-matched simulation. Then, correction weights for the p_T distribution of the pion are calculated. These weights are subsequently used to correct the `nTracks` and muon p_T distributions in all the MC samples used for the angular fit templates.

To calculate the correction weights, data/MC ratio histograms were created in the `nTracks` and pion p_T variables. The histograms were fitted with cubic spline functions. The corrections can be subsequently applied to any other MC sample by assigning each event a weight equal to the spline value at that particular point.

Figure 5.18 shows the `nTracks` and pion p_T distributions before and after the reweighting procedure. The weights successfully align the distributions in MC to more closely match the data. The ratio histograms and spline functions used to calculate the weights are shown in Fig. 5.19. Furthermore, the effect of the

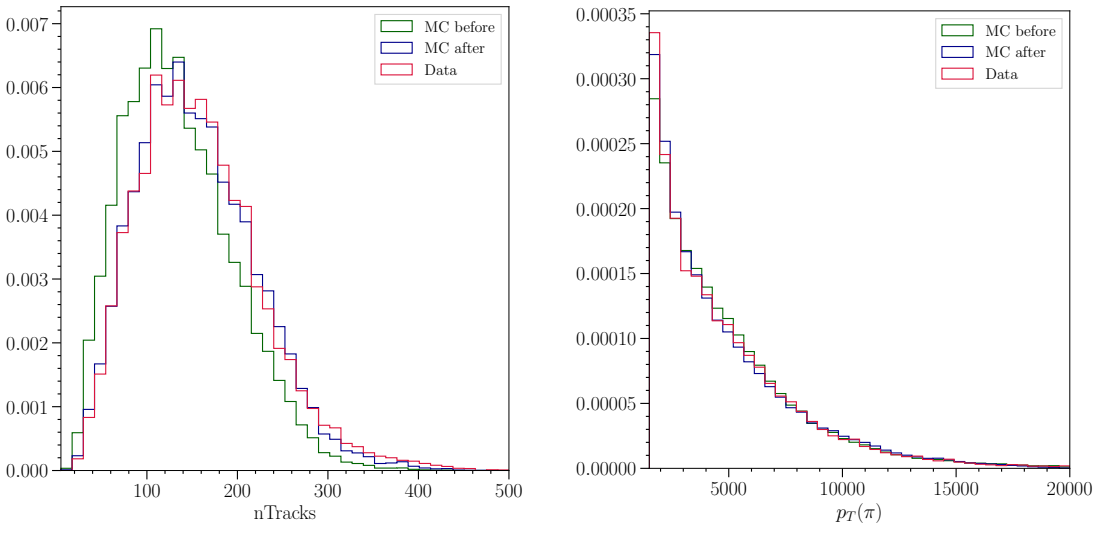


Figure 5.18: Comparison of $B^0 \to D^{*-}\pi^+$ sWeighted data and truth-matched MC in `nTracks` and pion p_T. The effect of the correction weights is shown by plotting the MC with and without the weights.

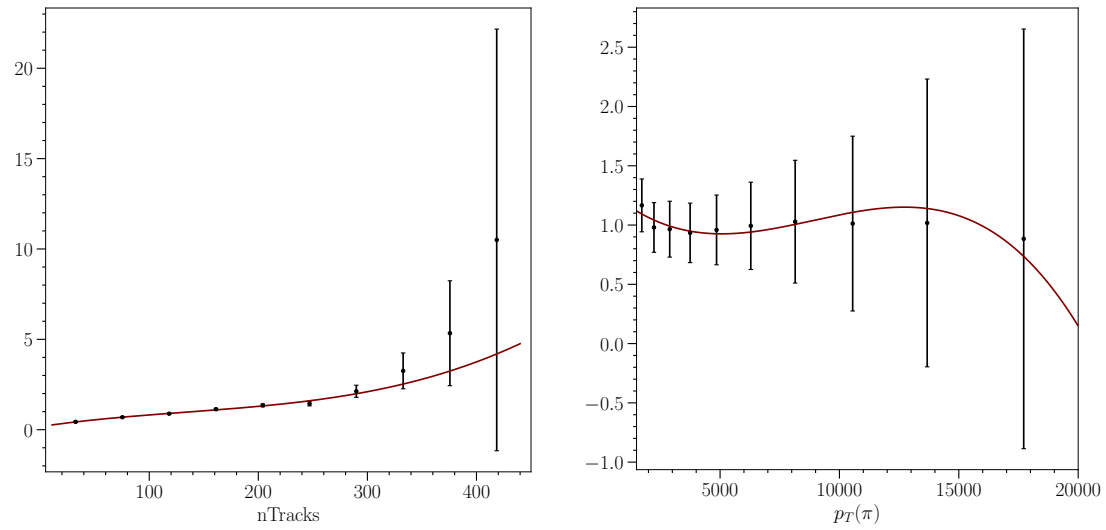


Figure 5.19: Ratio histograms (black points) and fitted cubic spline functions (red line) used to derive the `nTracks` and pion p_T correction weights.

combined `nTracks` and p_T weights on the $B^0 \to D^{*-}\mu^+\nu_\mu$ simulation are shown in Fig. 5.20. The shape change is relatively modest and the distributions remain smooth after the reweighting.

Derived correction to $\cos\theta_D$

The next step is to compare the sWeighted data with the truth-matched $B^0 \to D^{*-}\pi^+$ simulation, and derive data/MC corrections. The method used was to make data/MC ratio histograms in the three decay angles. This is shown in Fig. 5.21 for $\cos\theta_D$. Each ratio histogram was then fitted with a cubic spline function, to get a continuous correction. This was done in a sequential manner; a spline function was created for the first variable, then each event was assigned a weight equal to its spline value at that point. These weights were then applied when creating the ratio histograms in the next variable, which was then also fitted with a spline function. Eventually, a single set of weights was derived, which can be applied to the $B^0 \to D^{*-}\mu^+\nu_\mu$ simulation samples, to correct data/MC differences.

The results of the control study can be seen in Figure 5.22. The top left plot shows one ratio histogram with the continuous cubic spline function used to derive the corrections. The top right plot shows the effect of the control mode corrections

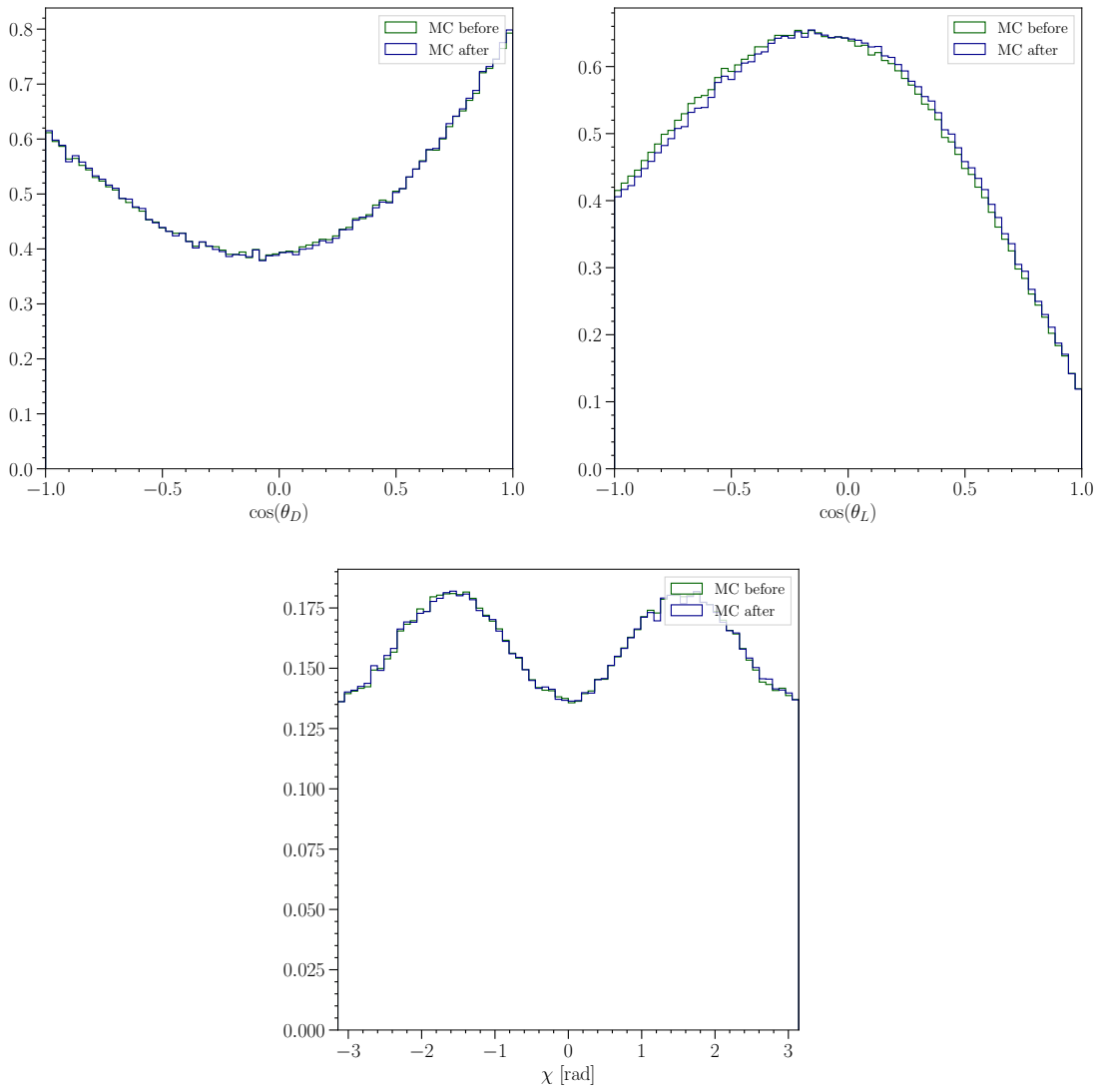


Figure 5.20: Decay angles in the $B^0 \to D^{*-}\mu^+\nu_\mu$ MC, before and after the `nTracks` and p_T weights applied.

on the $\cos\theta_D$ shape in $B^0 \to D^{*-}\mu^+\nu_\mu$. The correction weights shift the MC higher at $\cos\theta_D$ close to 0 and shift the MC lower at $\cos\theta_D$ close to ± 1. The other two decay angles are also shown for completeness.

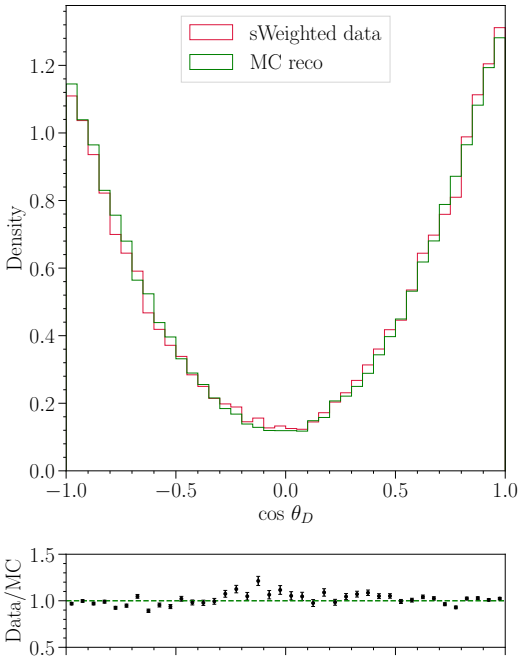


Figure 5.21: $\cos\theta_D$ in $B^0 \to D^{*-}\pi^+$ data and MC with data/MC ratio histogram shown below. The data/MC ratio histogram was used to derive the correction splines.

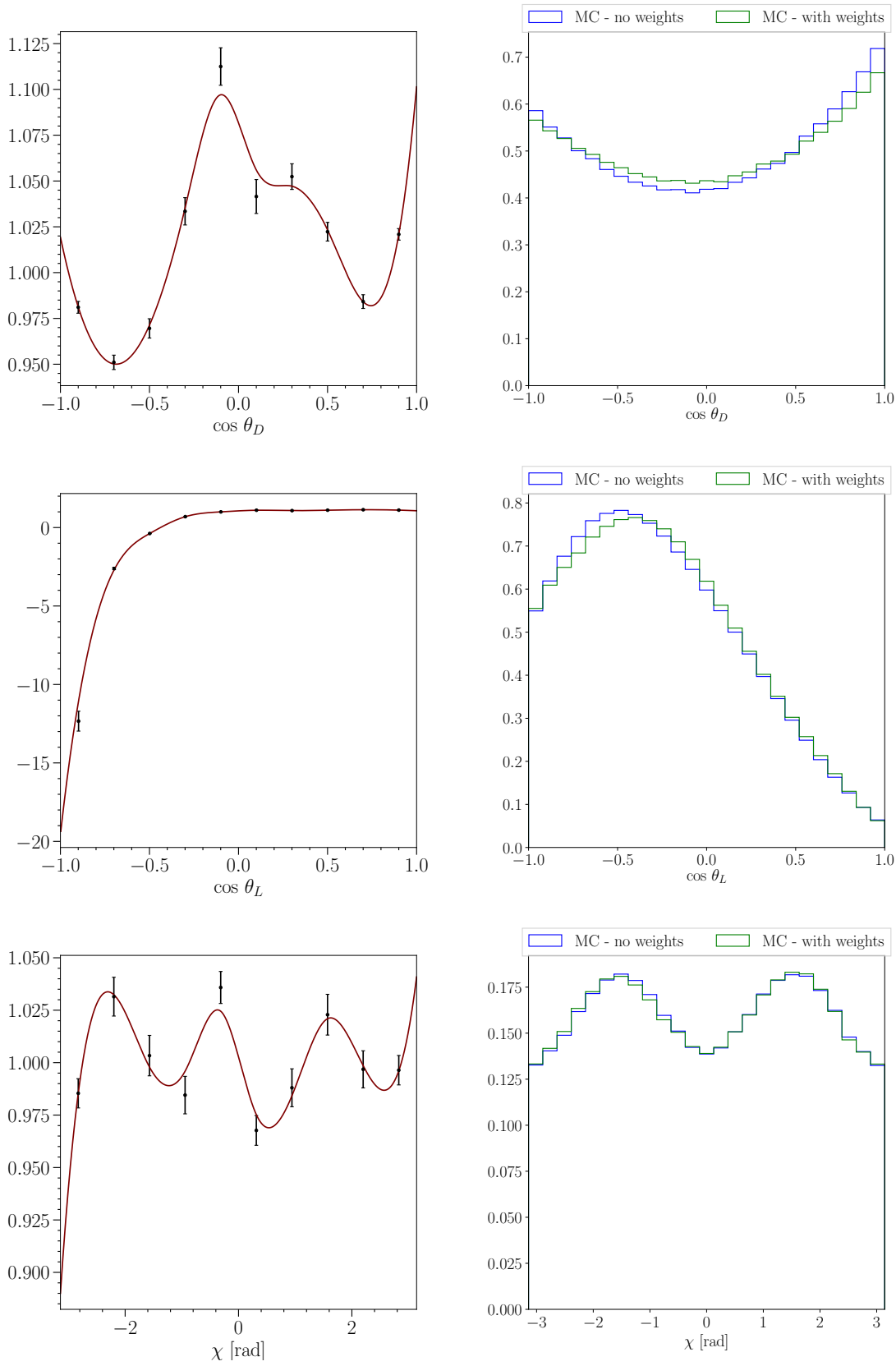


Figure 5.22: (Left) Correction functions for decay angles, taken from $B^0 \to D^{*-}\pi^+$ events. (Right) $B^0 \to D^{*-}\mu^+\nu_\mu$ MC, with and without control mode corrections applied.

5.5 Extracting the generator-level I_X from simulation

5.5.1 Introduction

This section describes the generation of a truth-level $B^0 \to D^{*-}\mu^+\nu_\mu$ MC sample, using the same decay model as the LHCb MC that is later used to create the templates in the four-dimensional angular fit. A three-dimensional parametric fit is performed to the truth-level sample, yielding 12 "true" I_X values, that describe the decay model in the LHCb simulation. The values obtained from the fit are validated by applying the Method of Moments [107].

These "true" I_X values are later used to cancel out the model dependence when building the 12 angular templates. In addition, they are used to test the template fit, by comparing them to values obtained by performing the template fit on signal-only LHCb MC.

5.5.2 Parametric fit

The RapidSim package [108] can be used to create generator-level Monte Carlo samples. In particular, a 4π acceptance can be applied to the final-state particles, and the neutrino momentum is included in the calculation of the B^0 meson kinematics. Another advantage of RapidSim is that it allows large samples to be generated with little computing time - millions of events can be generated in a few minutes.

RapidSim can be built with the EvtGen package [109], allowing the same decay models to be used as in the full LHCb MC. This means that it is possible to reliably measure the I_X coefficients that define the model used to generate the LHCb MC. In the absence of a tool that allows a user to manually specify the I_X coefficients when generating a simulation sample, this RapidSim study is the best way to extract the I_X values from the LHCb MC. In theory, it may be possible to calculate the I_X coefficients from a particular EvtGen model, given the form factor parameters used, but in practice this is not trivial.

Parameter	Result
I_{1s}	0.3515 ± 0.0004
I_{2c}	-0.5177 ± 0.0009
I_{2s}	0.1162 ± 0.0006
I_3	-0.1759 ± 0.0006
I_4	-0.3049 ± 0.0006
I_5	0.2374 ± 0.0006
I_{6c}	0.0129 ± 0.0009
I_{6s}	-0.2929 ± 0.0006
I_7	-0.0010 ± 0.0008
I_8	0.0002 ± 0.0008
I_9	0.0000 ± 0.0006
I_{1c}	0.5352 ± 0.0006

Table 5.7: Fit results for the parametric fit to the truth-level RapidSim sample. I_{1c} is derived using the normalisation condition in Eq. (4.16)

Equation (3.46) is used to create the unbinned fit PDF. A three-dimensional parametric shape in $(\cos\theta_D, \cos\theta_L, \chi)$ space is created for each of the 12 angular functions, and each corresponding I_X coefficient is a floating parameter in the fit. The fit uses a negative log likelihood method, and is carried out using the `zfit` package.

The fit was performed to a sample of 3 million RapidSim events. An alternative procedure was tested, where a smaller dataset was randomly sampled from the full RapidSim dataset and multiple fits were carried out. The average results for the I_X coefficients from these fits were compared to the results obtained from a single fit to the full dataset, and were found to be consistent with the single fit. The one dimensional fit projections are shown in Fig. 5.23. There is good agreement between the MC and the fit shape in all three dimensions.

The fit results for each parameter are shown in Tab. 5.7. The results obtained from this parametric fit are a crucial component of the template fit procedure. They are used in two separate ways:

- The results are used in the denominator of the weight calculation in Eq. (4.13) for the templates, to create the total model used in the LHCb MC (since

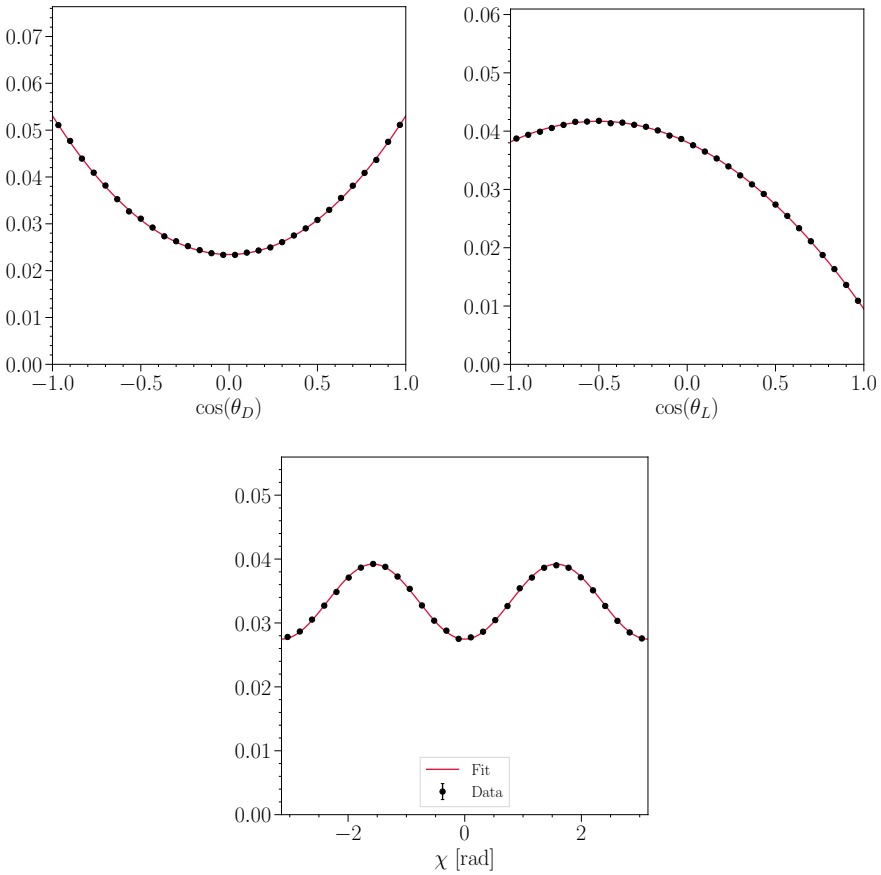


Figure 5.23: One dimensional fit projections for the parametric fit to the truth level sample.

this model is the same as the one used to generate the RapidSim sample). Dividing by this model is the process that ensures the templates are model independent.

- They are used to validate the signal-only template fit, by investigating if performing this fit using the LHCb $B^0 \to D^{*-}\mu^+\nu_\mu$ MC as the data sample yields results that are consistent with those from the RapidSim fit. If so, this is a reliable indication that the template fit successfully accounts for the reconstruction effects that are present in the LHCb MC, but not in the RapidSim MC. This study is shown in Sec. 5.7.1.

5.5.3 Fit validation with toys

To ensure the unbinned fit is stable, 2500 toy fits were performed. The toy fits were done by first creating a three-dimensional PDF, where the values of the I_X coefficients were assigned to their values shown in Tab. 3.1. From this PDF, a toy data sample was generated, with a yield matching the size of the RapidSim sample used in the default fit. For each iteration, the toy data was randomly resampled, then fitted as described in the section above. The starting values of the fit parameters were also randomly fluctuated for each fit. Of the 2500 toy fits, 95% of them converged.

Pull distributions were calculated for the 11 fitted I_X coefficients, as well as I_{1c}, which is derived from the normalisation condition. The pull for each parameter is defined as:

$$\text{pull}(I_X) = \frac{(I_X^{fit} - I_X^{init})}{\Delta I_X} \tag{5.4}$$

where I_X^{init} is the starting value for each parameter, I_X^{fit} and ΔI_X are the fitted values and fit errors respectively. The distributions were fitted with gaussians, and the results are plotted in Fig. 5.24. The pull mean and width for each observable are summarised in Tab. 5.8. They show no significant signs of bias and the fit can accurately estimate the uncertainties.

5.5.4 Method of Moments

The Method of Moments, as outlined in Ref. [107], is a technique that allows angular observables to be estimated from data without the use of a fit. It is used in this section as a cross-check of the results of the truth-level parametric fit. The definition of the angular moments proceeds from the full normalised angular decay rate shown in Eq. (3.46), which can be written more compactly as:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}\Omega} = \frac{9}{32\pi} \sum_X I_X \cdot f_X(\cos\theta_D, \cos\theta_L, \chi) \ , \tag{5.5}$$

where $X \in \{1c, 1s, 2c, 2s, 3, 4, 5, 6c, 6s, 7, 8, 9\}$, f_X are the 12 angular functions in Eq. (3.46), and $\mathrm{d}\Omega = \mathrm{d}\cos\theta_D\,\mathrm{d}\cos\theta_L\,\mathrm{d}\chi$. For each angular term, the corresponding

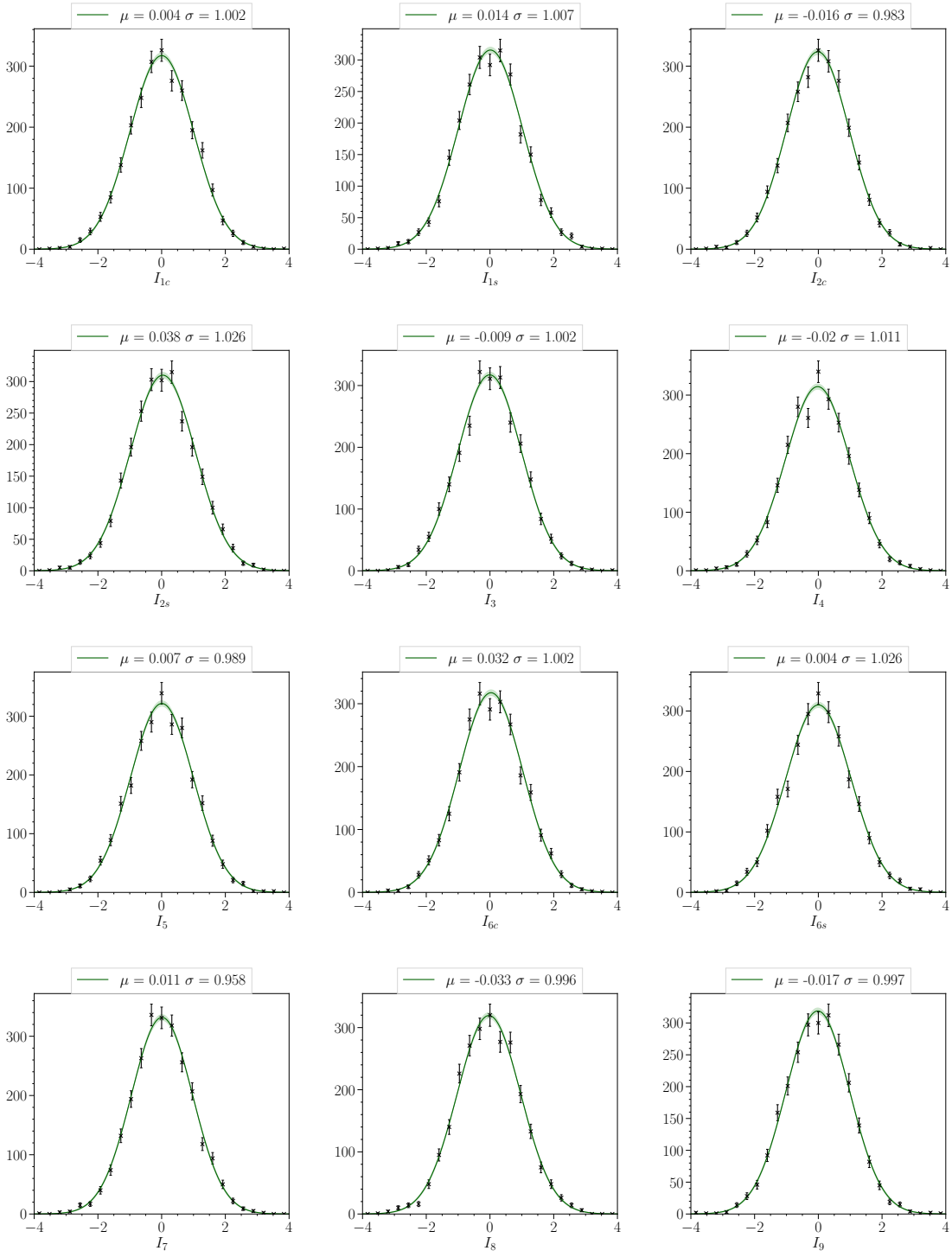


Figure 5.24: Fitted pull distributions from 3D parametric toy fits. The black points are the toy fit results, the green line is the fitted shape. The fitted mean and width are shown above each plot.

	μ	σ
I_{1c}	0.004 ± 0.022	1.002 ± 0.015
I_{1s}	0.014 ± 0.020	1.007 ± 0.015
I_{2c}	-0.016 ± 0.020	0.983 ± 0.014
I_{2s}	0.038 ± 0.021	1.026 ± 0.015
I_{6c}	0.032 ± 0.020	1.002 ± 0.015
I_{6s}	0.004 ± 0.021	1.026 ± 0.015
I_3	-0.009 ± 0.020	1.002 ± 0.014
I_4	-0.020 ± 0.021	1.011 ± 0.015
I_5	0.007 ± 0.020	0.989 ± 0.014
I_7	0.011 ± 0.020	0.958 ± 0.014
I_8	-0.033 ± 0.020	0.996 ± 0.014
I_9	-0.017 ± 0.020	0.997 ± 0.014

Table 5.8: Pull distribution results, showing the fitted mean and width for each coefficient

moment, M_X, is defined as:

$$\begin{aligned} M_X &= \int \frac{\mathrm{d}\Gamma}{\mathrm{d}\Omega} f_X(\cos\theta_D, \cos\theta_L, \chi) \\ &= \int \frac{9}{32\pi} \sum_j I_j \cdot f_j(\cos\theta_D, \cos\theta_L, \chi) \cdot f_X(\cos\theta_D, \cos\theta_L, \chi) \ . \end{aligned} \tag{5.6}$$

The M_X is the first moment of each angular function, and can be considered as the function's expectation value. By calculating the integrals in Eq. (5.6), the

moments can be related to the I_X as:

$$\begin{aligned}
I_{1s} &= \frac{35}{16}M_{1s} - \frac{5}{8}M_{2c} + \frac{15}{16}M_{2s} - \frac{7}{8} \\
I_{2c} &= -\frac{25}{8}M_{1s} + \frac{15}{2}M_{2c} - \frac{15}{8}M_{2s} + \frac{5}{2} \\
I_{2s} &= \frac{25}{16}M_{1s} - \frac{15}{8}M_{2c} + \frac{45}{16}M_{2s} - \frac{5}{8} \\
I_3 &= \frac{25}{8}M_3 \\
I_4 &= \frac{25}{8}M_4 \\
I_5 &= \frac{5}{2}M_5 \\
I_{6s} &= 3M_{6s} - 2M_{6c} \\
I_{6c} &= 8M_{6c} - 2M_{6s} \\
I_7 &= \frac{5}{2}M_7 \\
I_8 &= \frac{25}{8}M_8 \\
I_9 &= \frac{25}{8}M_9 \ ,
\end{aligned} \tag{5.7}$$

Note that the constraint from the normalisation condition in Eq.(4.16) has also been used here. The M_X can be easily calculated from the RapidSim MC sample, with no additional inputs. The process is equivalent to averaging each angular function over all events, i.e.

$$M_X = \frac{1}{N}\sum_{i=1}^{N} f_X(\cos\theta_D^i, \cos\theta_L^i, \chi^i) \ , \tag{5.8}$$

where N is the total number of events in the sample. The uncertainty on M_X is given by:

$$\delta M_X = \frac{\sigma_X}{\sqrt{N}} \ , \tag{5.9}$$

where σ_X is the standard deviation of the angular function evaluated for all events. The I_X values calculated from the Method of Moments are shown in Tab. 5.9. A comparison with the parametric fit results can be seen in Fig. 5.25. The agreement is excellent, the two sets of values have a χ^2/dof of 0.15. This indicates that truth-level I_X obtained from the parametric fit are reliable and unbiased.

Coefficient	Value
I_{1c}	0.5349 ± 0.0011
I_{1s}	0.3517 ± 0.0006
I_{2c}	-0.5181 ± 0.0019
I_{2s}	0.1164 ± 0.0010
I_{6c}	0.0137 ± 0.0015
I_{6s}	-0.2926 ± 0.0009
I_3	-0.1759 ± 0.0007
I_4	-0.3055 ± 0.0008
I_5	0.2382 ± 0.0008
I_7	-0.0016 ± 0.0008
I_8	0.0004 ± 0.0009
I_9	0.0004 ± 0.0007

Table 5.9: I_X values and uncertainties from the Method of Moments. I_{1c} has been calculated using Eq. (4.16).

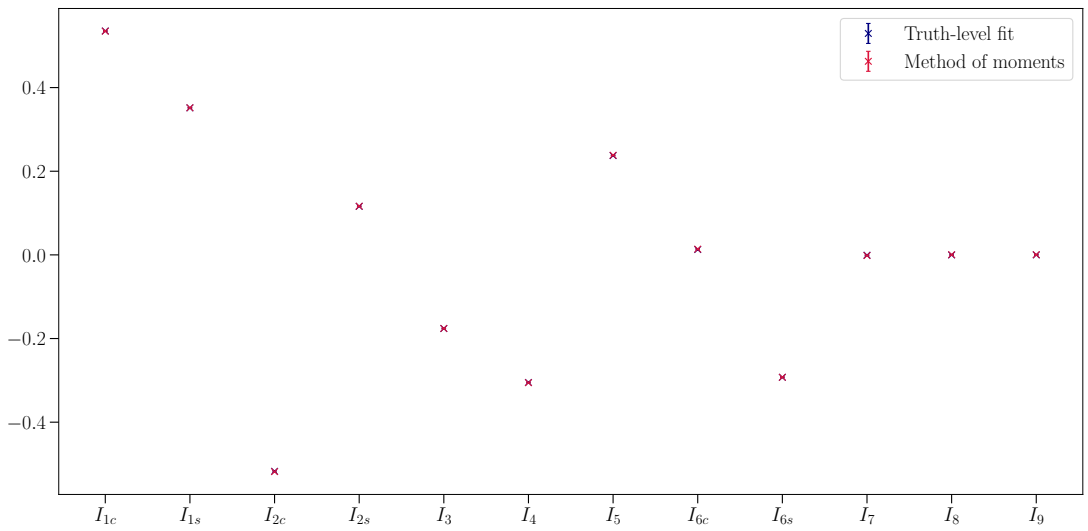


Figure 5.25: Comparison of the I_X values from the parametric fit and from the Method of Moments. I_{1c} has been calculated using Eq. (4.16).

5.6 Model independence of templates

5.6.1 Template shapes under model variation

To verify the model independence of the templates used to describe the signal in the angular fit, a study was performed using RapidSim MC samples (in this section "MC" refers to RapidSim simulation, unless explicitly stated otherwise). As mentioned earlier, RapidSim allows EvtGen to be used to specify the decay model. $B^0 \to D^{*-}\mu^+\nu_\mu$ samples can be generated with the default HQET2 parameters [5] (those used in the full LHCb simulation), as well as alternative HQET2 parameters.

The parametric fit to the true angular distribution is performed to determine the I_X coefficients for the default and varied models. These I_X coefficients are used in the denominator of the weight calculation to derive the templates for the angular fit. Explicitly, two sets of 12 templates were created. One set uses MC generated with a default model, and the template weights use the default I_X coefficients in their calculation. The second set of templates are created from the MC with the alternative model, and the weights use the I_X coefficients from this alternative model. This division by the total model essentially cancels out the model dependence of the angular fit templates in each case. Therefore, the templates created from samples with different decay models should be identical, within statistical uncertainty, since the weights used in the templates account for the respective models.

The 12 angular templates created with two different models can then be compared. The default decay model was chosen, and an alternative model was applied by varying the default HQET2 form factors in the RapidSim configuration by around 20% of their default values. The 20% variation is much larger than the current uncertainties of the form factor parameters, as estimated by the Heavy Flavour Averaging Group (HFLAV) [57]. The exact form factor values are shown in Tab. 5.10. The angular distributions of the $B^0 \to D^{*-}\mu^+\nu_\mu$ samples in both cases can be seen in Fig. 5.26. The plots show that the change in form factors leads to a clear shift in the angular shapes.

[5] Note that in EvtGen, the decay model is called HQET2, however the form-factor parameterisation actually follows the CLN scheme [65]. It will be referred to as HQET2 here.

Sample	ρ^2	h_{A1}	R_1	R_2	R_0
Default	1.122	0.908	1.270	0.852	1.15
Alternative	1.346	1.090	1.524	1.022	1.380

Table 5.10: HQET2 form factors [57] used in the default and alternative models to generate RapidSim MC samples

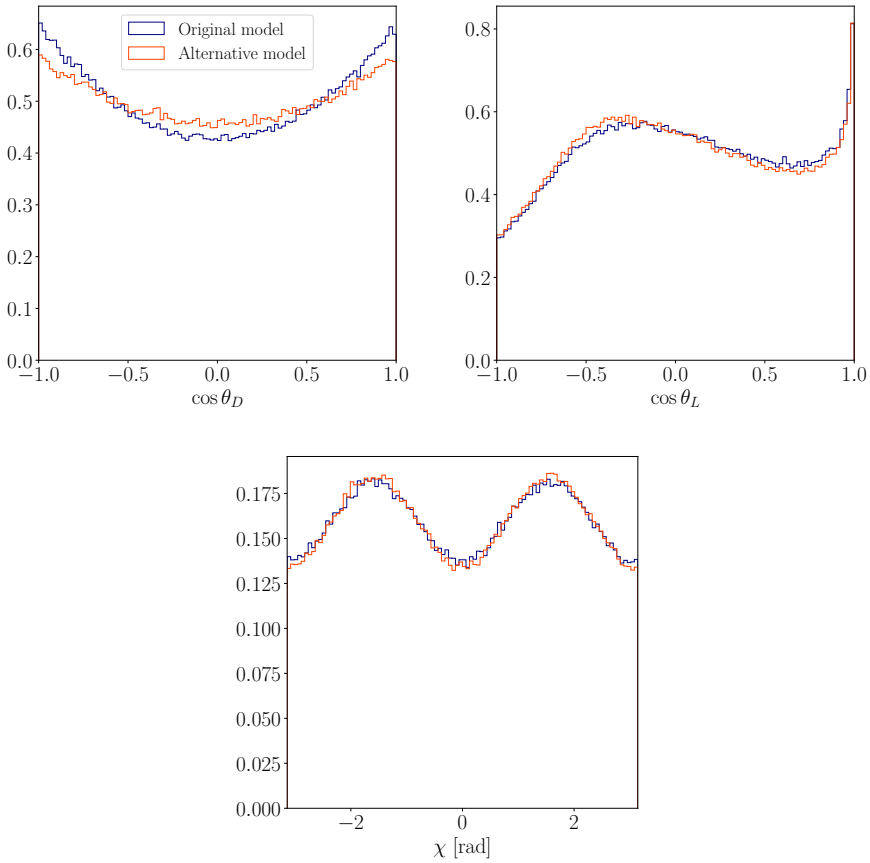


Figure 5.26: Angular distributions for default and alternative HQET2 decay models

To quantify the consistency of the 12 templates for the two models, the templates were compared bin by bin, and a pull was calculated:

$$\text{pull}_X(i) = \frac{T_X(i) - T'_X(i)}{\sqrt{E_X(i)^2 + E'_X(i)^2}} \tag{5.10}$$

where $X \in \{1c, 1s, 2c, 2s, 3, 4, 5, 6c, 6s, 7, 8, 9\}$, i is the bin number, T represents the template histogram values for each bin, and E is the corresponding uncertainty. The primed and unprimed quantities refer to the default and varied decay models.

This yields a set of pull values for each of the 12 template terms. A histogram of the pull values can be plotted for each term. These histograms should have a gaussian shape with a mean of 0 and width of 1 if the two samples are statistically equivalent. The pull plots are shown in Fig. 5.27. The pull distributions are consistent with unit gaussians centred at 0, showing that despite using two significantly different decay models, the resulting template shapes are consistent.

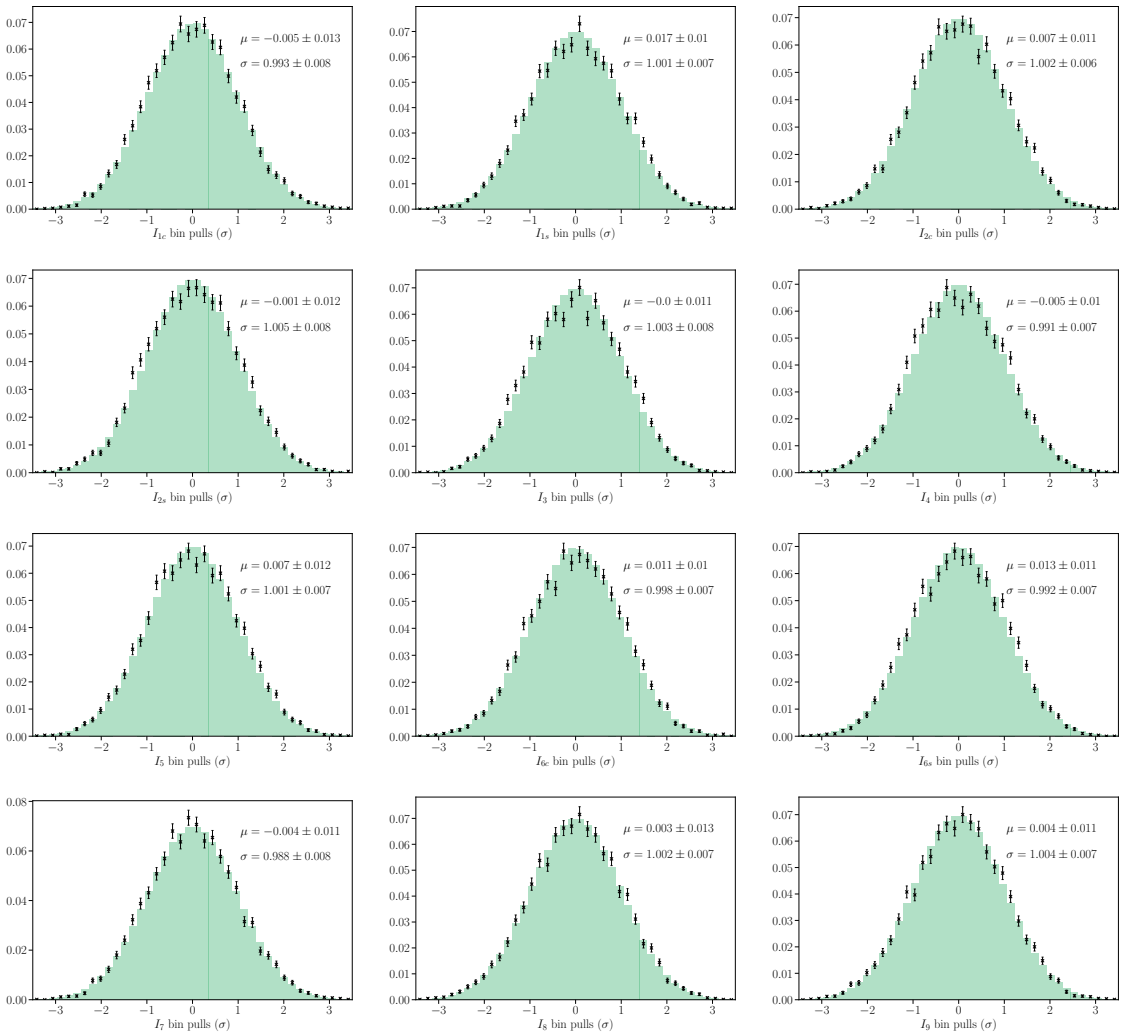


Figure 5.27: Distribution of per-bin pulls between the two models for each template. The green histogram is a unit gaussian centred at 0, the black points show the pulls with their uncertainties. The mean and width of each pull distribution is shown.

5.6.2 Potential biases from q^2 acceptance effects

Since this analysis measures the q^2-integrated I_X, there is a potential loss of information when the four-dimensional $(\cos\theta_D, \cos\theta_L, \chi, q^2)$ decay rate is described by templates in the three decay angles without considering q^2.

Specifically, the templates are required to adequately describe all reconstruction and acceptance effects present in the data. However, since the templates are only sensitive to effects in the decay angles (and m_{corr}), it is possible that effects that bias the q^2 distribution will not be fully captured. If the decay model in the simulation is not representative of the data, the data and simulation will have different q^2 shapes. This means that the simulation will not necessarily respond to a q^2 biasing effect in the same way as the data. Therefore, the size of the potential bias is determined by two factors: the difference between the simulated and realistic decay model; and the severity of the acceptance variation in q^2.

A toy study was initially carried out to investigate both of these effects, then the templates used in the analysis were studied to assess the effects directly.

Toy study

Firstly, a simplified decay model was constructed with two components and four variables. The model can be written as:

$$M(x, y, z, q^2) = J_1(q^2)\cos^2 x \sin^2 y\, e^{3z/2} + J_2(q^2)\sin^2 x \cos^2 y\, e^{-3z/2}\ , \tag{5.11}$$

where the q^2-dependent functions J_1 and J_2 have the form:

$$J_1(q^2) = e^{-\lambda q^2} - 0.25\ , \tag{5.12}$$

$$J_2(q^2) = \frac{1}{q^2 + 5} + 1.05\ , \tag{5.13}$$

with $\lambda = \log(2)/6$ and $q^2 \in [0, 10.8]$.

This model can be used to create q^2-integrated templates in an analogous way to the creation of the signal templates for the main analysis. The first step is to generate events that have (x, y, z, q^2) distributions described by the model. One

million events were generated for this study. Then $J_1(q^2)$ and $J_2(q^2)$ are integrated over the full q^2 range to yield:

$$\hat{J}_1 = \int_0^{10.8} J_1(q^2)\, dq^2 \approx 0.321 \,, \tag{5.14}$$

$$\hat{J}_2 = \int_0^{10.8} J_2(q^2)\, dq^2 \approx 0.943 \,. \tag{5.15}$$

The events are split into two samples; one is used to create the templates, and the other is used to create raw histograms which are then compared to the summed templates. Two templates $h_1(x, y, z)$ and $h_2(x, y, z)$ are created by assigning per-event weights derived from the model functions associated with each of the two terms.

The aim of the study is to assess the agreement between the raw angular shapes and the angular shapes from the templates, when the effect of reconstruction on q^2 is applied to both. The weighted sum of the templates is then calculated as:

$$T(x, y, z) = \hat{J}_1 h_1(x, y, z) + \hat{J}_2 h_2(x, y, z) \,. \tag{5.16}$$

If the templates can successfully account for the q^2 biasing, then $T(x, y, z)$ will have a similar shape in the x, y, z variables to simple histograms in those variables constructed from the raw events, i.e. the templates would accurately recover the original shapes with no additional biases.

However, since this analysis claims to make a model-independent measurement of the angular coefficients, it is important to also test if the template weights adequately cancel out the model used to generate the events. A q^2 biasing effect could impact the templates' ability to do this, since different models have different underlying q^2 distributions, so could respond differently to the effect. An alternative model can be created by altering the $J_1(q^2)$ function, leading to an alternative integrated coefficient ($\hat{J}_1'$). Templates can be created from this alternative model (h_1', h_2'). Now if h_1' and h_2' are model-independent, the original model shape can be recovered by performing the sum:

$$T(x, y, z) = \hat{J}_1 h_1'(x, y, z) + \hat{J}_2 h_2'(x, y, z) \,, \tag{5.17}$$

i.e. combining the *alternative* templates with the *default* coefficients should give the same results as combining the *default* templates with their coefficients. Furthermore, the relevant question for this study is whether this condition holds when an acceptance effect in q^2 is applied to both samples.

This is investigated by generating a set of samples based on different model variations and applying different q^2 acceptance effects. Then, templates were created from these varied samples and summed with the appropriate coefficients. The summed template shape was compared to a raw histogram in (x, y, z) created from the original sample (with the same q^2 acceptance effect applied). The agreement is assessed by a χ^2 test, where the difference between the template and the raw histogram is summed over every bin.

The procedure is repeated for a range of model variations and acceptance functions, effectively performing a two-dimensional grid scan. A KS test is performed on the samples, comparing the q^2 shapes of the default and varied model. The same is done to compare the samples with and without the acceptance effect applied. The two resulting KS test statistics quantify the size of the model and acceptance variations, respectively. A higher test statistic indicates a larger variation between the two samples. Figure 5.28 shows the results of the two-dimensional scan, with the consistency of the templates displayed as a function of the two KS test statistics.

As expected, the larger model and acceptance variations lead to worse agreement between the templates and the raw histograms. However, the plot shows a region where the templates have no significant discrepancy with the raw histograms. This indicates that for *reasonable* acceptance functions and *realistic* model choices for the simulation, integrating over q^2 does not cause any additional bias. More specifically, the range of acceptable model and acceptance variations can be defined by the region in Fig. 5.28 that has a $\chi^2/\text{dof} < 1.1$.

Real effects in templates

It is now important to determine which regime this analysis lies in. To do this, the real q^2 acceptance function present in data was extracted by comparing the q^2 shape

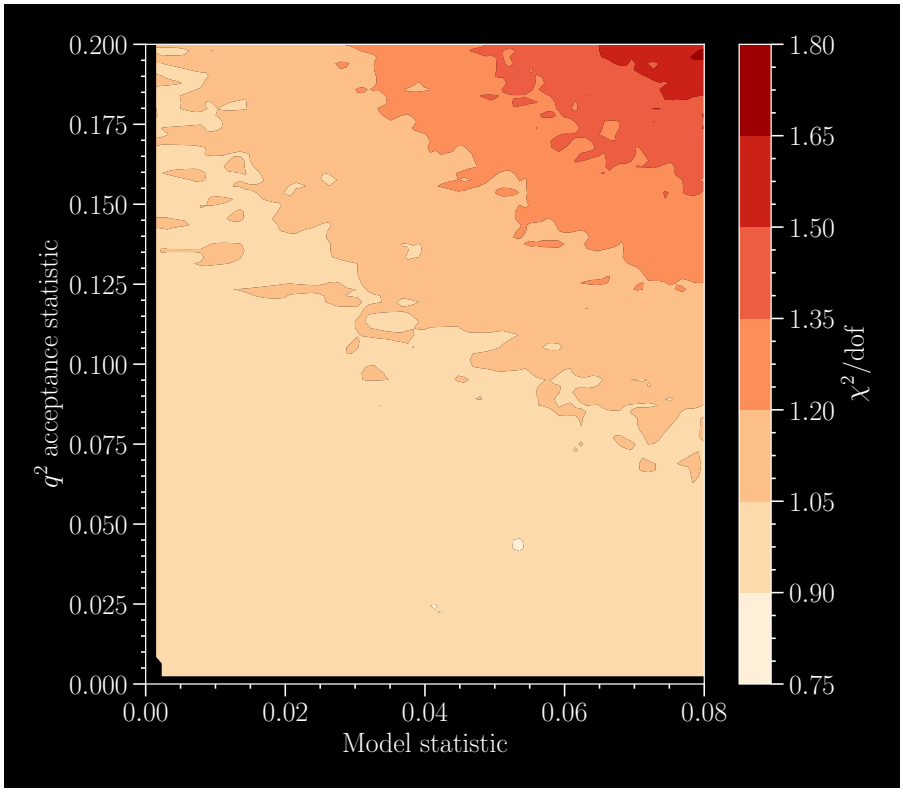


Figure 5.28: Results of toy study: a contour plot showing the variation in agreement between the raw shape and templates with the model shift and acceptance effect. The x and y axes show the KS test statistic corresponding to the applied model and acceptance variations respectively. Darker colours indicate a higher χ^2/dof so a worse agreement.

in a generator-level MC sample to the shape in MC after all selection is applied. The previous section showed that the model variation between simulation and data is the other relevant parameter. Clearly the decay model that best describes the data cannot be known *a priori*. However, the form-factor parameters in the simulation (shown in the top row of Tab. 5.10) are based on the latest global averages, so their associated uncertainties can be used to determine a realistic model variation.

For the acceptance variation, a realistic function can be constructed from LHCb simulation. This is done by creating a ratio of the truth-level and reconstructed q^2 distributions in the simulation. The ratio is then smoothed by fitting with a spline function. The resulting function is an accurate representation of the acceptance variation with q^2 in real data.

For this study, the samples described in Sec. 5.6.1 were used again. Importantly, this represents a much larger model variation than is reasonably expected between

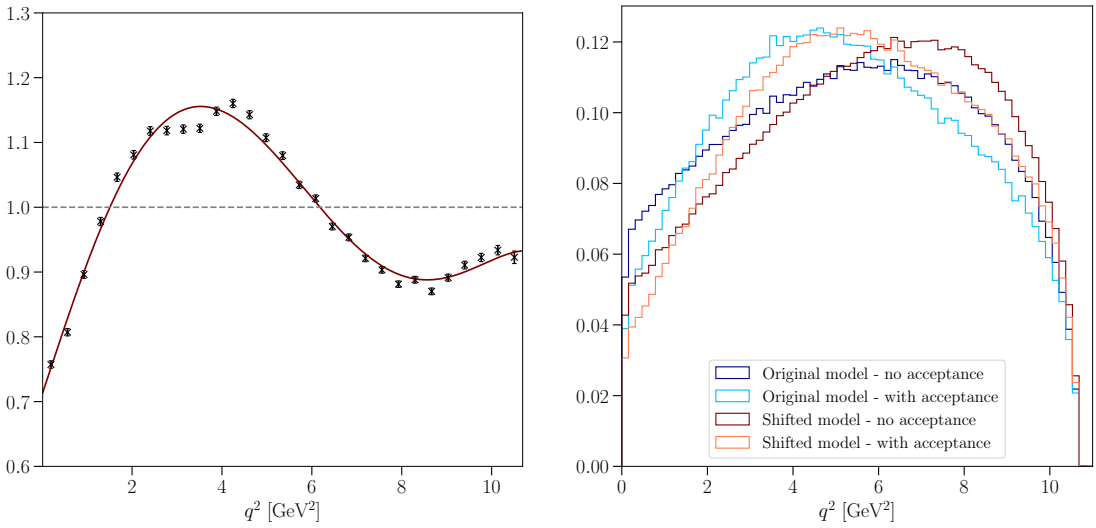


Figure 5.29: Left: Acceptance function derived from LHCb simulation. This was created by making a ratio histogram of the reconstructed and true q^2 distributions, and fitting this ratio with a spline function. Right: q^2 shape from simulation samples generated with models described in Tab. 5.10. For both models, the shapes with and without the acceptance function from data are shown.

simulation and data. As done in the previous section, the templates were created from the varied model, and the agreement with the raw histograms made from the default model was evaluated. The realistic acceptance function was applied to both samples. The model and acceptance variations are illustrated in Fig. 5.29, as well as the acceptance function.

The plots in Fig. 5.30 show the comparison between the summed templates and the raw histograms. There is good agreement in all the variables. This indicates that the templates can account for realistic q^2 effects and remain model independent. The templates generated from an alternative model can be summed with the default I_X to successfully recover the default shape, all in the presence of a q^2 acceptance effect. The consistency of the two shapes plotted in Fig. 5.30 was checked by calculating a χ^2/dof for each dimension, and all variables had a $\chi^2/\text{dof} < 1$, confirming the excellent agreement visible in the plots. The model difference used here is much larger than can be reasonably expected in the real data, and the acceptance function is realistic so this study is a robust check of the potential q^2 biases in this analysis.

In order to compare this to the toy study from the previous section, the KS

test statistics for the applied model and acceptance variations can be calculated. The values are 0.02 for the model statistic and 0.03 for the acceptance statistic. From Fig. 5.28, this point clearly lies within the $\chi^2/\text{dof} < 1.1$ region, supporting the claim that the templates can account for this level of variation in q^2.

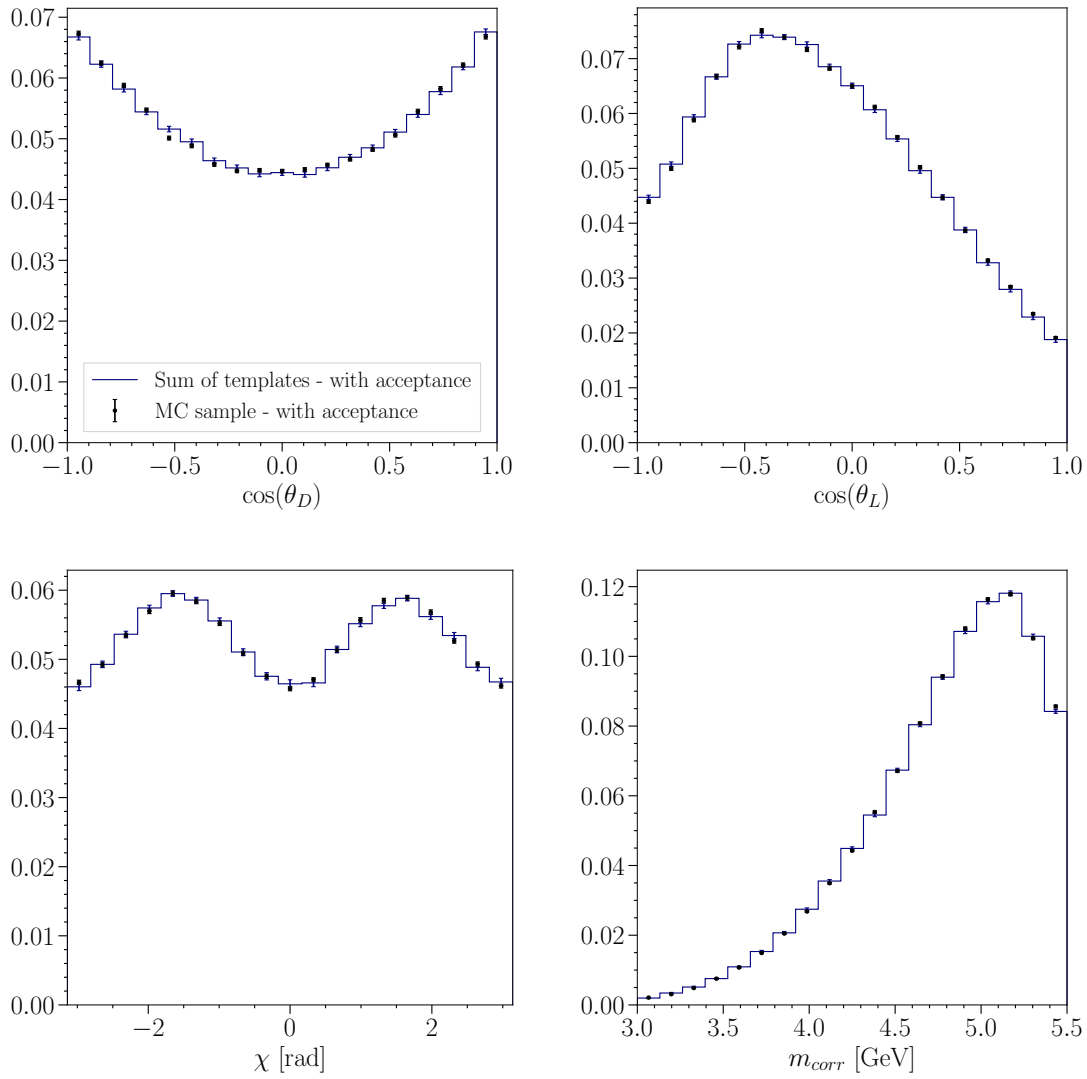


Figure 5.30: The blue line shows one dimensional projections of templates created from the alternative HQET2 model and summed with the true I_X of the default HQET2 model, where the events in the templates have the data acceptance applied. This is compared to the black points, which is simply the generator-level simulation with the original HQET2 model, with the data acceptance function applied.

However, if the q^2 acceptance variation in data was larger, this could introduce biases when the templates are created. To demonstrate this, the study was repeated

using an unrealistically large acceptance function instead of the real function from data. Figure 5.31 is the equivalent of Fig. 5.29 in this case.

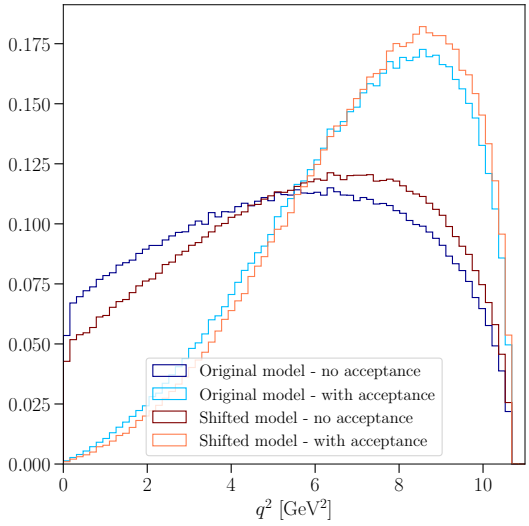


Figure 5.31: q^2 shape from RapidSim MC samples generated with models described in Tab. 5.10. For both models, the shapes with and without the large acceptance function are shown.

The template projections in the large acceptance case are shown in Fig. 5.32. In this case, there are significant differences between the sum of templates and the raw MC shape, when both have the large acceptance function applied. The χ^2/dof values for each projection are all > 1. This indicates that the template method in this analysis is valid when the variation of the acceptance in q^2 is smooth and not too large, as is the case in this analysis. However, if the variation is more extreme, the method begins to show discrepancies. This behaviour confirms what was found in the toy study in Sec. 5.6.2.

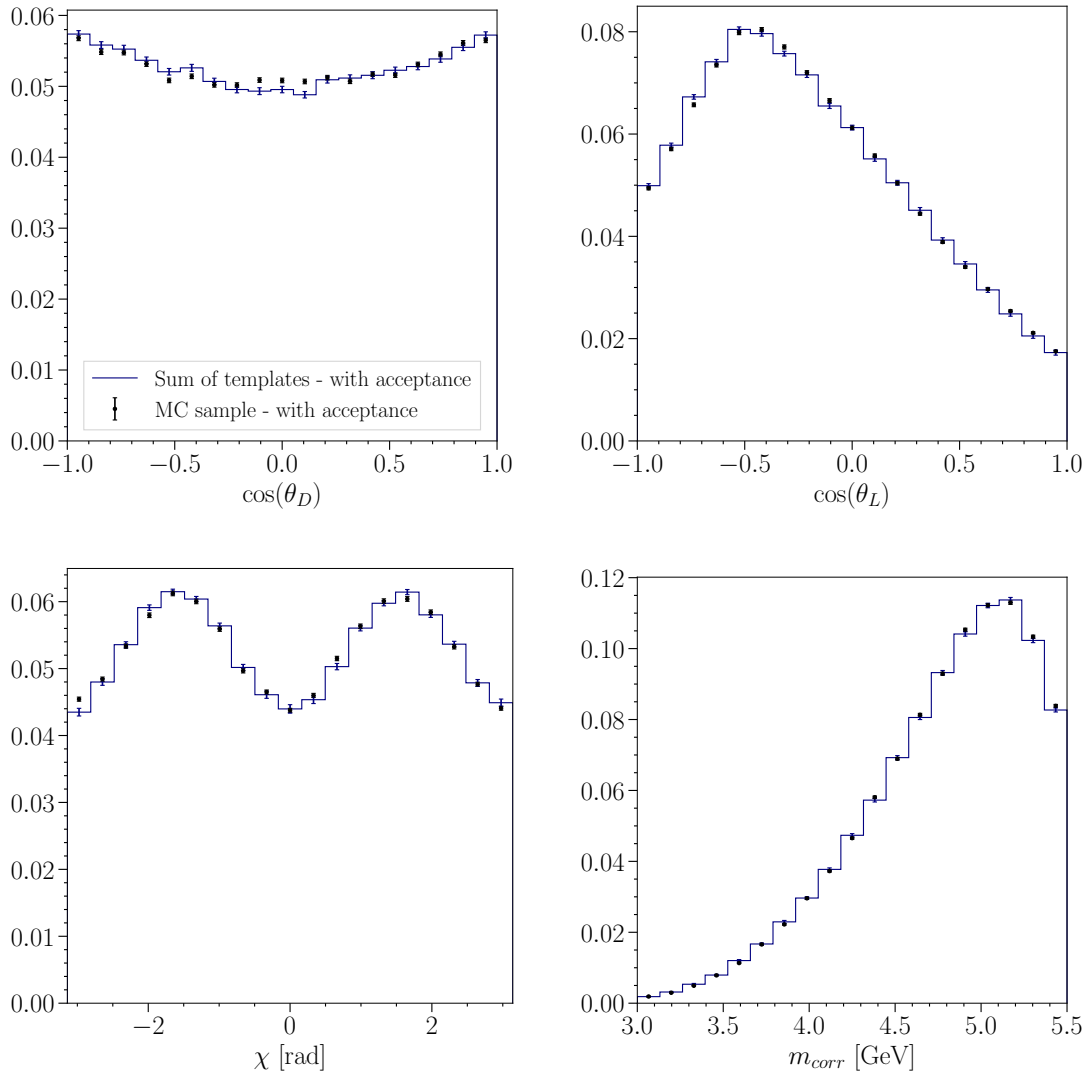


Figure 5.32: The blue line shows one dimensional projections of templates created from the alternative HQET2 model and summed with the true I_X of the default HQET2 model, where the events in the templates have the large acceptance applied. This is compared to the black points, which is simply the RapidSim MC generated with the original HQET2 model, with the large acceptance function applied.

5.7 Template fit results and validation

This section presents the results of the template fit. Firstly, a signal-only template fit is performed using a simulation sample to test the reliability of the method. Additionally, a toy study is performed to test the polynomial correction procedure (outlined in Sec. 4.3.4). Then the full results from the template fit to data are shown. The validation of the template fit with a toy study is shown, as well

as several other cross-checks.

5.7.1 Validation with simulated data

Before applying the four-dimensional angular fit procedure to the LHCb data sample, the reliability of the fit results was assessed by performing the fit on many samples of pseudodata. The signal MC only fit was performed using the 3D PDF in Eq. 4.18.

The agreement between the parametric fit and the template fit was investigated to ensure the template fit can recover the I_X coefficients measured by the truth-level fit, and that it is not biased by reconstruction or acceptance effects. Signal MC events are used to create the 12 templates, and to create the fitted data sample. Clearly, the same signal MC events should not be used to create the template sample and toy data sample as the two samples must be statistically independent. Therefore the full MC dataset was randomly sampled for each toy fit, with 80% of the events used to make the templates, and the remaining 20% used for the fitted data. These fractions were chosen to ensure the template statistics were sufficiently higher than those in the toy data, as in the real template fit to data.

The results for all 12 coefficients are shown in Figure 5.33. The distribution of template fit results is centred around the parametric fit results for all the coefficients. This indicates that the template fit is not biased, and the weights used in the template creation can adequately describe the event migration from the true to reconstructed angular space.

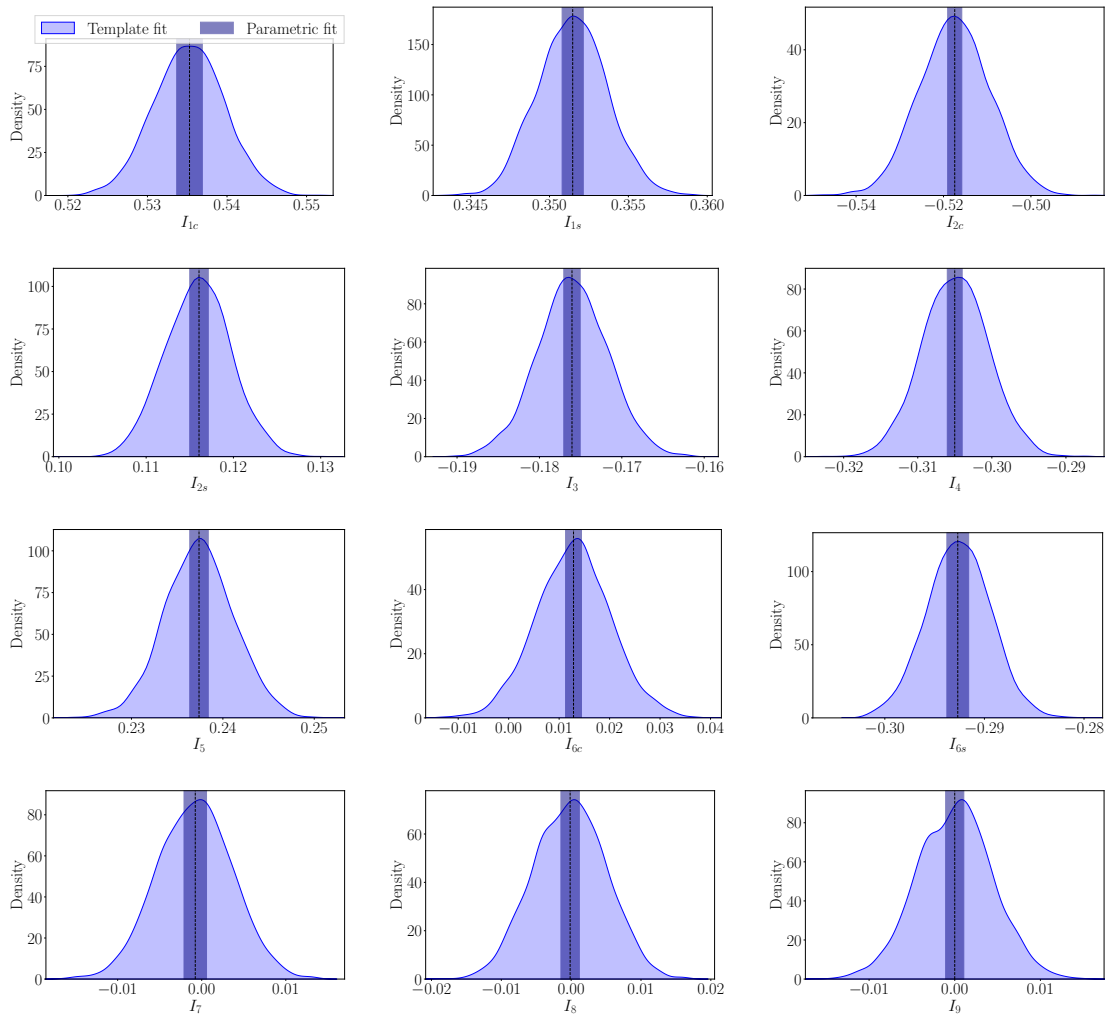


Figure 5.33: Comparison of template and parametric toy fits for the 12 I_X coefficients. The vertical band represents the $\pm 1\sigma$ region for the parametric fit results. The template fit results are shown by a kernel density estimate plot, made using the seaborn package [110].

5.7.2 Validation of polynomial corrections

It is also important to study the reliability of the polynomial correction shapes that are obtained in the angular fit. To do this, a toy data sample was created from the total PDF, and scaled to match the yield of the default data sample. This PDF was then multiplied by a polynomial histogram, created from two arbitrary one-dimensional polynomial shapes in $\cos\theta_L$ and χ. 1000 toy fits were run, where the toy data sample was Poisson varied for each fit.

The distribution of fitted polynomial shapes can be compared to the shape that was applied to the toy data to check if they are consistent, as shown in Fig. 5.34.

Alternatively, the distribution of the fitted polynomial parameters can be plotted and compared to their initial values, as in Figs. 5.35 and 5.36. Note that although the fitted shapes are determined by the fitted polynomial parameters, the effect of the parameters on the resulting shape is not obvious by just looking at the values for the polynomial parameters. It is possible to obtain results where combinations of parameters can be shifted by large amounts, but these effects cancel out, leaving the polynomial shape almost unchanged. For this reason, it is preferable to look at the polynomial shapes rather than just the polynomial parameters, especially as these parameters have no physical significance.

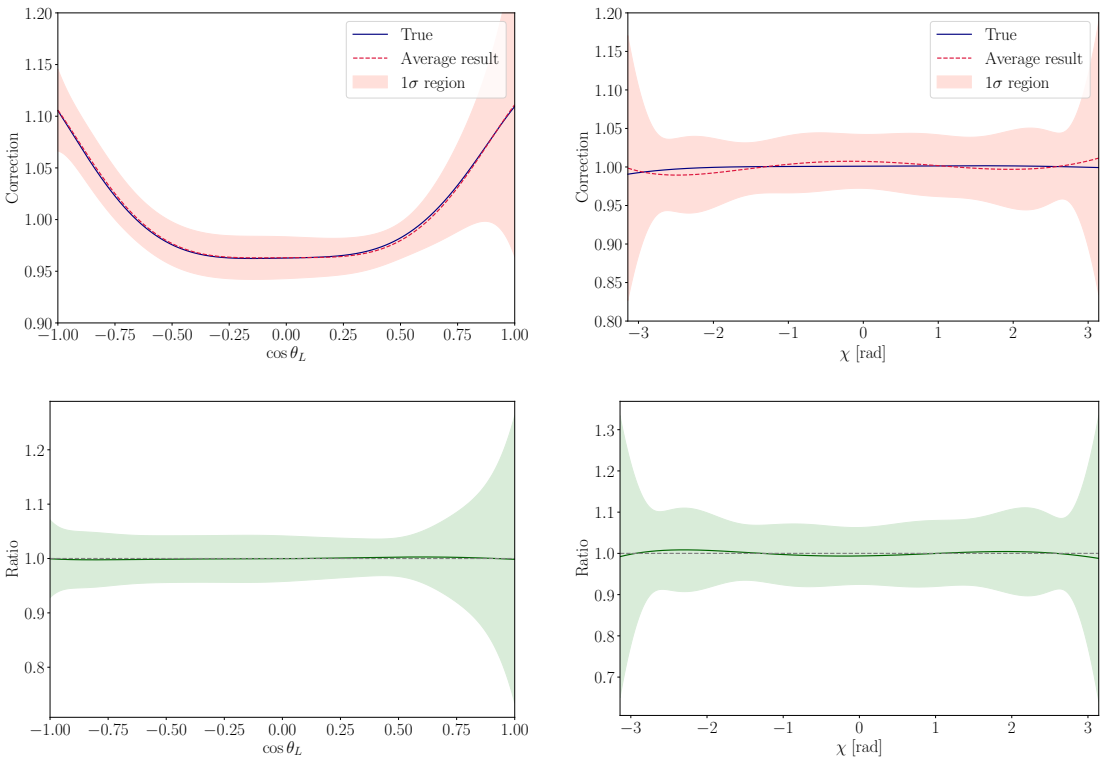


Figure 5.34: Fitted polynomial shapes from the toy fits. The upper plots show the average shape of the measured polynomial (red), compared with the shape applied to the toy data (blue). The lower plots show the ratio between the red and blue lines

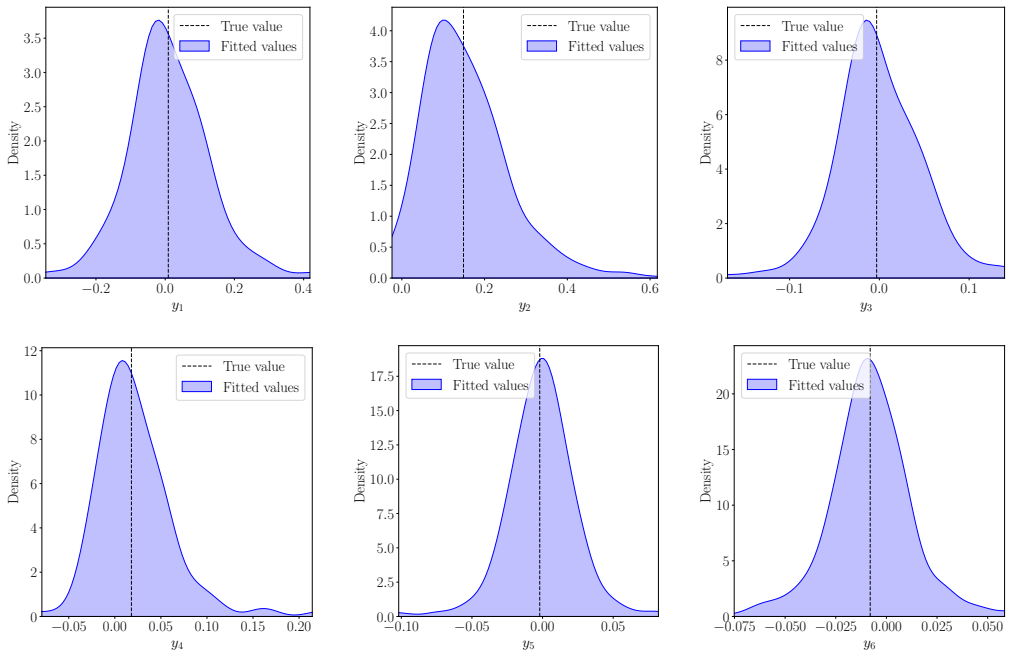


Figure 5.35: Fitted $\cos\theta_L$ polynomial parameters

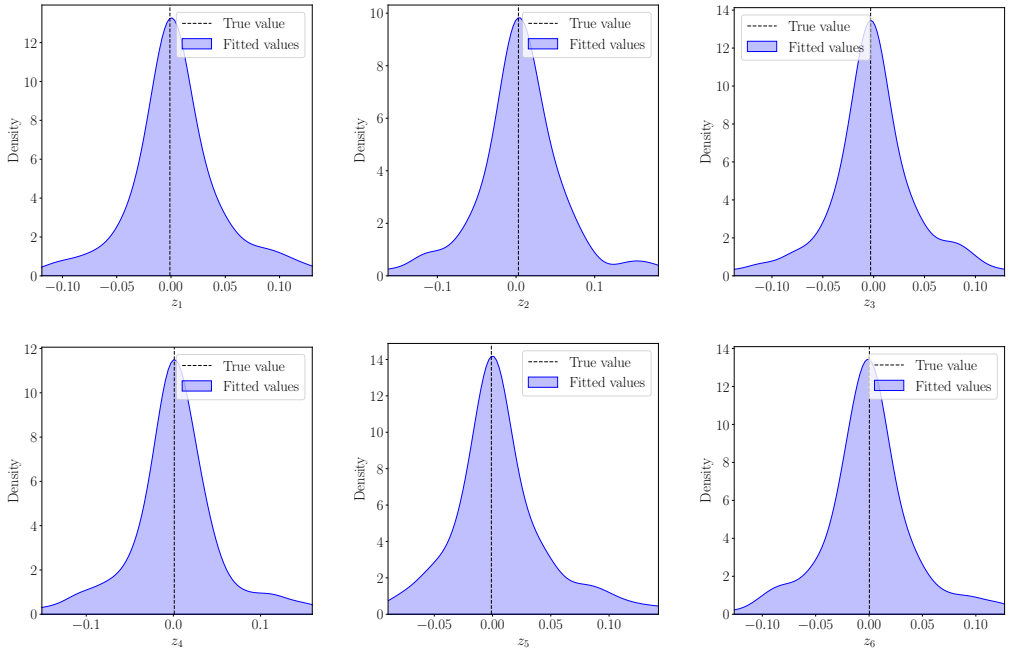


Figure 5.36: Fitted χ polynomial parameters

5.7.3 Results from template fit to real data

The one dimensional predictions from the four-dimensional template fit to data are shown in Fig. 5.37. The signal has approximately a 90% contribution and the

largest background component is from $B \to D^{**}\mu\nu$ decays. The ratios of data and template histograms confirm that the agreement between the fitted result and the data is excellent across the four-dimensional phase space, apart from a 10–20% discrepancy for the small fraction of events with $m_{corr} < 4600\,\text{MeV}$. The overall χ^2/dof, summed over all the bins, is 1.12.

The measured I_X coefficients are shown in Tab. 5.11. Using Eq. (4.16), I_{1c} is measured as:

$$I_{1c} = 0.539 \pm 0.007$$

Figure 5.38 shows a comparison between the measured values and the SM predictions shown in Tab. 3.1; all measured I_X values are consistent with the theoretical predictions. As mentioned earlier, the I_X coefficients can be used to calculate other physics parameters. In particular, the forward-backward asymmetry and longitudinal D^* polarisation are found to be:

$$\mathcal{A}_{FB} = -0.216 \pm 0.011$$

$$F_L^{D^*} = 0.530 \pm 0.002 \quad ,$$

where the uncertainties are statistical only. This is the most precise measurement of these quantities in a semileptonic B decay.

5.7.4 Cross-checks

To validate the performance and stability of the angular fit, several checks were carried out. Firstly, the template fit was run separately for each of the 2016, 2017 and 2018 data samples. The results from these fits were then compared to the default results to investigate if there are any significant biases in a particular data sample. The results are shown in Fig. 5.39. Overall the agreement with the default results is good for each data year. The agreement can be quantified by calculating an overall χ^2 for the ΔI_X values. The χ^2/dof values obtained are: 0.35 for 2016, 0.46 for 2017, and 0.41 for 2018. This indicates that the I_X results are overall consistent across all data years, within statistical uncertainty.

Parameter	Value	Statistical uncertainty
I_{1s}	0.3507	0.0024
I_{2c}	-0.5005	0.0131
I_{2s}	0.1115	0.0039
I_3	-0.1929	0.0031
I_4	-0.2934	0.0036
I_5	0.2481	0.0035
I_{6c}	0.0233	0.0180
I_{6s}	-0.2991	0.0061
I_7	0.0056	0.0023
I_8	0.0030	0.0025
I_9	-0.0018	0.0028
I_{1c}	0.5394	0.0070
$f_{D^*K^0_S\mu\nu}$	0.0063	0.0017
$f_{D^*\tau\nu}$	-0.0004	0.0017
$f_{D^*\pi\mu\nu}$	0.0471	0.0026
$f_{D^*\pi\pi\mu\nu}$	0.0023	0.0016
f_{sig}	0.9038	0.0033
y_1	-0.0161	0.0706
y_2	0.0283	0.0597
y_3	-0.0134	0.0469
y_4	0.0025	0.0301
y_5	-0.0038	0.0147
y_6	-0.0049	0.0054
z_1	-0.0044	0.0027
z_2	0.0076	0.0035
z_3	-0.0073	0.0025
z_4	0.0045	0.0032
z_5	-0.0037	0.0024
z_6	0.0017	0.0027

Table 5.11: Fit results from template fit to real data. The I_X values, background fractions, and polynomial parameters are shown separately. y_i and z_i are the polynomial parameters for $\cos\theta_L$ and χ, respectively.

Furthermore, the angular fit was performed separately for Magnet Up and Magnet Down data samples. The results are shown in Fig. 5.40. The overall χ^2/dof between the Mag Up and Mag Down results is 0.74. This indicates that

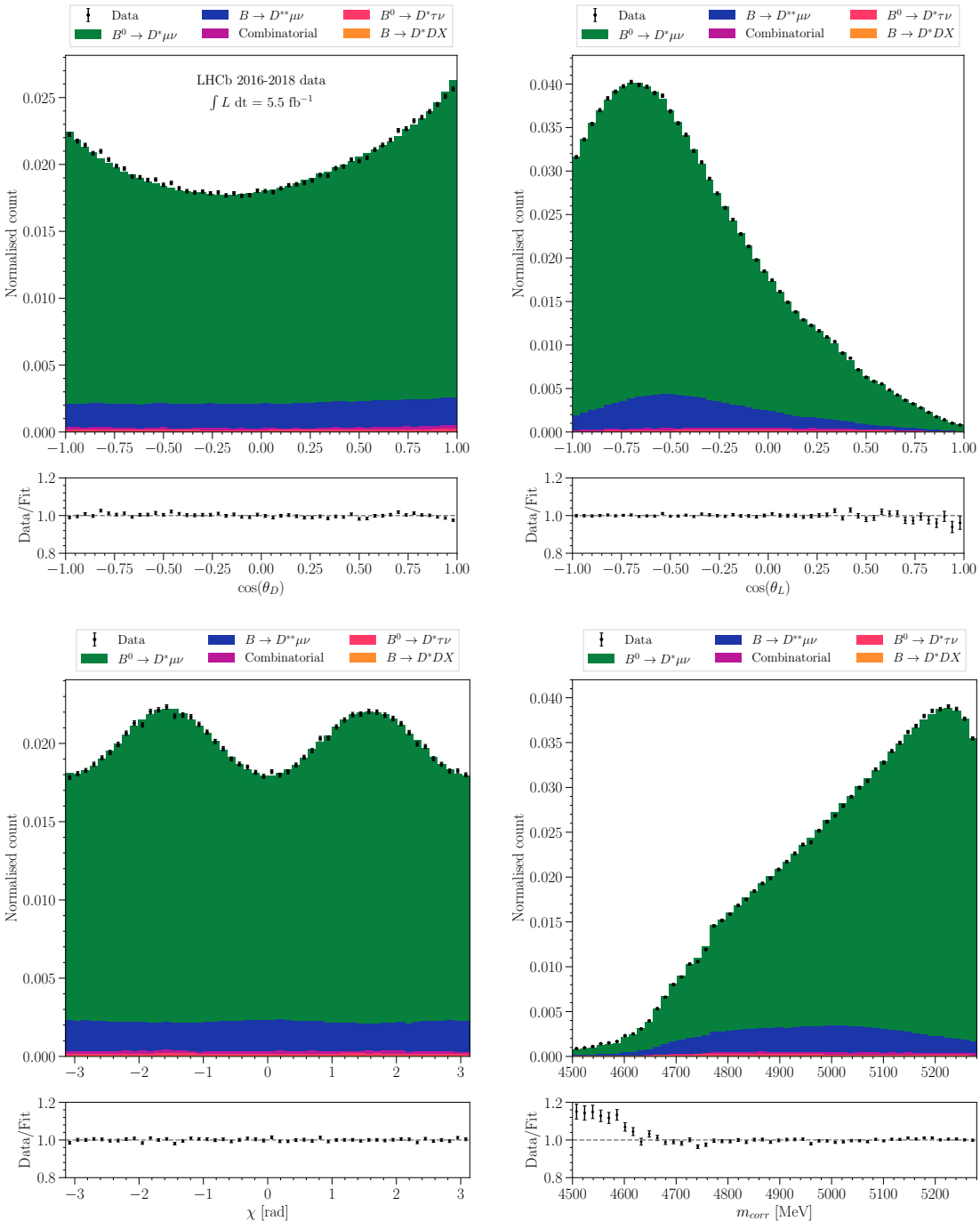


Figure 5.37: One-dimensional histograms constructed using fit results. Note that the binning scheme here is not the same as the binning scheme used for the fit templates

the Mag Up and Mag Down results are consistent, so this cross-check shows no signs of significant biases in the fit.

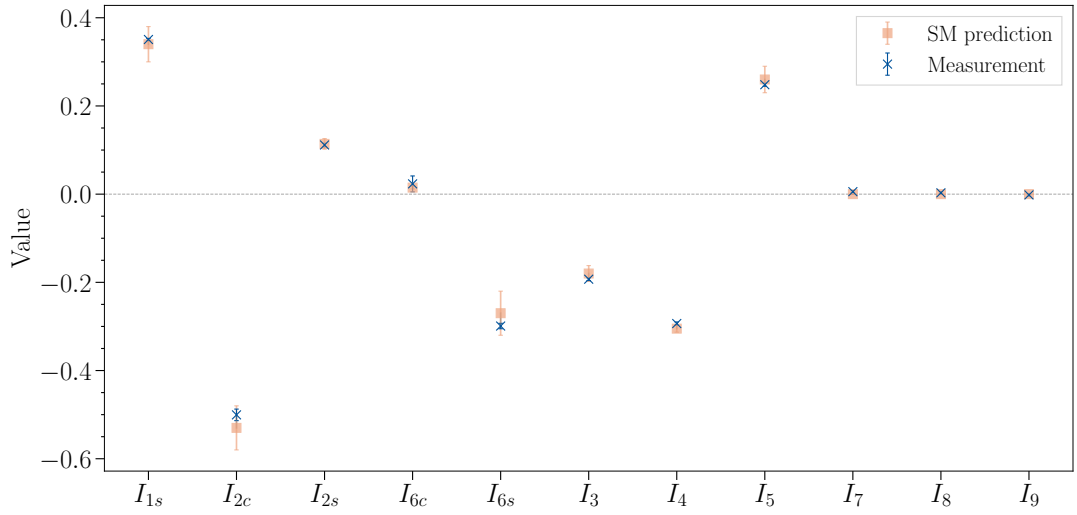


Figure 5.38: Comparison of the SM predictions and measured values for the I_X. The uncertainties associated with the measured values are statistical only.

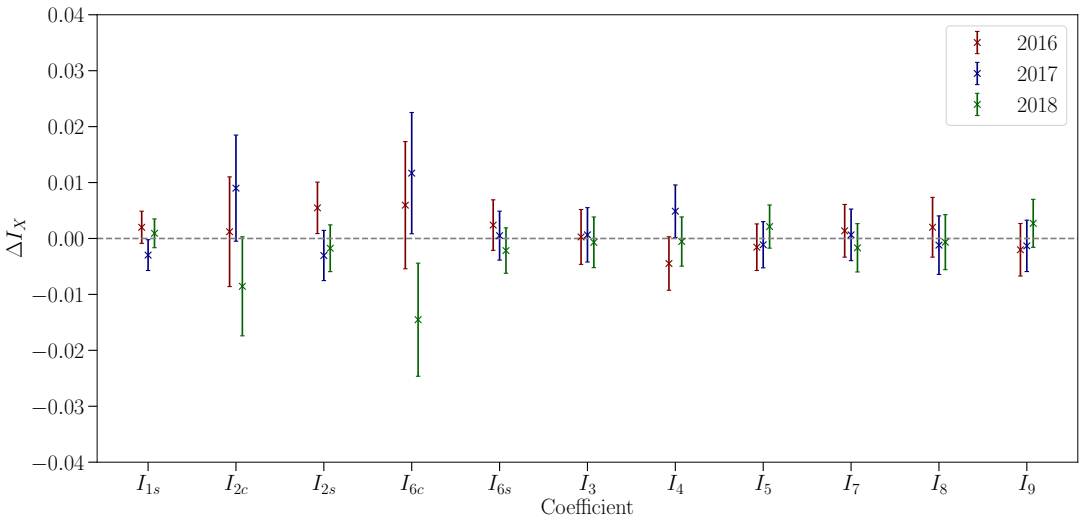


Figure 5.39: Difference in I_X results when the fit is run for a single data year with respect to the default fit.

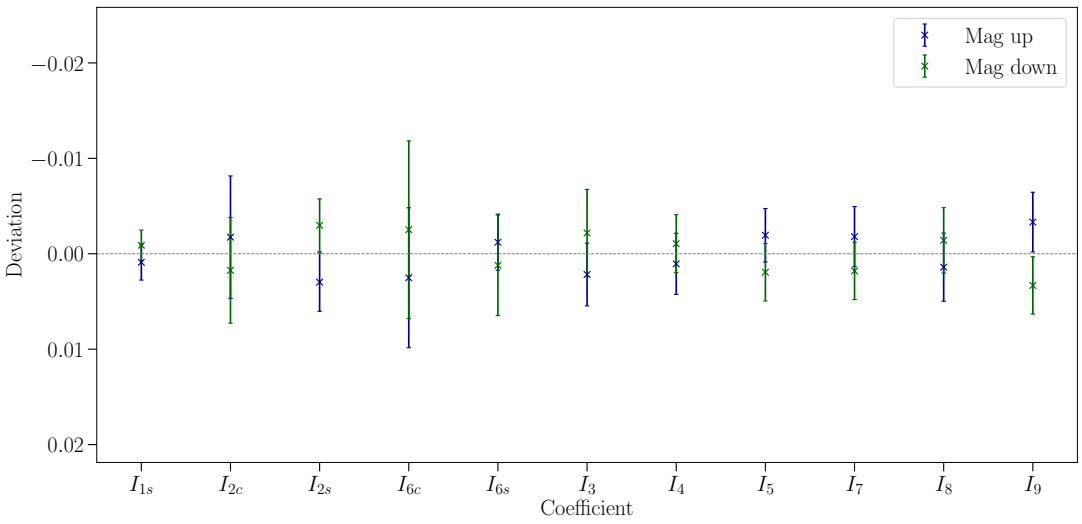


Figure 5.40: Angular fit results split by magnet polarity. For each I_X, the value relative to the averaged result for Mag Up and Mag Down is shown.

5.8 Systematic uncertainties

The results from this analysis have associated systematic uncertainties arising from various aspects of the method. In the following section, the main sources of systematic uncertainty are presented along with details of their evaluation. The results are summarised in Tab. 5.12.

5.8.1 Template statistics

The uncertainty on the fit parameters from finite template statistics is estimated using bootstrapping. For each template, 1000 fits were performed. The template under study was created by randomly sampling from the full background sample, allowing individual events to be sampled multiple times. All other templates were created in the same way as in the default fit. The fits were carried out on the full data sample. The distributions of results for the fit parameters were fitted with gaussian PDFs. The fitted width of a single gaussian shape was taken as the systematic uncertainty for a particular parameter, from a particular template. This systematic uncertainty is relatively small.

5.8.2 Fixed components in angular fit

Combinatorial fraction

The fraction of the combinatorial background (f_{comb}) is fixed in the angular fit. Its value is calculated by comparing the size of the wrong-sign data sample to the right-sign, after the full selection is applied. Then this value is corrected by taking the average ratio of the wrong-sign to right-sign in the high m_{corr} sideband.

A systematic uncertainty was estimated by first estimating the statistical uncertainty on the combinatorial fraction used in the angular fit. This was done by combining the statistical uncertainty for the wrong-sign yield (0.0099) with the uncertainty of the weighted average used as the correction factor (0.1097). This gives an overall uncertainty of $\sqrt{0.0099^2 + 0.1097^2} = 0.110$.

Then, the angular fit was run on the full data sample many times, with a fixed f_{comb}. For each fit, the value of f_{comb} was chosen by randomly sampling

from a gaussian distribution, where the mean of this distribution was set to the value of f_{comb} used in the real angular fit, and the width of the distribution was taken as the uncertainty on f_{comb} calculated above. The widths of the parameters from these fits were taken as the systematic uncertainty from fixing f_{comb}. This is a relatively small uncertainty.

$B \to D^{**}\mu\nu$ fractions

A similar process was done for the fixed internal fractions of the $B \to D^{**}\mu\nu$ template. For each $B \to D^{**}\mu\nu$ component, the angular fit was run 1000 times, with the fraction chosen by randomly sampling from a gaussian centered at its nominal value, and a width taken from its uncertainty according to the latest values from the PDG [75]. The fitted widths of the distributions of the parameters from these fits are taken as the systematic uncertainty. This is a dominant systematic uncertainty.

5.8.3 Tracker-only MC corrections

To align the tracker-only simulation with the full simulation, and account for the trigger selection, spline functions were used to reweight the tracker-only samples when creating the angular fit templates, as described in Section 5.4.1. Two systematic uncertainties were estimated from this method, one arising from the finite sample sizes of the MC samples used to create the ratio histograms, and the other from the choice of model for the spline function.

Finite sample sizes

This uncertainty was estimated by bootstrapping the fullsim and tracker-only MC samples 1000 times. Ratio histograms were created for each of the 1000 samples, and cubic spline functions were fitted to these. This yields 1000 spline functions, so each event has a distribution of 1000 weights. The standard deviation of these weights was calculated for each event.

Then the angular fit was performed 1000 times, and each time the correction weights applied to the MC templates were randomly fluctuated by sampling from a gaussian with a mean set to the default weight value, and the width set to

the standard deviation obtained from the 1000 spline functions. The width of the distributions of the fit parameters from these fits is taken as the systematic uncertainty. This is one of the smallest systematic uncertainties.

Spline model

Another systematic uncertainty was estimated from the choice of a cubic model for the spline functions. This was done by running the tracker simulation correction process using 4th and 5th order polynomials instead of a cubic function. For each event, the difference in correction weight from the cubic spline and alternative model was taken as an uncertainty. This uncertainty was used as the width of a gaussian distribution for sampling the correction weights in the angular fit, which was run 1000 times, similarly to the previous study. Again, the width of the distributions of the fit parameters was used to assign the systematic uncertainty for the model choice. This is also a very small contribution.

5.8.4 Model independence of templates

Although the weight procedure for the angular templates involves dividing by the total model used in the simulation, there will still be some small residual model dependence remaining in the templates. This model dependence is due to two main factors.

Uncertainty in weight calculation

Firstly, the total model used in the weight calculation is defined using the results of the parametric fit to the truth-level RapidSim sample. These results have associated uncertainties, so the total model used for the weights may not be exactly the same as that used in the simulation. Therefore, when the I_X results are used in the weight calculation, they may not exactly cancel the model used in the simulation leading to small biases. This effect is minimised by using a large MC sample for the RapidSim fit, to try and reduce the statistical uncertainties on the fitted I_X values.

A systematic uncertainty is assigned by randomly varying the truth-level I_X values obtained from the RapidSim fit within their uncertainties 1000 times, then

creating 1000 sets of signal templates using these fluctuated I_X values in the weight calculation. These signal templates were then used in the template fit to data, giving a distribution of 1000 values for each parameter. The width of these distributions is taken as the systematic uncertainty. This systematic uncertainty is fairly significant.

Correlation between m_{corr} and q^2

Secondly, the template fit uses m_{corr} as a variable, which is correlated with q^2. All the q^2 dependence in Eq. 3.46 is contained in I_X coefficients. We create the signal templates to represent the angular functions in Eq. 3.46, which have no q^2 dependence. However, the four-dimensional templates contain some q^2 dependence indirectly, through the dependence on m_{corr}. Therefore, the model dependence does not cancel out completely in m_{corr}.

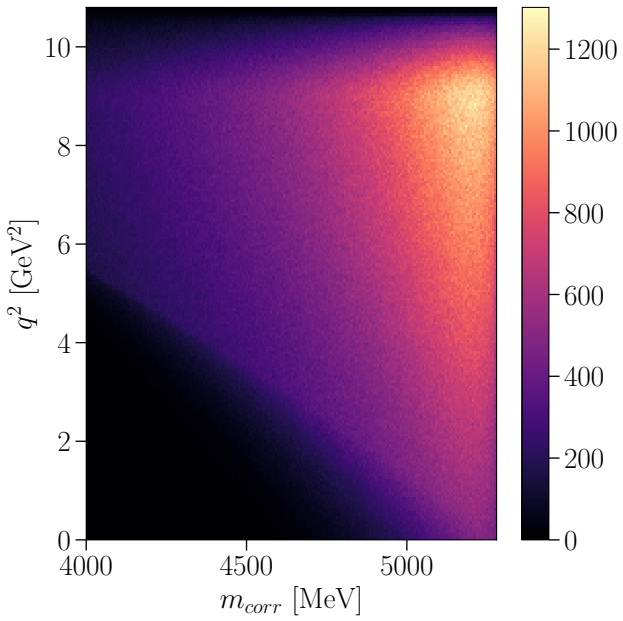


Figure 5.41: 2d histogram showing weak correlation between q^2 and m_{corr}; lighter colours indicate a higher density of events. The correlation coefficient is -0.272

To assign a systematic uncertainty for this, a RapidSim sample with an alternative decay model was created by varying the form factors in EvtGen by 20% of their default values. Then the m_{corr} shape of this sample was compared to a RapidSim sample created with the default EvtGen model. A ratio histogram was

created by dividing the m_{corr} yields of the two samples, which was fitted with a cubic spline. The value of the resulting spline function at a given point in m_{corr} can be taken as a correction weight for an event with that m_{corr} value.

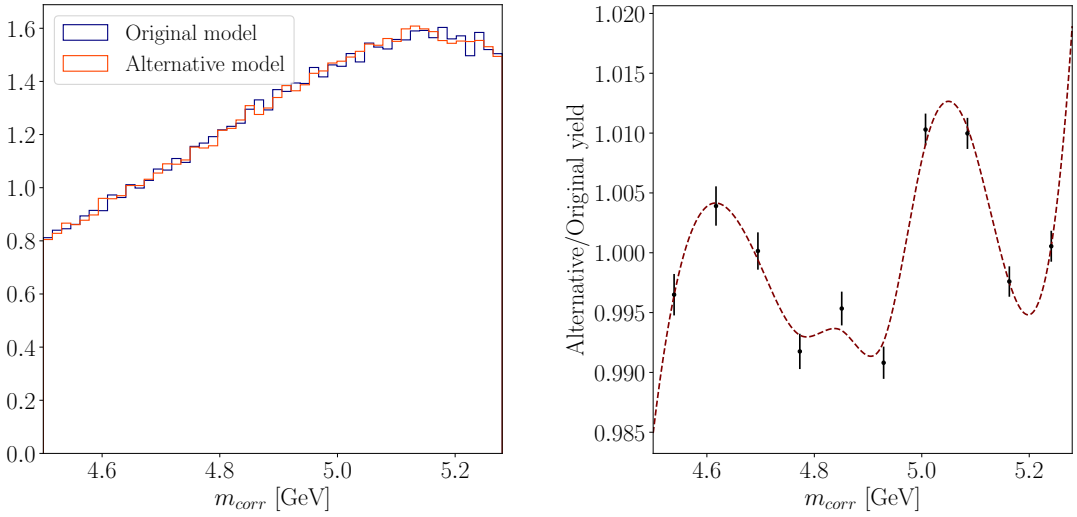


Figure 5.42: m_{corr} shapes from RapidSim for the default and alternative models (left), and the ratio histogram fitted with a cubic spline function (right)

These correction weights were then applied to the 12 signal templates in the four-dimensional angular fit to data. The difference between the fitted results for the parameters in this fit and the default angular fit is taken as the systematic uncertainty. The size of this uncertainty varies a lot across the I_X.

5.8.5 Fake D^0, fake D^* fit

The 2d fit to the D^0 mass and ΔM is used to calculate a set of sWeights, which are then applied to the angular fit, in order to subtract fake D^0 and fake D^* backgrounds.

The systematic uncertainty arising from applying the sWeights was estimated by first bootstrapping the data in the D^0 mass and ΔM fit 1000 times. The results from each of the bootstrapped data fits were used to calculate a set of sWeights on the original data sample. These alternative sWeights were then applied to the angular fit. The standard deviation of the distribution of angular fit results from these 1000 fits was taken as the systematic uncertainty. This is the smallest systematic uncertainty considered here.

5.8.6 D^{**} model uncertainty

In the angular fit, unlike the 12 signal templates, events in the background templates are not assigned weights to cancel out the decay model used in the simulation. Therefore, the fit results assume that the background models are correct, and there is a systematic uncertainty associated with this assumption. The background from $B \to D^{**}\mu\nu$ decays is by far the largest, so the systematic uncertainty for the decay models in the other background modes is negligible.

To estimate the systematic uncertainty, the HAMMER package [111] was used to randomly vary the Wilson coefficients in the $B \to D^{**}\mu\nu$ samples. This process was done 200 times, yielding 200 sets of HAMMER weights. For each iteration, each Wilson coefficient was allowed to vary by 20%. The conservative level of variation was allowed because there is currently so little information available about the angular structure of the $B \to D^{**}\mu\nu$ decays. The HAMMER weights were applied to the $B \to D^{**}\mu\nu$ templates and the angular fit to data was performed for each of the 200 weights. The effect of these weights on the fit variables is shown in Fig. 5.43.

The width of the distribution of fitted results from these 200 angular fits was taken as the systematic uncertainty from the background model assumption. This is a dominant systematic uncertainty; it is the largest contribution for several I_X.

5.8.7 $B^0 \to D^{*-}\pi^+$ corrections

For the control study with $B^0 \to D^{*-}\pi^+$ decays, discussed in Section 5.4.2, the uncertainty on the spline function used for the correction weight arising from finite MC statistics was estimated by bootstrapping the $B^0 \to D^{*-}\pi^+$ sample 1000 times and calculating a spline function for each sample. The distribution of spline functions is shown in Fig. 5.44.

Then the angular fit was performed with the default data sample 1000 times. Each time a correction weight was applied to the MC templates using one of the spline functions. The distributions of fit parameters were fitted with a gaussian function. The fitted width of the gaussian function was taken as the systematic

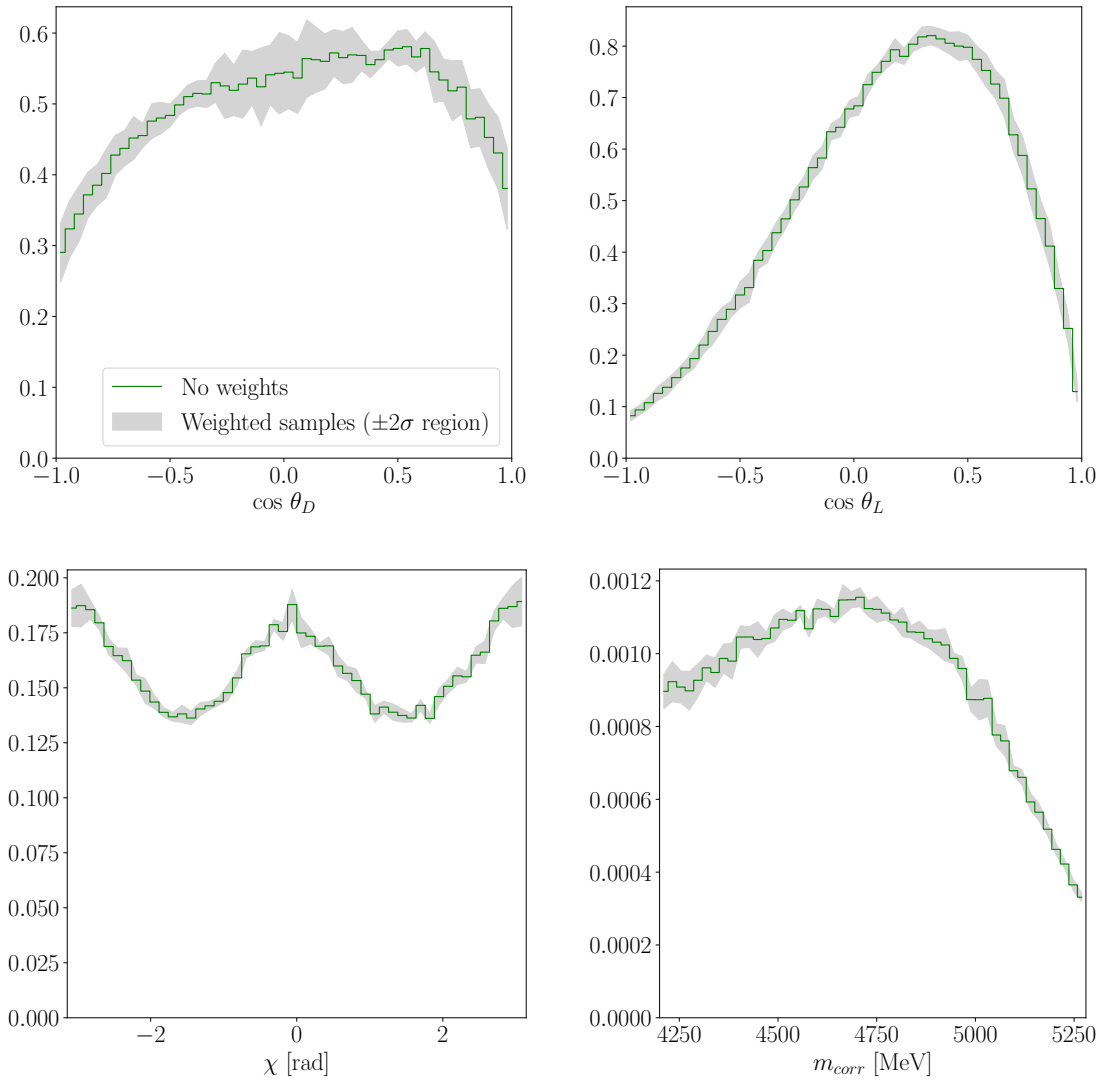


Figure 5.43: $B \to D^{**}\mu\nu$ MC sample without HAMMER weight is shown in green. The grey band shows the 2σ region for the MC sample with the HAMMER weights applied

uncertainty for each of the parameters. This is a large contribution to the overall systematic uncertainty.

5.8.8 Binning scheme in template fit

The binning used for the template fit is defined in Tab. 4.4. Alternative binning schemes were tested. In one case, the number of bins in each variable was decreased by two, and in another case it was increased by two. For each of the binning schemes, 1000 fits were run, where the data sample was bootstrapped for each

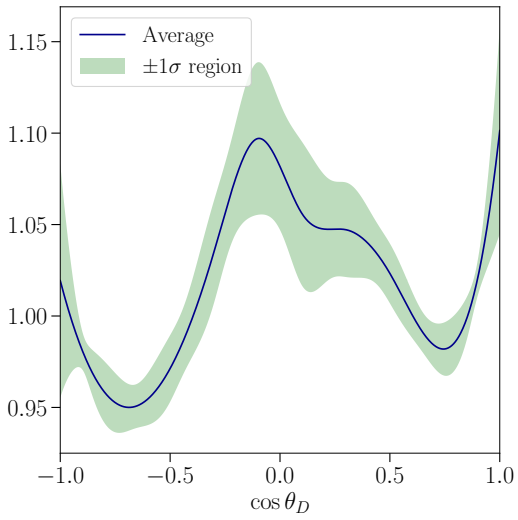


Figure 5.44: Distribution of spline functions resulting from bootstrapping $B^0 \to D^{*-}\pi^+$ MC sample 1000 times. The blue line is the default spline function and the green band shows the $\pm 1\sigma$ region of the distribution

fit. The distributions of parameters were each fitted with gaussian functions. The difference between the means of these parameters for the two alternative binning schemes was taken as the systematic uncertainty from the chosen binning scheme in the angular fit. This systematic uncertainty is relatively small.

5.8.9 Residual muon misID background

The data selection removes the vast majority of background from particles misidentified as muons. However, it is possible that a small fraction of these events remain after all selection is applied. The effect of this can be studied using the $B^0 \to D^{*-}\pi^+$ control sample used for simulation corrections in Sec. 5.4.2, since this sample is selected using an identical preselection to the signal, apart from the PID requirement on the muon.

To estimate the fraction of candidates with misidentified muons in the data sample, the $B^0 \to D^{*-}\pi^+$ sample was used to estimate the offline selection efficiency of these candidates. The misidentified muon candidates can be assumed to behave in the same way as those in the $B^0 \to D^{*-}\pi^+$ sample. So the full offline selection is applied to that sample, and the measured efficiency is taken as the selection efficiency

of the misidentified muons candidates. This was combined with the efficiency of the `isMuon` algorithm [112]. The `isMuon` efficiency is measured in bins of muon p_T, so the weighted average of the efficiency can be calculated over the measured muon p_T spectrum. This is then multiplied by the offline selection efficiency to estimate the fraction of misidentified muons, yielding an estimate of maximum 0.5%.

The $B^0 \to D^{*-}\pi^+$ sample was then used to create a template for the angular fit. This "fake muon" component was added to the template fit, with a fixed fraction of 0.5%. This fit was run 1000 times with bootstrapping, and the same was done for the default angular fit (without the fake muon template). The distributions of I_X results were fitted with a gaussian. For each I_X the difference in means of the gaussians for the default (no fake muon template) and alternative (with fake muon template) results was taken as the systematic uncertainty for residual background from misidentified muons. For most of the I_X this is a relatively small contribution.

5.8.10 L0 trigger variation

To emulate the trigger response in the tracker-only simulation used for the fit templates, the 2016 fullsim MC samples are used, as outlined in Sec. 5.4.1. These corrections are applied to samples for all data-taking years. However, due to changes in L0 trigger thresholds, the trigger efficiencies can vary by year. Therefore, a systematic uncertainty is assigned to account for these differences.

The variation in trigger response is studied by comparing hadron p_T and multiplicity variables in data. For each data-taking year, a control sample was selected by applying the full selection apart from the BDT cuts, and requiring $m^2_{miss} < 0.5$ GeV^2 to reduce the impact of backgrounds. Then the shape of the 2016 sample was compared to the combined 2017+2018 sample in the variables: $(p_T(\pi), p_T(K), p_T(D^0), \texttt{nTracks}, \texttt{nSPDHits})$.

Ratio histograms were made in each variable by dividing the yield of the 2017+2018 sample by the yield of the 2016 sample. The histograms were fitted with cubic spline functions. Each event can be assigned a weight equal to the spline value at that point. The process is carried out sequentially, and the final

set of per-event weights is calculated by taking the product of the individual sets of weights for each variable.

These weights were applied to the 2017+2018 MC samples used to create the angular fit templates. Then the fit was run using these alternative templates, and results were compared to the default fit results. The difference between the results is taken as a systematic uncertainty. This is a relatively large systematic uncertainty.

5.8.11 Total systematic uncertainties

The systematic uncertainty for each parameter is shown in Tab. 5.12. The total systematic uncertainty is calculated by summing each contribution in quadrature. The largest sources of systematic uncertainty are the model uncertainty for the $B \to D^{**}\mu\nu$ template; the $B \to D^*\pi$ control study; the fixed $B \to D^{**}\mu\nu$ fractions; and the L0 trigger variation. For future iterations of this analysis, further study of $B \to D^{**}\mu\nu$ could potentially reduce two of the dominant systematic uncertainties.

I_7, I_8, and I_9 have smaller systematic uncertainties than the other coefficients. The dominant systematic uncertainties for these coefficients are the uncertainty on the truth level I_X and the template fit binning scheme.

Source	I_{1s}	I_{2c}	I_{2s}	I_{6c}	I_{6s}	I_3
Fixed combinatorial component	0.00017	0.00073	0.00015	0.00082	0.00075	0.00036
Fixed $B \to D^{**}\mu\nu$ fractions	0.00157	0.00391	0.00342	0.00309	0.00217	0.00197
Spline model for fullsim/tracker-only corrections	0.00005	0.00014	0.00017	0.00015	0.00012	0.00010
Stats for fullsim/tracker-only corrections	0.00005	0.00012	0.00009	0.00012	0.00005	0.00007
Uncertainty for truth-level I_X	0.00077	0.00196	0.00222	0.00228	0.00107	0.00107
sWeights from $m(D^0), \Delta M$ fit	0.00001	0.00003	0.00004	0.00005	0.00002	0.00005
$B \to D^{**}\mu\nu$ decay model	0.00308	0.00528	0.00836	0.00613	0.00493	0.00229
m_{corr} and q^2 correlation	0.00091	0.00582	0.00380	0.00503	0.00175	0.00030
$B \to D^*\pi$ control study	0.00413	0.00744	0.00229	0.00377	0.00010	0.00265
Binning scheme in angular fit	0.00133	0.00300	0.00081	0.00219	0.00164	0.00140
Residual muon misID background	0.00066	0.00234	0.00104	0.00373	0.00074	0.00029
L0 trigger variation	0.00227	0.00544	0.00002	0.00697	0.00250	0.00194
Template statistics	0.00021	0.00077	0.00039	0.00094	0.00039	0.00049
Total	0.00615	0.01347	0.01040	0.01268	0.00659	0.00485

Source	I_4	I_5	I_7	I_8	I_9
Fixed combinatorial component	0.00043	0.00038	0.00001	0.00003	0.00008
Fixed $B \to D^{**}\mu\nu$ fractions	0.00168	0.00169	0.00047	0.00028	0.00034
Spline model for fullsim/tracker-only corrections	0.00018	0.00018	0.00005	0.00005	0.00006
Stats for fullsim/tracker-only corrections	0.00006	0.00005	0.00006	0.00006	0.00006
Uncertainty for truth-level I_X	0.00149	0.00079	0.00085	0.00080	0.00064
sWeights from $m(D^0), \Delta M$ fit	0.00002	0.00001	0.00006	0.00001	0.00002
$B \to D^{**}\mu\nu$ decay model	0.00326	0.00174	0.00023	0.00023	0.00034
m_{corr} and q^2 correlation	0.00073	0.00063	0.00000	0.00002	0.00002
$B \to D^*\pi$ control study	0.00033	0.00203	0.00004	0.00003	0.00003
Binning scheme in angular fit	0.00080	0.00076	0.00132	0.00008	0.00178
Residual muon misID background	0.00090	0.00001	0.00020	0.00003	0.00045
L0 trigger variation	0.00227	0.00631	0.00022	0.00016	0.00027
Template statistics	0.00043	0.00041	0.00037	0.00040	0.00040
Total	0.00483	0.00719	0.00172	0.00098	0.00207

Table 5.12: Total systematic uncertainties

6

VELO online monitoring and tomography

6.1 VELO upgrade

As part of the extensive upgrade to the LHCb experiment between 2019 and 2022, the VELO used in Run 1 and 2 was entirely replaced by a brand new detector. The new VELO is designed to operate at a luminosity of 2×10^{33} cm^{-2} s^{-1}. Like the original detector, the new VELO consists of two halves, placed in vacuum inside the LHC beam pipe, either side of the proton-proton collision region. The two sides are conventionally labelled as the A (*Ascenseur*) and C (*Cryogénique*) sides, referring to their positions in the LHCb underground cavern. Each half contains 26 modules. The modules are separated by 25 mm in the central region and more widely separated near the ends [113]. The layout is shown in Fig. 6.1.

A VELO module is made of two silicon pixel sensors mounted on each side. The sensors are 200 µm thin and have a pixel pitch of 55 µm. Each sensor is bump-bonded to three ASICs [115] based on TimePix3 technology [116], which are responsible for the sensor readout. The ASICs have 256×256 pixels with a pixel size of 55×55 µm. They operate at a readout rate of up to 40 MHz (compared to 1.1 MHz for the old detector), and a bandwidth of up to 20.4 Gb/s.

For silicon biasing, high voltage (HV) is delivered to the sensors through thin HV tapes, which can be operated up to maximum of 1000 V. The sensors are

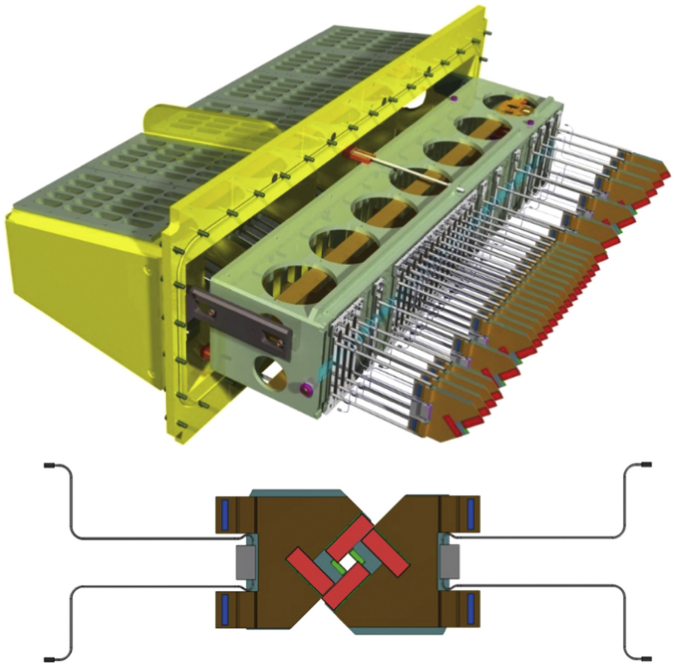


Figure 6.1: Top: Layout of one half of the upgrade VELO. The 26 modules are aligned along the beam direction. Bottom: A pair of modules shown in the closed position used for data taking. The sensors are shown in red. The modules are in different z-positions along the beam line to minimise gaps in the acceptance. [114]

mounted onto a custom-made silicon substrate with integrated microchannels. The microchannels are designed to run directly underneath the primary heat sources of the module. The sensors must be cooled and kept below -20°C to avoid thermal runaway induced by leakage currents from radiation damage. The mechanism employed for this is two-phase CO_2 cooling. CO_2 at -30°C is passed through the microchannels. The microchannels are shaped such that the region near the inlet pipes has a higher pressure, which is then followed by an area where the channels are widened, reducing the pressure. This causes the CO_2 to boil, thereby cooling the module surface above. The layout of a VELO module is shown in Fig. 6.2.

The VELO modules operate in a "secondary vacuum" that is weaker than the "primary vacuum" inside the LHC beam pipe. The vacua are separated by a thin aluminium box, one enclosing each half, commonly referred to as the RF foil. A picture of the RF foil is shown in Fig. 6.3. They are designed to withstand a pressure difference of up to 10 mbar. The RF foil also protects the electronics from currents induced by the LHC beam. The foils were each constructed

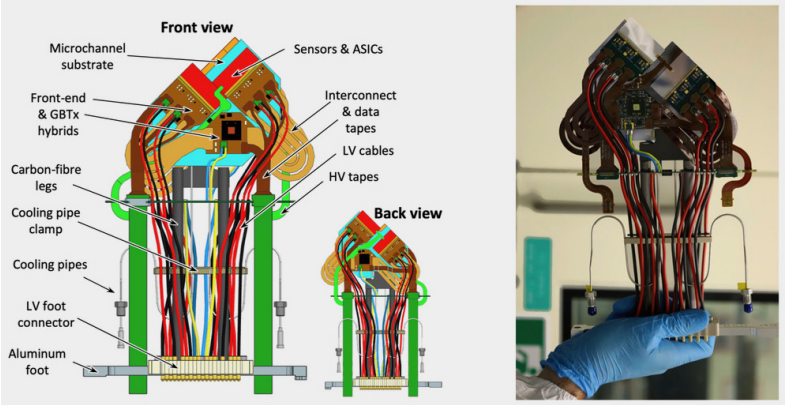


Figure 6.2: Left: Diagram of a single VELO module, with each component labelled. Right: photo of a fully assembled VELO module [117]

from a single block of aluminium weighing over 300 kg. The blocks were firstly machined down to a thickness of 500 µm. Then they were milled and etched to 150 µm at the thinnest part.



Figure 6.3: Photo of the two RF foils. The shape is corrugated, allowing the modules to be placed inside with a clearance of less than 1 mm [118]

Another feature of the upgrade VELO is SMOG2 [119]. This allows LHCb to operate as a fixed-target experiment by injecting gas into the beam pipe. The upgraded system differs from the original SMOG through the use of a storage cell to contain the injected gas. The cell is split into two halves, each mounted to one VELO half. As the VELO closes, the two cell halves also come together to create a closed structure. This enables much higher gas densities and luminosities than the old system.

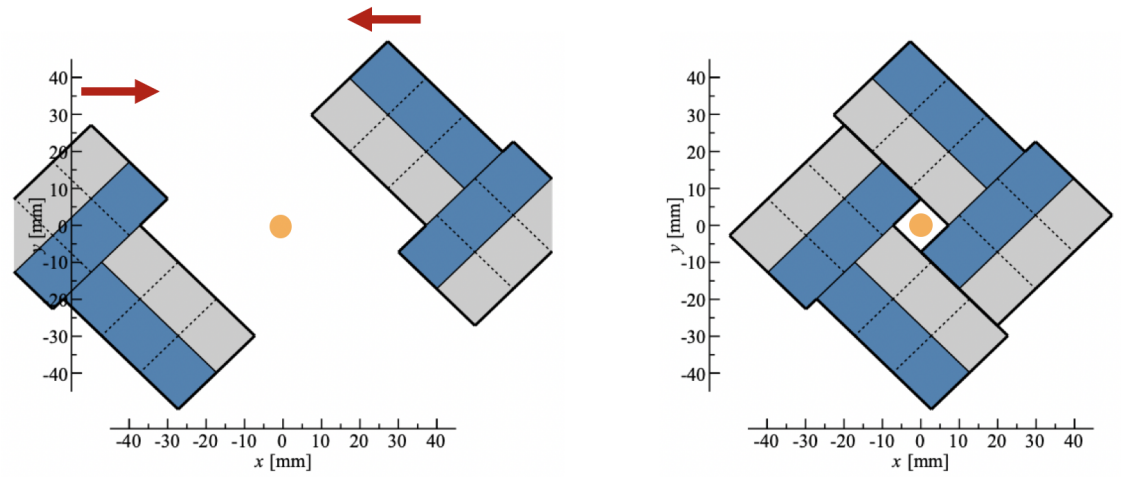


Figure 6.4: Position of a VELO module pair in the x, y plane in the open (left) and closed (right) positions. The yellow dot indicates the beamspot. Based on diagram from Ref. [113]

6.2 VELO closing procedure

During data taking in Run 3, the two halves of the VELO are designed to sit just 3.5 mm from the colliding proton beams. This is smaller than the corresponding distance (5.1 mm) in Run 1 and 2. During beam setup and injection, the position of the LHC beams can fluctuate by a few mm. For this reason, it is unsafe to place the VELO halves at this distance for the entire beam cycle. Furthermore, the final position of the beams during collisions can vary between fills. This means that if the VELO position were permanently fixed, it would not be perfectly centred around the collision region for every fill. Physics data-taking commences only after the proton beams are at their peak energy and they have been narrowed and aligned by the magnets. This condition is referred to as "stable beams". To protect the VELO, before the LHC operators declare stable beams, the two halves are retracted horizontally to a safe distance of 27 mm from the collision region. This will be referred to as the "open position". The open and closed positions are illustrated in Fig. 6.4.

Then, after the stable beams condition is achieved, each VELO half moves towards the collision region in several stages. The VELO motion is facilitated through a set of stepping motors, and the position is measured by resolvers and potentiometers. The two halves move independently in the horizontal (x) direction.

In addition, they can move in the vertical (y) direction, but in this case the entire centre frame is moved, so the two halves must always move together. The movement of the VELO halves is referred to as the *closing procedure*.

As mentioned above, it is not possible *a priori* to determine the exact interaction position of the colliding beams for a particular fill. Therefore, the VELO measures the position of the collision region in x and y, by reconstructing vertices from the pp collisions. The reconstruction is performed independently by each VELO half. In this way, each half can measure the distance between itself and the collision region, and thus determine the distance it is required to move, so that the VELO can be precisely centred around the beams at the desired distance.

During the closing procedure, several safety criteria need to be monitored:

- The trends of the measured vertex positions are used to check that the collision region is not drifting significantly.
- A set of beam position monitors, installed at both ends of LHCb, are used to check if the individual beams are drifting.
- The LHC beam conditions monitor provides a warning of a potential LHC malfunction outside LHCb.
- The HV bias current is used to check if there is an unusual level of activity on the sensors.

In addition to the above, it is useful to monitor the RF foil temperature. If this increases significantly during the closing procedure, it may indicate the detector is dangerously close to the beams.

6.3 Vertex monitoring for VELO closing

The measured PV positions are the most important input for the VELO closing procedure. Each half can provide an independent measurement of the PV position relative to itself, then the target position for both halves is defined as the average of the two measurements.

The vertex measurements are performed using an algorithm that runs in the LHCb online monitoring framework, known as a *monitoring task*. This monitoring task runs independently and in parallel for both VELO halves. PVs are reconstructed using standard LHCb tracking and vertexing algorithms. Crucially, the PV reconstruction does not use information about the global position of each VELO half (the positions from the resolvers and potentiometers are not used). This ensures the measured vertex positions are minimally biased, and reduces the potential impact of any mechanical misalignment. Each half reconstructs vertices using only tracks which have hits on that half, i.e. vertices reconstructed by the A side do not use any tracks that have hits only in the C side[1].

When the vertices are reconstructed, histograms of the measured x and y positions are filled. After 1000 entries, the histograms are fitted with gaussian functions. Some simple checks for fit quality are then performed:

- The fit χ^2/dof is checked.
- The fitted width is compared to the RMS width.
- The fit is performed again with a restricted range, and the results are compared to the original fit.

If the fit quality passes these tests, the fitted parameters (the mean and width of the fitted gaussian function) are published to an LHCb online database [120]. From there, they can be read by the VELO closing algorithm, which monitors the other safety criteria, decides if it is safe to perform the next step of the motion, and calculates the target position. Furthermore, the histograms themselves are published to an online database, and can be viewed in a web-based monitoring application [121], which was newly designed for Run 3, so that the shift crew can check for any issues. During normal data-taking conditions, the entire procedure takes a few minutes. A diagram outlining the procedure is shown in Fig. 6.5.

[1]Note that due to the overlap of the VELO modules in the two halves, a small fraction of tracks produce hits in both halves. These are known as *overlap tracks*. A single overlap track can be used by both halves in their respective PV reconstruction.

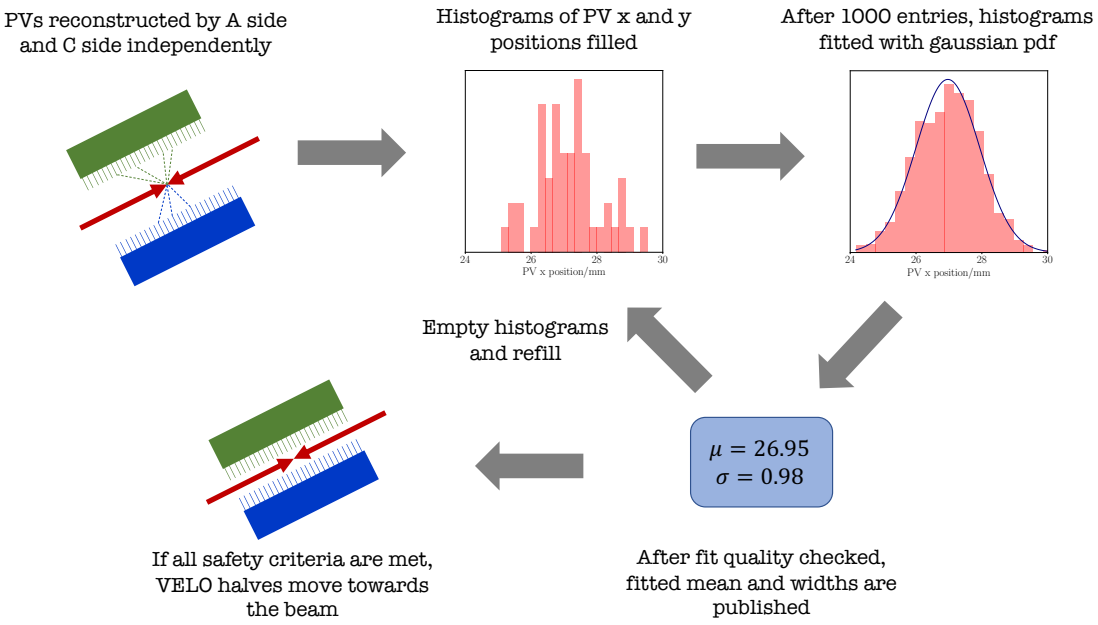


Figure 6.5: Overview of PV monitoring task for VELO closing procedure

In addition to PV x and y positions, other useful histograms are published. For example, plotting the trends in the PV positions could potentially indicate if one VELO half is moving unexpectedly. Also, a histogram recording the time taken for each PV measurement is plotted as it could reveal issues with data taking. Two dimensional histograms of the PV positions can be used to view the beamspot shape, or to study vertices from hadronic interactions with material, which will be discussed in the next section.

After the VELO is closed, data taking for physics commences, but the monitoring task continues running. This provides regular updates of the PV positions, acting as a safety system to ensure the beam does not drift significantly from its original position. If the beam does drift, or there is a significant interruption to the data acquisition, the monitoring task will detect this and trigger the VELO halves to move out fully.

6.4 RF foil metrology with vertex reconstruction

6.4.1 RF foil metrology from first VELO closing

The biggest danger when closing the VELO for the first time was potential damage to the RF foil. Since it encloses the VELO modules, the foil is the closest part of the VELO to the beam. In addition, it was possible that the foil had deformed from its original design shape. This is made more likely because of its incredibly thin shape, and the fact that it had a complex installation procedure, involving transport from Amsterdam to Liverpool, then to CERN, before being lowered $\sim 100\,\mathrm{m}$ underground into the LHCb cavern. If the foil is too close to the beams during the closing procedure, it can be permanently damaged, and the separation of the primary and secondary vacua could be lost. Moreover, the small separation between the module edges and the RF foil implies that if the foil were damaged, there would also be a strong possibility of damage to the VELO modules.

Therefore, before closing the VELO for the first time, it was vital to create a map of the RF foil shape to determine the miniumum distance between the foil and the beams. This can be achieved by selecting vertices arising from hadronic interactions with the foil material. Plotting the distribution of these vertex positions reveals the position of the RF foil edges, and thus determines how far the VELO halves can be moved safely. The vertex positions in the x, y plane – effectively looking along the beam direction – are shown in Fig. 6.6. Figure 6.7 shows the radial distance of the vertices from the beamline as a function of z. In effect, the VELO itself is used to perform a high resolution scan of its enclosing box.

Figure 6.6 combines vertices reconstructed independently from the A and C sides on one plot. To translate to a single coordinate system, fits were performed to find the beamspot positions, as measured by each half. Then these positions were used to transform the C side vertices to the A side coordinate system. This ensures that no external information about the VELO position is used, so the measurements are minimally biased. The plots show that the foil edges are well aligned, indicating that there was no significant rotation during assembly or installation. In Fig. 6.7,

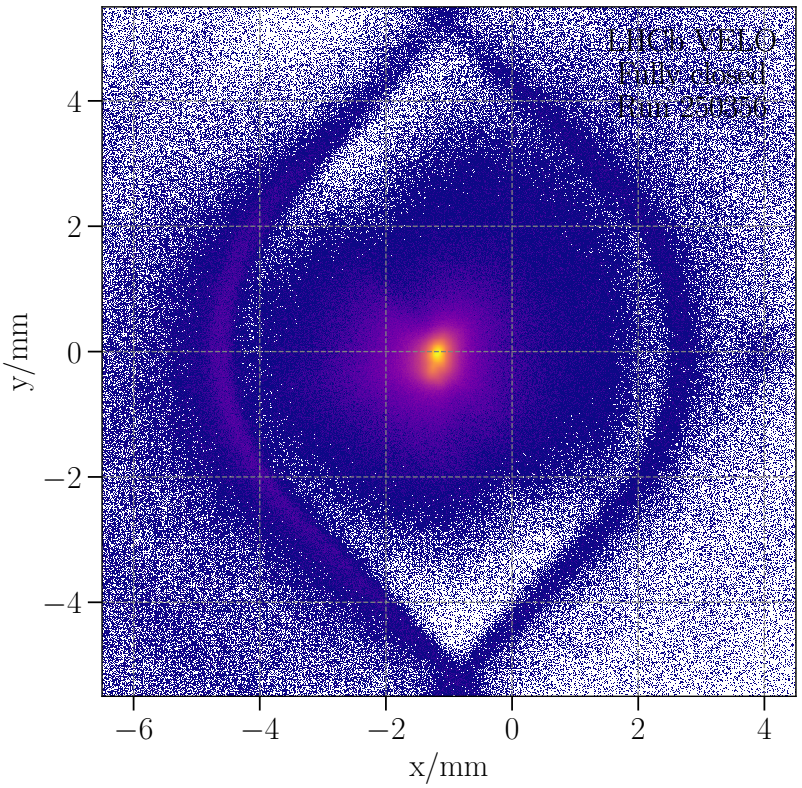


Figure 6.6: Histogram in the x, y plane of around 100 million vertices recorded in October 2022. The colours show the density on a log scale, with lighter colour indicating a higher number of events. The yellow spot in the centre is the collision region, and the diffuse purple area around it shows secondary vertices, e.g. B and D meson decays. The edges on the outer part of the plot show the inner edges of the RF foil.

the peaks in the vertex distribution line up perfectly with the module positions according to the design specification [113], showing no significant misalignment in the z direction. Also, most of the RF foil peaks fulfil the design clearance from the beams of $r = 3.5\,\text{mm}$. However, in the region $300\,\text{mm} < r < 600\,\text{mm}$, a small inward deformation is visible.

In addition to material interactions from the RF foil, secondary vertices can be reconstructed from further along the beampipe. For example, Fig. 6.8 shows vertices from interactions with the SMOG2 cell. The plot confirms that the cell is fully closed, and that the halves are well aligned.

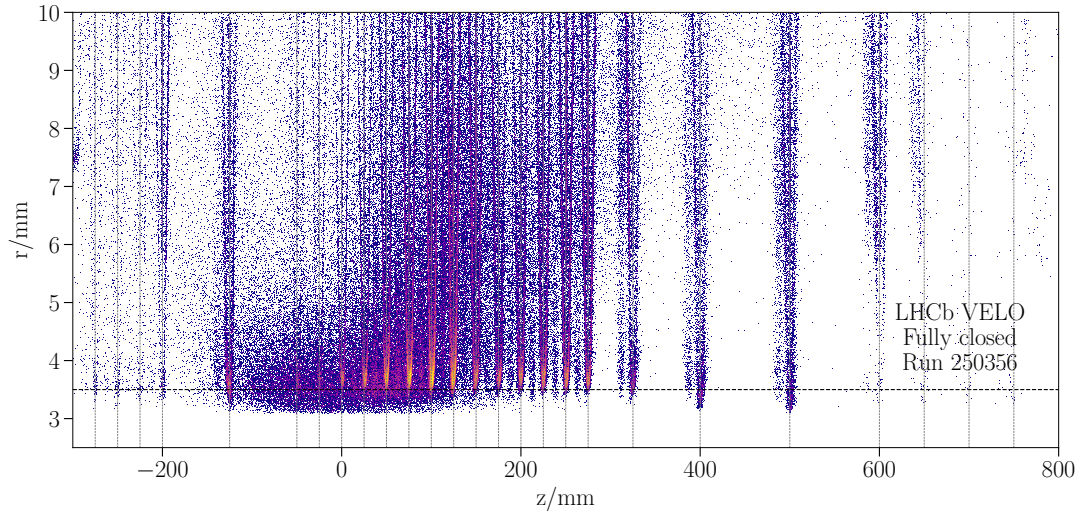


Figure 6.7: r, z position of vertices recorded in October 2022 from the VELO A side, where $r = \sqrt{x^2 + y^2}$, and z is along the beamline. Vertices from the collision region have been removed to more clearly show the structure of the RF foil. The vertical grey lines show the design position of the 26 A side modules. The horizontal black line at $r = 3.5\,\text{mm}$ indicates the design clearance between the foil and the beams.

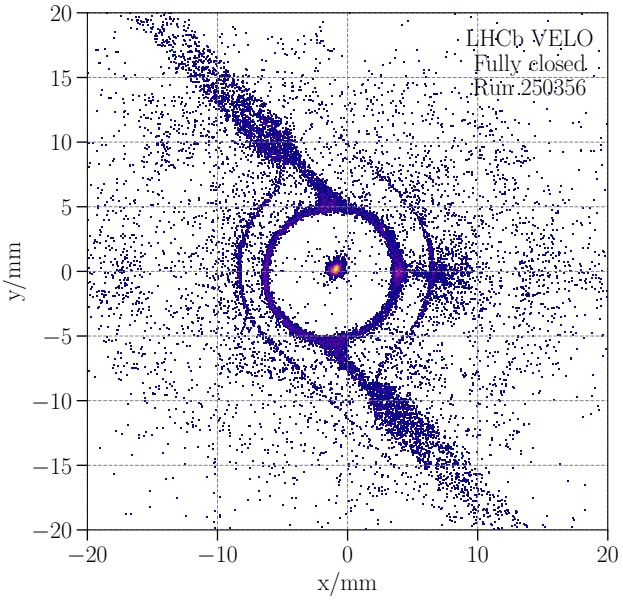


Figure 6.8: Vertices in the x, y plane in the region $-500\,\text{mm} < z < -300\,\text{mm}$ showing material interactions with the SMOG2 cell.

6.4.2 VELO aperture measurement

The plots of the VELO RF foil presented in the previous section are useful for checking whether the foil had any large deformations that would prevent the VELO from fully closing. However, before fully closing the VELO for the first time, a more accurate measurement of the aperture size was desired, especially since Fig. 6.7 indicated a possible bulge in the high z region. In this regard, a study was carried out using data taken with the VELO halves positioned 2 mm from their final position.

To calculate the aperture size, the exact positions of the RF foil edges in Fig. 6.6 can be estimated by fitting histograms of the vertex positions, as the interactions from the foil edges produce a higher density of vertices. The positions of the four edges can then be used to calculate the aperture size. A more detailed view of the VELO aperture can be obtained by repeating this process in discrete regions of z. This will show the variation of the aperture size along the VELO. Furthermore, due to the rotated L-shape geometry of the modules (e.g. see Fig. 6.4), vertices from material interactions on one half can also be reconstructed by the other half. This means that with enough data, both VELO halves can reconstruct the position of all four foil edges. Therefore, the edge position measurements can be made separately with both halves and compared as a cross-check.

The data was first split into 26 discrete regions in z, each region corresponding to the area covered by one module pair. A rotation of 45° was then applied to the x, y vertex positions, to account for the VELO geometry. The rotated coordinates a_1, a_2 are defined as:

$$\begin{aligned} a_1 &= -\frac{1}{\sqrt{2}}(x - y) \ , \\ a_2 &= -\frac{1}{\sqrt{2}}(x + y) \ , \end{aligned} \tag{6.1}$$

for the A side, and:

$$\begin{aligned} a_1 &= \frac{1}{\sqrt{2}}(x + y) \ , \\ a_2 &= \frac{1}{\sqrt{2}}(x - y) \ , \end{aligned} \tag{6.2}$$

for the C side. One dimensional histograms in the a_1 and a_2 variables were created and the four maximum likelihood fits were carried out to find the four edge positions (two vertical and two horizontal edges in the a_1, a_2 coordinates). The fit model was created by summing a gaussian function and a third order polynomial. The mean value of the gaussian function was taken as the edge position. The procedure was implemented for each of the 26 data samples, and for the A side and C side data separately.

Due to the geometry of the VELO halves, one of the four edges produces much fewer vertices than the others. This can be seen in Fig. 6.9. To solve this issue, the corresponding *outer* edge was fitted instead, as this yielded much more vertices. Then, to translate the outer edge measurement to the inner edge position, a translation of 12.5 mm was applied, which is the perpendicular distance between the two edges, as confirmed by an independent metrology measurement.

An example of the fits for one region in z is shown in Fig. 6.10. Using the fit results for the edge positions, the aperture size l is calculated by averaging the distances between the two pairs of opposite edges. Then the aperture radius R is estimated as:

$$R = \frac{l - 2\sqrt{2} - 0.2}{2} \text{ mm} . \tag{6.3}$$

The factor of $2\sqrt{2}$ is subtracted to account for the VELO being partially open, as each half was 2 mm from its final closed position. 0.2 mm is also subtracted to account for the the foil thickness.

The values for the aperture radius are summarised in Fig. 6.11. There is reasonably good agreement between the A and C side measurements. A clear trend is visible, the radius is consistently around 3.5 mm in the range $-100\,\text{mm} < z < 200\,\text{mm}$, but then decreases sharply to a minimum value of just over 3.0 mm around $z = 500\,\text{mm}$. The radius then increases towards 3.5 mm at higher regions of z. These measurements are consistent with the small deformation in the foil visible in Fig. 6.7. From this study, it was concluded that the minimum VELO aperture

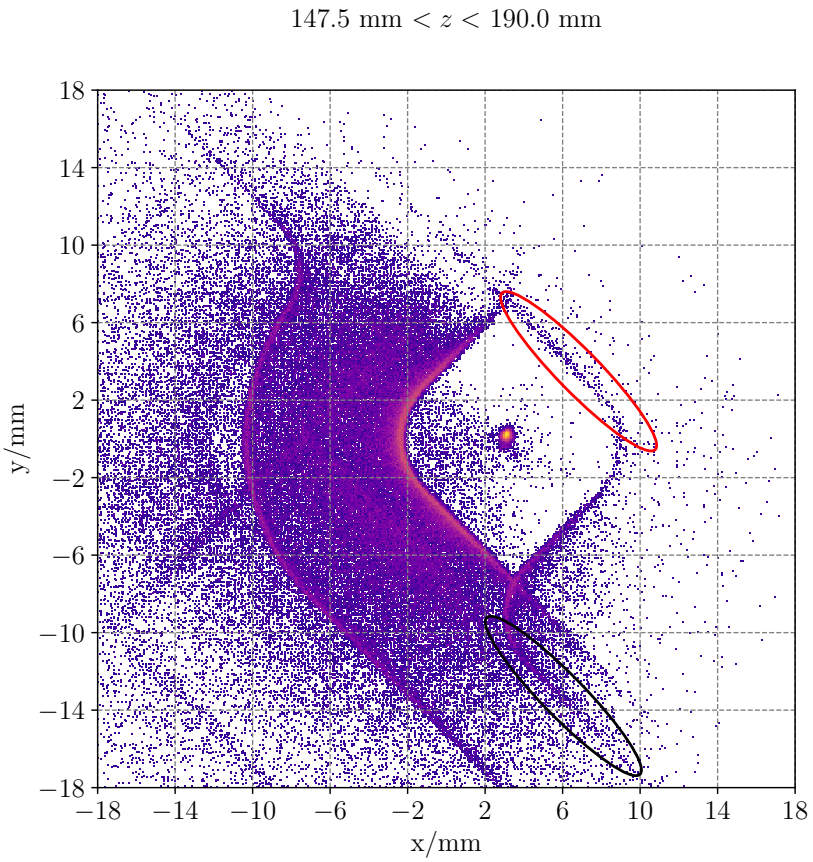


Figure 6.9: Vertices in the x, y plane reconstructed by the C side for one section in z. The inner edge of the A side (highlighted in red) has a low density of vertices, so the outer edge (highlighted in black) was fitted instead.

was just over 3.0 mm. This value was reported to the LHC Machine Protection Panel, who agreed it was safe to fully close the VELO.

6.4.3 RF foil metrology after vacuum incident

On 10/01/23, a failure of the VELO pressure safety system occurred [122]. Consequently, the pressure in the secondary vacuum increased to over 10 times the safe limit. This led to a significant inward (towards the beam) deformation of the RF foil. To assess the extent of the damage, data was taken with the VELO in the fully open position and vertices from material interactions were selected. Then equivalent plots to Figs. 6.6 and 6.7 were made to visualise the shape of the foil after the damage. They are shown in Fig. 6.12.

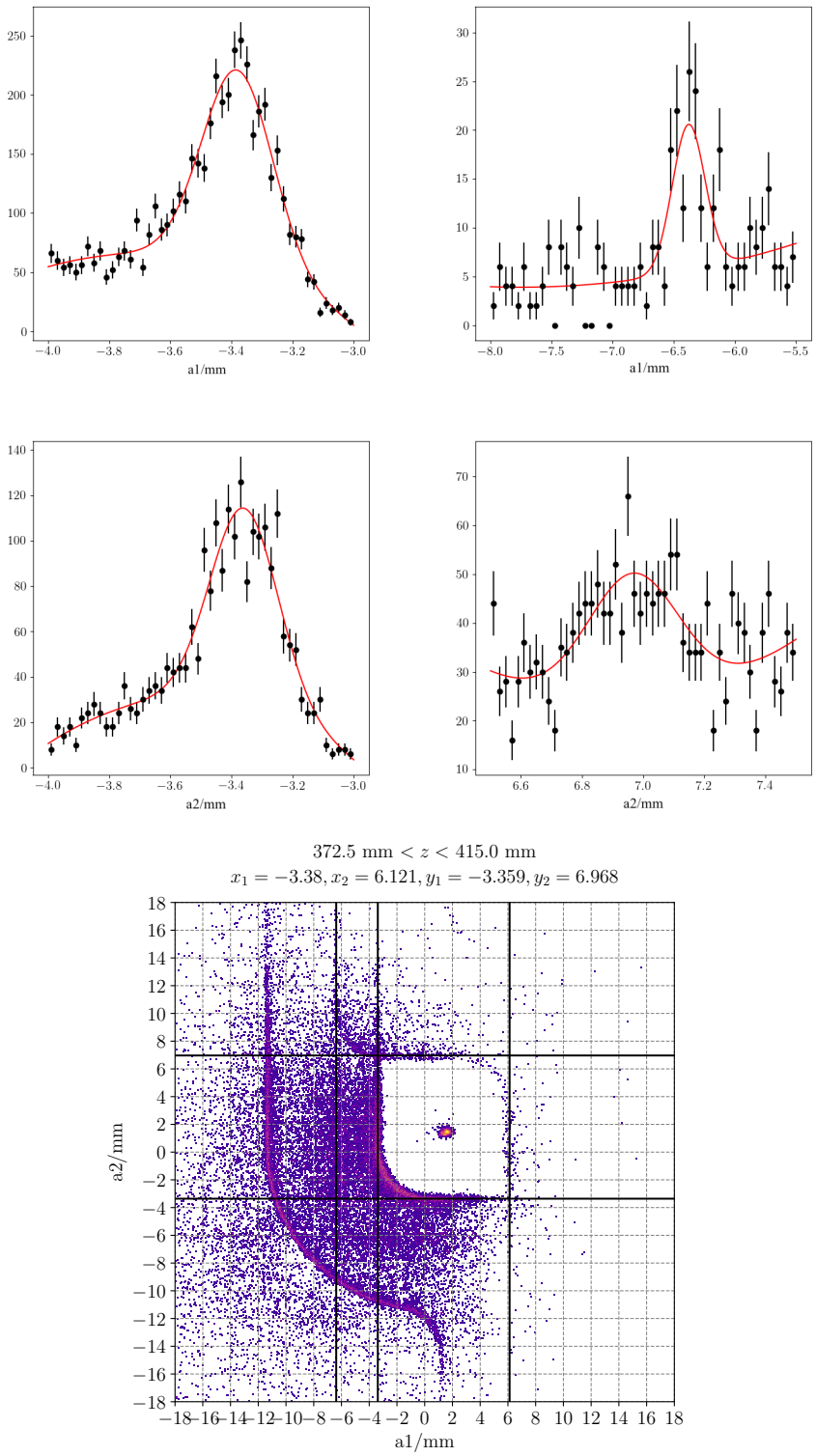


Figure 6.10: An example of the fitting procedure used to determine the foil edges for 372.2 mm $< z <$ 415.0 mm. The upper and middle plots show the individual fits. The fitted mean was taken as the position of the edge. The black points are histograms of the reconstructed vertex positions in the a_1 and a_2 variables. The red line shows the fitted shape. The lower plot is an a_1, a_2 projection of the vertices, with the fitted edge positions superimposed as black lines. The leftmost vertical line is the fitted inner edge and the rightmost vertical line shows the fitted position translated to the corresponding outer edge.

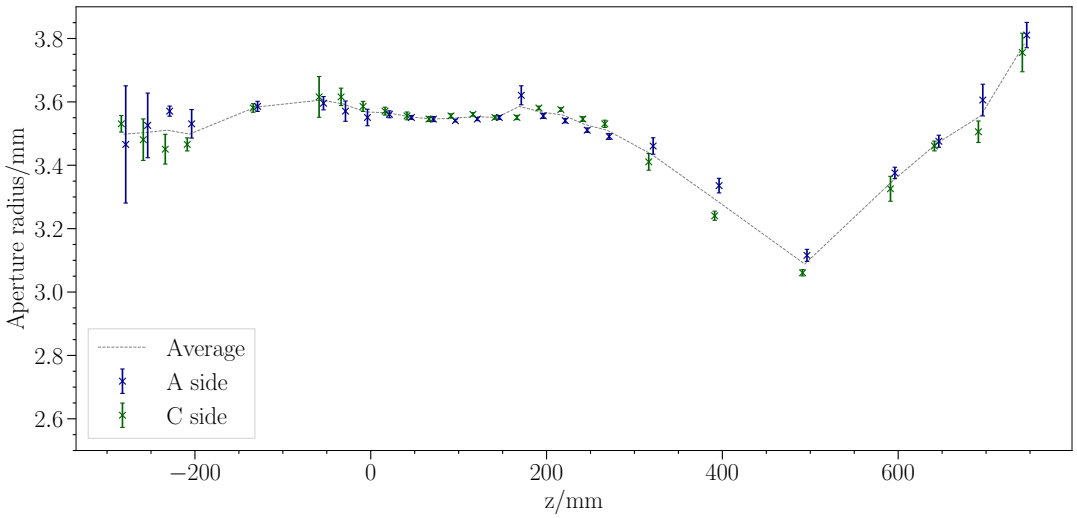


Figure 6.11: Measured aperture radius for 26 regions of z. Independent values from the A and C side data are shown.

During this data taking period, gas was injected using the SMOG2 cell. This greatly increased the number of tracks due to the presence of beam-gas collisions in addition to beam-beam collisions. Also, the beam-gas collisions yield greater resolution away from the collision region (at high z). This is important since there are relatively few tracks in this region from beam-beam collisions, and this was the area expected to show the most significant damage.

The plots in Fig. 6.12 show a significant buckling of the RF foil on both sides at high z. The size of the deformation can be seen by comparing Fig. 6.12 to Fig. 6.7. Interestingly, the damage appears to be most significant in the regions where the foil corrugations are more widely spaced. This indicates that the widely spaced corrugations are mechanical weak points. Figure 6.7 supports this idea, since the small deformation seen here before the vacuum incident was also in the region $400\,\text{mm} < z < 600\,\text{mm}$.

Following analysis of the tomography scans, it was determined that the damage to the VELO was too severe to close in during 2023 data taking. Instead the VELO would remain in the fully open position and the RF foils would be replaced during the technical stop at the end of the year. At the time of writing, the VELO RF

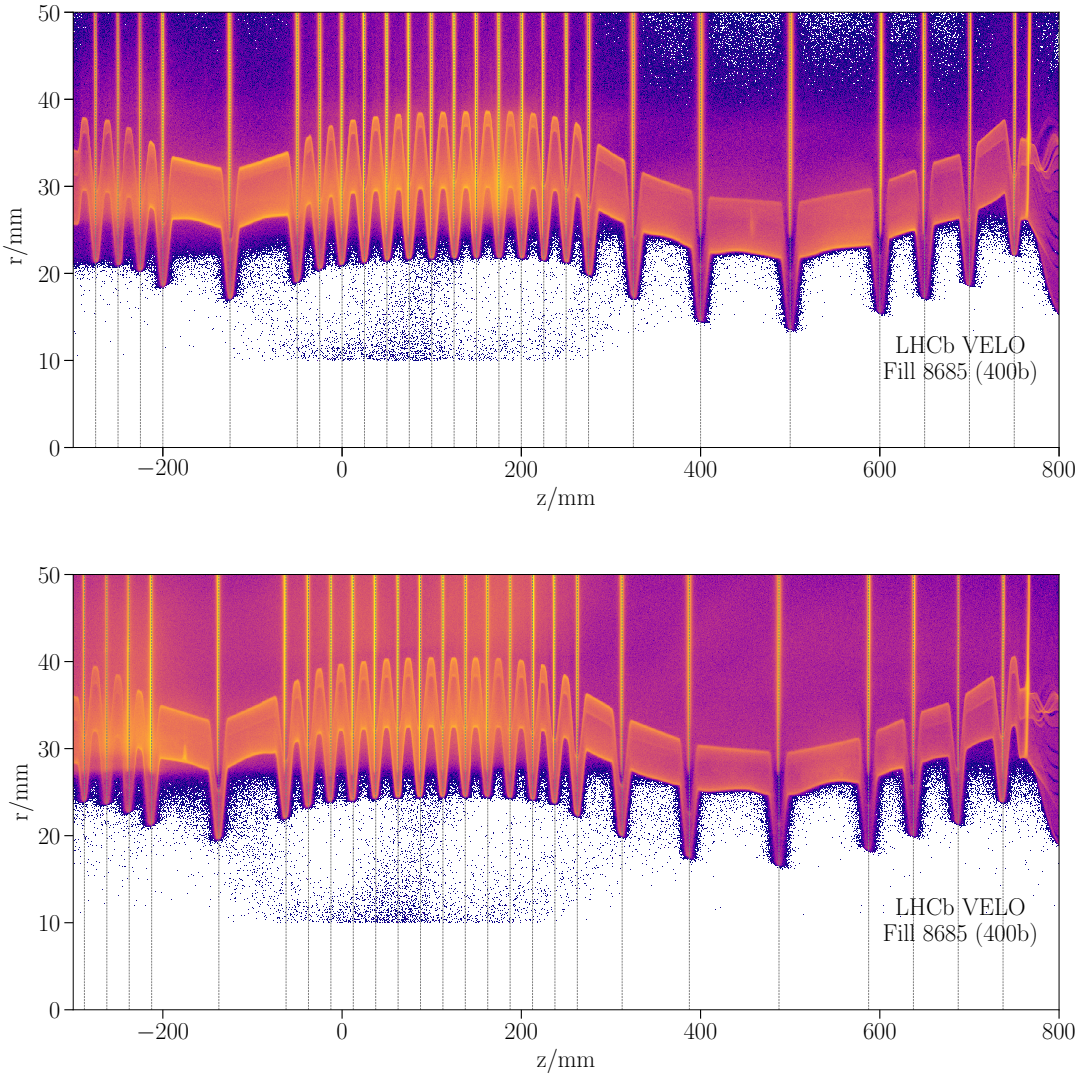


Figure 6.12: r, z projections of vertices from material interactions using data taken after the VELO vacuum incident. The upper plot shows 94 million vertices reconstructed by the A side and the lower plot shows 122 million vertices from the C side. 400b refers to the number of proton bunches in each beam during this fill.

foils have been successfully replaced so full physics data taking can be resumed with the VELO in the fully closed position.

7 Conclusions

7.1 Summary

The work presented in this thesis is the first full angular analysis of $B^0 \to D^{*-}\mu^+\nu_\mu$ decays. This was implemented using a novel method, involving binned multi-dimensional histograms that are agnostic to a particular choice of decay model. The measured values for the 12 angular observables have good agreement with the theoretical predictions. The results are summarised below, where the first uncertainty is statistical, and the second is systematic. All measured values are

consistent with the SM predictions shown in Tab. 3.1.

$$
\begin{aligned}
I_{1s} &= 0.3507 \pm 0.0024 \pm 0.0062 \\
I_{2c} &= -0.5005 \pm 0.0131 \pm 0.0135 \\
I_{2s} &= 0.1115 \pm 0.0039 \pm 0.0104 \\
I_{6c} &= 0.0233 \pm 0.0180 \pm 0.0127 \\
I_{6s} &= -0.2991 \pm 0.0061 \pm 0.0066 \\
I_{3} &= -0.1929 \pm 0.0031 \pm 0.0049 \\
I_{4} &= -0.2934 \pm 0.0036 \pm 0.0048 \\
I_{5} &= 0.2481 \pm 0.0034 \pm 0.0072 \\
I_{7} &= 0.0056 \pm 0.0023 \pm 0.0017 \\
I_{8} &= 0.0030 \pm 0.0025 \pm 0.0010 \\
I_{9} &= -0.0018 \pm 0.0028 \pm 0.0021 \\
I_{1c} &= 0.5394 \pm 0.0070 \pm 0.0149 \\
\mathcal{A}_{FB} &= -0.2156 \pm 0.0110 \pm 0.0069 \\
F_{L}^{D^*} &= 0.5297 \pm 0.0024 \pm 0.0107
\end{aligned}
$$

The results of this work can provide useful inputs for semileptonic form-factor fits. This could lead to improved constraints on NP effects through the calculation of Wilson coefficients [123]. The search for NP in semileptonic decays is motivated by the current tension between experimental measurements and SM predictions of $R(D^0)$ and $R(D^*)$. Alternatively, purely SM fits can be performed by combining the measured angular observables with other inputs such as lattice QCD data. These SM fits can be used to improve the predicted values for $R(D^0)$ and $R(D^*)$, as well as other parameters.

7.2 Looking towards the future

This work serves as a proof of principle for a model-independent angular analysis of a semileptonic process. Accordingly, this lays a foundation for many forthcoming studies in the field.

Since LHCb started Run 3 data taking last year, it will soon be possible to update the analysis with more data. The higher statistics (both in simulation and real data) will reduce the uncertainties of the angular observables, with the eventual goal of obtaining comparable or higher precision compared with the theoretical predictions.

A potential extension to this measurement would be to make separate measurements of the I_X for B^0 and $\overline{B}^0$ decays. This would give access to potential CP-violating effects in $B^0 \to D^{*-}\mu^+\nu_\mu$ processes. Currently, no evidence of direct CP violation has been measured in the $B^0 \to D^{*-}\mu^+\nu_\mu$ decay. In addition, this new measurement could reveal the influence of additional NP processes that would not be observed with the CP-averaged results.

In addition, the analysis methodology could be modified to allow the measurement of the I_X in bins of q^2, as suggested in Ref. [89]. This would be analogous to other LHCb angular analyses, such as $B \to K^*\mu\mu$ [62, 63]. The variation of the I_X functions with q^2 is predicted in the SM so results of a binned measurement would yield much greater sensitivity when included in form-factor fits.

Moreover, although there are compelling reasons to study $B^0 \to D^{*-}\mu^+\nu_\mu$ processes, the prospects of observing NP are generally thought to be much stronger in the 3rd lepton generation, i.e. in $B^0 \to D^{*-}\tau^+\nu_\tau$ decays. A complementary measurement of the I_X in $B^0 \to D^{*-}\tau^+\nu_\tau$ decays is ongoing. However, a more interesting prospect is the simultaneous angular analysis of the muonic and the tauonic modes with a sample of $B^0 \to D^{*-}\tau^+\nu_\tau$ data where the τ decays to a muon. This would involve modifying the fit setup used here by replacing the single component modelling $B^0 \to D^{*-}\tau^+\nu_\tau$ with 12 templates, in the same way as for the signal. The result of such a measurement would be 12 ratios, I_X^τ/I_X^μ. These ratios would benefit from the same advantages as measurements of $R(D^0)$ and $R(D^*)$, namely the cancellation of theoretical and some systematic uncertainties. Clearly, 12 angular observable ratios would be a far more powerful test of NP than branching-fraction ratios alone.

In summary, as demonstrated by the work presented in this thesis, angular analyses of semileptonic decays are viable at LHCb and they have the potential to

yield interesting and significant results. Therefore the further use of this technique is strongly motivated in Run 3 and beyond.

Appendices

A

Definition of functions used in mass fits

Crystal ball function: A function consisting of a gaussian component with a power law tail. It is defined as:

$$f\left(m; m_0, \sigma, \alpha_L, n_L, \alpha_R, n_R\right) = \begin{cases} A_L \cdot \left(B_L - \frac{m-m_0}{\sigma_L}\right)^{-n_L}, & \text{for } \frac{m-m_0}{\sigma_L} \leq -\alpha_L \\ \exp\left(-\frac{1}{2} \cdot \left[\frac{m-m_0}{\sigma_L}\right]^2\right), & \text{for } -\alpha_L < \frac{m-m_0}{\sigma_L} \leq 0 \\ \exp\left(-\frac{1}{2} \cdot \left[\frac{m-m_0}{\sigma_R}\right]^2\right), & \text{for } 0 < \frac{m-m_0}{\sigma_R} \leq \alpha_R \\ A_R \cdot \left(B_R + \frac{m-m_0}{\sigma_R}\right)^{-n_R}, & \text{otherwise} , \end{cases} \tag{A.1}$$

where

$$\begin{aligned} A_i &= \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \cdot \exp\left(-\frac{|\alpha_i|^2}{2}\right) , \\ B_i &= \frac{n_i}{|\alpha_i|} - |\alpha_i| \ . \end{aligned} \tag{A.2}$$

RooDstD0BG: Dedicated pdf shape that is commonly used to model the background of $m(D^*) - m(D^0)$ distributions. It is defined as:

$$\text{RooDSTD0}\left(m \mid m_0, A, B, C\right) = \left(1 - \exp\left(-\frac{m-m_0}{C}\right)\right) \cdot \left(\frac{m}{m_0}\right)^A + B \cdot \left(\frac{m}{m_0} - 1\right) , \tag{A.3}$$

where A, B, C and m_0 are parameters in the fit.

B
Helicity amplitudes

B.1 Leptonic amplitudes

The leptonic helicity amplitudes for each potential coupling type are defined as:

$$\begin{aligned} L^{\lambda_\ell}_{V-A,\lambda}\left(q^2,\chi,\theta_\ell\right) &= \eta_\mu(\lambda)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\gamma^\mu\left(1-\gamma_5\right)\nu\right|0\right\rangle, \\ L^{\lambda_\ell}_{S-P}\left(q^2,\chi,\theta_\ell\right) &= \left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\left(1-\gamma_5\right)\nu\right|0\right\rangle, \\ L^{\lambda_\ell}_{T-T5,\lambda\lambda'}\left(q^2,\chi,\theta_\ell\right) &= -L^{\lambda_\ell}_{T-T5,\lambda'\lambda} = -i\eta_\mu(\lambda)\eta_\nu\left(\lambda'\right)\left\langle \ell\left(\lambda_\ell\right)\bar{\nu}\left|\bar{\ell}\sigma^{\mu\nu}\left(1-\gamma_5\right)\nu\right|0\right\rangle. \end{aligned} \tag{B.1}$$

Using the polarization vectors, Dirac spinors, and gamma matrices in the Weyl representation, the explicit formulas for the vector type amplitudes are obtained:

$$\begin{aligned} L^{+}_{V-A,+}\left(q^2,\chi,\theta_\ell\right) &= \pm\sqrt{2}m_\ell\beta_\ell\sin\theta_\ell e^{-2i\chi}, \\ L^{+}_{V-A,-}\left(q^2,\chi,\theta_\ell\right) &= \pm\sqrt{2}m_\ell\beta_\ell\sin\theta_\ell, \\ L^{+}_{V-A,0}\left(q^2,\chi,\theta_\ell\right) &= 2m_\ell\beta_\ell\cos\theta_\ell e^{-i\chi}, \\ L^{+}_{V-A,t}\left(q^2,\chi,\theta_\ell\right) &= -2m_\ell\beta_\ell e^{-i\chi}, \\ L^{-}_{V-A,\pm}\left(q^2,\chi,\theta_\ell\right) &= \sqrt{2q^2}\beta_\ell\left(1\pm\cos\theta_\ell\right)e^{\mp i\chi}, \\ L^{-}_{V-A,0}\left(q^2,\chi,\theta_\ell\right) &= -2\sqrt{q^2}\beta_\ell\sin\theta_\ell, \\ L^{-}_{V-A,t}\left(q^2,\chi,\theta_\ell\right) &= 0, \end{aligned} \tag{B.2}$$

where $\beta_l = \sqrt{1 - m_l^2/q^2}$. The scalar-type amplitudes are:

$$\begin{aligned} L^+_{S-P}(q^2, \chi, \theta_\ell) &= -2\sqrt{q^2}\beta_\ell e^{-i\chi} \ , \\ L^-_{S-P}(q^2, \chi, \theta_\ell) &= 0 \ . \end{aligned} \tag{B.3}$$

The tensor-type amplitudes are:

$$\begin{aligned} L^+_{T-T5,+0}\left(q^2, \chi, \theta_\ell\right) &= \sqrt{2q^2}\beta_\ell \sin\theta_\ell \, e^{-2i\chi} \ , \\ L^+_{T-T5,-0}\left(q^2, \chi, \theta_\ell\right) &= \sqrt{2q^2}\beta_\ell \sin\theta_\ell \ , \\ L^+_{T-T5,+-}\left(q^2, \chi, \theta_\ell\right) &= -L^+_{T-T5,0t} = 2\sqrt{q^2}\beta_\ell \cos\theta_\ell \, e^{-i\chi} \ , \\ L^+_{T-T5,+t}\left(q^2, \chi, \theta_\ell\right) &= \mp\sqrt{2q^2}\beta_\ell \sin\theta_\ell \, e^{-2i\chi} \ , \\ L^+_{T-T5,-t}\left(q^2, \chi, \theta_\ell\right) &= \mp\sqrt{2q^2}\beta_\ell \sin\theta_\ell \ , \\ L^-_{T-T5,\pm 0}\left(q^2, \chi, \theta_\ell\right) &= \pm\sqrt{2}m_\ell\beta_\ell \left(1 \pm \cos\theta_\ell\right) e^{\mp i\chi} \ , \\ L^-_{T-T5,+-}\left(q^2, \chi, \theta_\ell\right) &= -L^-_{T-T5,0t} = -2m_\ell\beta_\ell \sin\theta_\ell \\ L^-_{T-T5,\pm t}\left(q^2, \chi, \theta_\ell\right) &= -\sqrt{2}m_\ell\beta_\ell \left(1 \pm \cos\theta_\ell\right) e^{\mp i\chi}. \end{aligned} \tag{B.4}$$

B.2 Hadronic amplitudes

In a similar way to the leptonic amplitudes, the hadronic helicity amplitudes for each coupling type are defined as:

$$\begin{aligned} H^{\lambda_M}_{V(A),\lambda}\left(q^2\right) &= \left(g_{V(A)} \pm 1\right) \eta^*_\mu(\lambda) \left\langle M\left(\lambda_M\right) |\bar{c}\gamma^\mu\left(\gamma_5\right) b| \bar{B}\right\rangle, \\ H_{S(P),\lambda}\left(q^2\right) &= g_{S(P)} \left\langle M\left(\lambda_M = 0\right) |\bar{c}\left(\gamma_5\right) b| \bar{B}\right\rangle, \\ H^{\lambda_M}_{T(T5),\lambda\lambda'}\left(q^2\right) &= -H^{\lambda_M}_{T,\lambda'\lambda}\left(q^2\right) = i\, g_{T(T5)}\, \eta^*_\mu(\lambda)\eta^*_\nu\left(\lambda'\right) \left\langle M\left(\lambda_M\right) |\bar{c}\sigma^{\mu\nu}\left(\gamma_5\right) b| \bar{B}\right\rangle, \end{aligned} \tag{B.5}$$

These can be written explicitly in terms of the coupling strengths (g_X), the masses of the particles, q^2, and the form factors:

$$\begin{aligned} H^{\pm}_{V,\pm}\left(q^2\right) &= \mp\left(g_V + 1\right) \frac{\sqrt{\lambda_{BD^*}\left(q^2\right)}}{m_B + m_{D^*}} V\left(q^2\right), \\ H^{\pm}_{A,\pm}\left(q^2\right) &= -\left(g_A - 1\right)\left(m_B + m_{D^*}\right) A_1\left(q^2\right), \\ H_0\left(q^2\right) \equiv H^0_{A,0}\left(q^2\right) &= \left(g_A - 1\right) \frac{8 m_B m_{D^*}}{\sqrt{q^2}} A_{12}\left(q^2\right), \\ H_t\left(q^2\right) \equiv H^0_{A,t}\left(q^2\right) &= \left(g_A - 1\right) \sqrt{\frac{\lambda_{BD^*}\left(q^2\right)}{q^2}} A_0\left(q^2\right), \\ H_P\left(q^2\right) &\simeq -g_P \frac{\sqrt{\lambda_{BD^*}\left(q^2\right)}}{m_b + m_c} A_0\left(q^2\right), \\ H^{\pm}_{T,\pm 0}\left(q^2\right) &= \mp\frac{g_T}{g_{T5}} H^{\pm}_{T5,\pm t}\left(q^2\right) = \pm g_T \frac{m_B^2 - m_{D^*}^2}{\sqrt{q^2}} T_2\left(q^2\right), \\ H^{\pm}_{T,\pm t}\left(q^2\right) &= \mp\frac{g_T}{g_{T5}} H^{\pm}_{T5,\pm 0}\left(q^2\right) = \pm g_T \sqrt{\frac{\lambda_{BD^*}\left(q^2\right)}{q^2}} T_1\left(q^2\right), \\ H^0_{T,+-}\left(q^2\right) &= -\frac{g_T}{g_{T5}} H^0_{T5,0t}\left(q^2\right) = -g_T \frac{4 m_B m_{D^*}}{m_B + m_{D^*}} T_{23}\left(q^2\right), \\ H^0_{V,0}\left(q^2\right) &= H^0_{V,t}\left(q^2\right) = H_S\left(q^2\right) = H^0_{T,0t}\left(q^2\right) = H^0_{T5,+-}\left(q^2\right) = 0, \end{aligned} \tag{B.6}$$

The above amplitudes can be combined and written in the form:

$$\begin{aligned}
\widetilde{H}_{\pm}^{\lambda_M,+} &\equiv H_{V,\pm}^{\lambda_M} + H_{A,\pm}^{\lambda_M} - 2\frac{\sqrt{q^2}}{m_\ell}\left(H_{T,\pm t}^{\lambda_M} \pm H_{T,\pm 0}^{\lambda_M} + H_{T5,\pm t}^{\lambda_M} \pm H_{T5,\pm 0}^{\lambda_M}\right), \\
\widetilde{H}_{0}^{\lambda_M,+} &\equiv H_{V,0}^{\lambda_M} + H_{A,0}^{\lambda_M} - 2\frac{\sqrt{q^2}}{m_\ell}\left(H_{T,+-}^{\lambda_M} + H_{T,0t}^{\lambda_M} + H_{T5,+-}^{\lambda_M} + H_{T5,0t}^{\lambda_M}\right), \\
\widetilde{H}_{\pm}^{\lambda_M,-} &\equiv H_{V,\pm}^{\lambda_M} + H_{A,\pm}^{\lambda_M} - 2\frac{m_\ell}{\sqrt{q^2}}\left(H_{T,\pm t}^{\lambda_M} \pm H_{T,\pm 0}^{\lambda_M} + H_{T5,\pm t}^{\lambda_M} \pm H_{T5,\pm 0}^{\lambda_M}\right), \\
\widetilde{H}_{0}^{\lambda_M,-} &\equiv H_{V,0}^{\lambda_M} + H_{A,0}^{\lambda_M} - 2\frac{m_\ell}{\sqrt{q^2}}\left(H_{T,+-}^{\lambda_M} + H_{T,0t}^{\lambda_M} + H_{T5,+-}^{\lambda_M} + H_{T5,0t}^{\lambda_M}\right), \\
\widetilde{H}_{t}^{\lambda_M} &\equiv \widetilde{H}_{V,t}^{\lambda_M} + \widetilde{H}_{A,t}^{\lambda_M},
\end{aligned} \tag{B.7}$$

where λ_M denotes the D^* helicity state. Additional linear combinations can be created for convenience:

$$\begin{aligned}
\widetilde{H}_{\pm}^{+}\left(q^2\right) &\equiv H_{\pm}\left(q^2\right) - 2\frac{\sqrt{q^2}}{m_\ell}H_{T,\pm}\left(q^2\right), \\
\widetilde{H}_{0}^{+}\left(q^2\right) &\equiv H_{0}\left(q^2\right) - 2\frac{\sqrt{q^2}}{m_\ell}H_{T,0}\left(q^2\right), \\
\widetilde{H}_{\pm}^{-}\left(q^2\right) &\equiv H_{\pm}\left(q^2\right) - 2\frac{m_\ell}{\sqrt{q^2}}H_{T,\pm}\left(q^2\right), \\
\widetilde{H}_{0}^{-}\left(q^2\right) &\equiv H_{0}\left(q^2\right) - 2\frac{m_\ell}{\sqrt{q^2}}H_{T,0}\left(q^2\right), \\
\widetilde{H}_{t}\left(q^2\right) &\equiv H_{t}\left(q^2\right) + \frac{\sqrt{q^2}}{m_\ell^2}H_{P}\left(q^2\right),
\end{aligned} \tag{B.8}$$

with:

$$\begin{aligned}
H_{\pm}\left(q^2\right) &\equiv H_{V,\pm}^{\pm}\left(q^2\right) + H_{A,\pm}^{\pm}\left(q^2\right) \\
H_{T,\pm}\left(q^2\right) &\equiv H_{T,\pm t}^{\pm}\left(q^2\right) \pm H_{T,\pm 0}^{\pm}\left(q^2\right) + H_{T5,\pm t}^{\pm}\left(q^2\right) \pm H_{T5,\pm 0}^{\pm}\left(q^2\right) \\
H_{T,0}\left(q^2\right) &\equiv H_{T,+-}^{0}\left(q^2\right) + H_{T5,0t}^{0}\left(q^2\right)
\end{aligned} \tag{B.9}$$

Now the I_X functions can be expressed in a more compact form in terms of these redefined helicity amplitudes, as shown in Appendix C.

C

Definition of I_X coefficients in terms of helicity amplitudes

The full angular decay rate is given by:

$$\begin{aligned}\frac{d^4\Gamma}{dq^2 d(\cos\theta_D) d(\cos\theta_L) d\chi} = \frac{9}{32\pi}\Big\{ & I_{1c}\cos^2\theta_D + I_{1s}\sin^2\theta_D \\ & + [I_{2c}\cos^2\theta_D + I_{2s}\sin^2\theta_D]\cos 2\theta_L \\ & + [I_{6c}\cos^2\theta_D + I_{6s}\sin^2\theta_D]\cos\theta_L \\ & + [I_3\cos 2\chi + I_9\sin 2\chi]\sin^2\theta_L\sin^2\theta_D \\ & + [I_4\cos\chi + I_8\sin\chi]\sin 2\theta_L\sin 2\theta_D \\ & + [I_5\cos\chi + I_7\sin\chi]\sin\theta_L\sin 2\theta_D\Big\}\end{aligned} \tag{C.1}$$

The 12 I_X coefficients are defined as:

$$I_{1c} = 2N\left[\left|\widetilde{H}_0^-\right|^2 + \frac{m_\ell^2}{q^2}\left|\widetilde{H}_0^+\right|^2 + 2\frac{m_\ell^2}{q^2}\left|\widetilde{H}_t\right|^2\right] \tag{C.2}$$

$$I_{1s} = \frac{N}{2}\left[3\left(\left|\widetilde{H}_+^-\right|^2 + \left|\widetilde{H}_-^-\right|^2\right) + \frac{m_\ell^2}{q^2}\left(\left|\widetilde{H}_+^+\right|^2 + \left|\widetilde{H}_-^+\right|^2\right)\right] \tag{C.3}$$

$$I_{2c} = 2N\left[-\left|\widetilde{H}_0^-\right|^2 + \frac{m_\ell^2}{q^2}\left|\widetilde{H}_0^+\right|^2\right] \tag{C.4}$$

$$I_{2s} = \frac{N}{2}\left[\left|\widetilde{H}_+^-\right|^2 + \left|\widetilde{H}_-^-\right|^2 - \frac{m_\ell^2}{q^2}\left(\left|\widetilde{H}_+^+\right|^2 + \left|\widetilde{H}_-^+\right|^2\right)\right] \tag{C.5}$$

$$I_3 = -2N\mathcal{R}e\left[\widetilde{H}_+^-\widetilde{H}_-^{-*} - \frac{m_\ell^2}{q^2}\widetilde{H}_+^+\widetilde{H}_-^{+*}\right] = -2N\beta_\ell^2\mathcal{R}e\left[H_+H_-^* - 4H_{T,+}H_{T,-}^*\right] \tag{C.6}$$

$$\begin{aligned} I_4 &= N\mathcal{R}e\left[\left(\tilde{H}_+^- + \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\tilde{H}_+^+ + \tilde{H}_-^+\right)\tilde{H}_0^{+*}\right] \\ &= N\beta_\ell^2\mathcal{R}e\left[\left(H_+ + H_-\right)H_0^* - 4\left(H_{T,+} + H_{T,-}\right)H_{T,0}^*\right] \end{aligned} \tag{C.7}$$

$$I_5 = 2N\mathcal{R}e\left[\left(\tilde{H}_+^- - \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\tilde{H}_+^+ + \tilde{H}_-^+\right)\tilde{H}_t^*\right] \tag{C.8}$$

$$I_{6c} = 8N\frac{m_\ell^2}{q^2}\mathcal{R}e\left[\widetilde{H}_0^+\tilde{H}_t^*\right] \tag{C.9}$$

$$I_{6s} = 2N\left(\left|\tilde{H}_+^-\right|^2 - \left|\tilde{H}_-^-\right|^2\right) \tag{C.10}$$

$$I_7 = 2N\mathcal{I}m\left[\left(\tilde{H}_+^- + \widetilde{H}_-^-\right)\widetilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\widetilde{H}_+^+ - \widetilde{H}_-^+\right)\widetilde{H}_t^*\right] \tag{C.11}$$

$$\begin{aligned} I_8 &= N\mathcal{I}m\left[\left(\tilde{H}_+^- - \tilde{H}_-^-\right)\tilde{H}_0^{-*} - \frac{m_\ell^2}{q^2}\left(\widetilde{H}_+^+ - \widetilde{H}_-^+\right)\widetilde{H}_0^{+*}\right] \\ &= N\beta_\ell^2\mathcal{I}m\left[\left(H_+ - H_-\right)H_0^* - 4\left(H_{T,+} - H_{T,-}\right)H_{T,0}^*\right] \end{aligned} \tag{C.12}$$

$$I_9 = -2N\mathcal{I}m\left[\widetilde{H}_+^-\widetilde{H}_-^{-*} - \frac{m_\ell^2}{q^2}\widetilde{H}_+^+\widetilde{H}_-^{+*}\right] = -2N\beta_\ell^2\mathcal{I}m\left[H_+H_-^* - 4H_{T,+}H_{T,-}^*\right] \tag{C.13}$$

with

$$N \equiv N\left(q^2\right) = \mathcal{B}\left(D^* \to D\pi\right)\frac{G_F^2\left|V_{cb}\right|^2}{48(2\pi)^3 m_B^3}q^2\sqrt{\lambda_{BD^*}\left(q^2\right)}\left(1 - \frac{m_\ell^2}{q^2}\right)^2 \tag{C.14}$$

Bibliography

[1] E. Rutherford, *VIII. Uranium radiation and the electrical conduction produced by it*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **47** (1899) 109.

[2] H. Becquerel, *Déviation du rayonnement du radium dans un champ électrique*, Comptes rendus de l'académie des sciences **130** (1900) 809.

[3] H.-E. Mahnke, *Lise Meitner, β-decay and non-radiative electromagnetic transitions*, Notes and records of the Royal Society of London **76** (2022) 107.

[4] E. Fermi, *An attempt of a theory of beta radiation. 1.*, Z. Phys. **88** (1934) 161.

[5] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[6] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[7] CPLEAR collaboration, A. Angelopoulos *et al.*, *First direct observation of time reversal noninvariance in the neutral kaon system*, Phys. Lett. B **444** (1998) 43.

[8] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[9] S. L. Glashow, *The renormalizability of vector meson interactions*, Nucl. Phys. **10** (1959) 107.

[10] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[11] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, Nuovo Cim. **11** (1959) 568.

[12] Gargamelle Neutrino collaboration, F. J. Hasert *et al.*, *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*, Phys. Lett. B **46** (1973) 138.

[13] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. B **122** (1983) 103.

[14] UA2 collaboration, M. Banner *et al.*, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. B **122** (1983) 476.

[15] C. De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, `https://cds.cern.ch/record/2119882`. [Accessed: 2023-02-14].

[16] LHCb collaboration, *$\bar{b}b$ production angle plots*, `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`. [Accessed: 2023-02-14].

[17] LHCb collaboration, B. Adeva *et al.*, *Roadmap for selected key measurements of LHCb*, LHCB-PUB-2009-029, 2009.

[18] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[19] LHCb collaboration, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, Nature Phys. **11** (2015) 743, `arXiv:1504.01568`.

[20] LHCb collaboration, R. Aaij *et al.*, *Measurement of $|V_{cb}|$ with $B_s^0 \to D_s^{(*)-}\mu^+\nu_\mu$ decays*, Phys. Rev. D **101** (2020) 072004, `arXiv:2001.03225`.

[21] LHCb collaboration, *Measurement of the ratios of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$*, Phys. Rev. Lett. **131** (2023) 111802, `arXiv:2302.02886`.

[22] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, `arXiv:1708.08856`.

[23] LHCb collaboration, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in $B^0 - \overline{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, `arXiv:1409.8586`.

[24] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B_c^- meson production fraction and asymmetry in 7 and 13 TeV pp collisions*, Phys. Rev. D **100** (2019) 112006, `arXiv:1910.13404`.

[25] LHCb collaboration, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (2015) 1530022, arXiv:1412.6352.

[26] LHCb collaboration, R. Aaij *et al.*, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030, 2003.

[27] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

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

[29] J. Gassner, M. Needham, and O. Steinkamp, *Layout and Expected Performance of the LHCb TT Station*, LHCb-2003-140, 2004.

[30] LHCb collaboration, R. Aaij *et al.*, *Momentum scale calibration of the LHCb spectrometer*, JINST **19** (2024) P02008, arXiv:2312.01772.

[31] T. K. Gaisser, *Cosmic rays and particle physics*, Cambridge University Press, 1990.

[32] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. C **73** (2013) 2431, arXiv:1211.6759.

[33] C. Abellán Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, LHCb-DP-2020-001, 2020.

[34] LHCb collaboration, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, arXiv:1812.10790.

[35] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transactions of the ASME – Journal of Basic Engineering **82D** (1960) 35.

[36] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, LHCb-2007-015, 2007.

[37] R. Aaij *et al.*, *Optimization of the muon reconstruction algorithms for LHCb Run 2*, LHCb-PUB-2017-007, 2017.

[38] LHCb Collaboration, *LHCb performance numbers*, https://lhcb.web.cern.ch/speakersbureau/html/PerformanceNumbers.html, 2023. [Accessed: 2023-02-16].

[39] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[40] A. Celis, M. Jung, X.-Q. Li, and A. Pich, *Sensitivity to charged scalars in $B \to D^{(*)}\tau\nu_\tau$ and $B \to \tau\nu_\tau$ decays*, JHEP **01** (2013) 054, arXiv:1210.8443.

[41] M. Tanaka, *Charged Higgs effects on exclusive semitauonic B decays*, Z. Phys. C **67** (1995) 321, arXiv:hep-ph/9411405.

[42] S. Iguro and K. Tobe, *$R(D^{(*)})$ in a general two Higgs doublet model*, Nucl. Phys. B **925** (2017) 560, arXiv:1708.06176.

[43] S. M. Boucenna *et al.*, *Phenomenology of an $SU(2) \times SU(2) \times U(1)$ model with lepton-flavour non-universality*, JHEP **12** (2016) 059, arXiv:1608.01349.

[44] B. Bhattacharya *et al.*, *Simultaneous Explanation of the R_K and $R_{D^{(*)}}$ Puzzles: a Model Analysis*, JHEP **01** (2017) 015, arXiv:1609.09078.

[45] P. Asadi, M. R. Buckley, and D. Shih, *It's all right(-handed neutrinos): a new W' model for the $R_{D^{(*)}}$ anomaly*, JHEP **09** (2018) 010, arXiv:1804.04135.

[46] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. D **89** (2014) 095033, arXiv:1403.1269.

[47] A. Celis, J. Fuentes-Martin, M. Jung, and H. Serodio, *Family nonuniversal Z' models with protected flavor-changing interactions*, Phys. Rev. D **92** (2015) 015007, arXiv:1505.03079.

[48] M. Bauer and M. Neubert, *Minimal Leptoquark Explanation for the $R_{D^{(*)}}, R_K, and (g-2)_\mu$ Anomalies*, Phys. Rev. Lett. **116** (2016) 141802, arXiv:1511.01900.

[49] B. Dumont, K. Nishiwaki, and R. Watanabe, *LHC constraints and prospects for S_1 scalar leptoquark explaining the $\bar{B} \to D^{(*)}\tau\bar{\nu}$ anomaly*, Phys. Rev. D **94** (2016) 034001, arXiv:1603.05248.

[50] Y. Sakaki, M. Tanaka, A. Tayduganov, and R. Watanabe, *Testing leptoquark models in $\bar{B} \to D^{(*)}\tau\bar{\nu}$*, Phys. Rev. D **88** (2013) 094012, arXiv:1309.0301.

[51] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870.

[52] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavor universality using $B^0 \to D^{*-}\tau^+\nu_\tau$ decays with hadronic τ channels*, Phys. Rev. D **108** (2023) 012018, arXiv:2305.01463.

[53] BaBar collaboration, J. P. Lees *et al.*, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. D **88** (2013) 072012, arXiv:1303.0571.

[54] Belle collaboration, G. Caria *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, Phys. Rev. Lett. **124** (2020) 161803, arXiv:1910.05864.

[55] LHCb collaboration, R. Aaij *et al.*, *Observation of the decay $\Lambda_b^0 \to \Lambda_c^+\tau^-\overline{\nu}_\tau$*, Phys. Rev. Lett. **128** (2022) 191803, arXiv:2201.03497.

[56] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$*, Phys. Rev. Lett. **120** (2018) 121801, arXiv:1711.05623.

[57] HFLAV collaboration, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, Eur. Phys. J. **C81** (2021) 226, arXiv:1909.12524, updated results and plots available at https://hflav.web.cern.ch/.

[58] Belle collaboration, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ with one-prong hadronic τ decays at Belle*, Phys. Rev. D **97** (2018) 012004, arXiv:1709.00129.

[59] Belle collaboration, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017) 211801, arXiv:1612.00529.

[60] Belle collaboration, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. D **92** (2015) 072014, arXiv:1507.03233.

[61] LHCb collaboration, R. Aaij *et al.*, *Measurement of lepton universality parameters in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays*, Phys. Rev. D **108** (2023) 032002, arXiv:2212.09153.

[62] LHCb collaboration, R. Aaij *et al.*, *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

[63] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.

[64] P. Koppenberg, *Flavour anomalies*, `https://www.nikhef.nl/~pkoppenb/anomalies.html`. [Accessed: 2024-01-12].

[65] I. Caprini, L. Lellouch, and M. Neubert, *Dispersive bounds on the shape of $B \to D^{(*)} l\nu$ form factors*, Nucl. Phys. B **530** (1998) 153, `arXiv:hep-ph/9712417`.

[66] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Constraints on form-factors for exclusive semileptonic heavy to light meson decays*, Phys. Rev. Lett. **74** (1995) 4603, `arXiv:hep-ph/9412324`.

[67] HPQCD collaboration, J. Harrison, C. Davies, and M. Wingate, *Lattice QCD calculation of the $B_{(s)} \to D^*_{(s)} \ell\nu$ form factors at zero recoil and implications for $|V_{cb}|$*, Phys. Rev. D **97** (2018) 054502, `arXiv:1711.11013`.

[68] I. I. Balitsky, V. M. Braun, and A. V. Kolesnichenko, *Radiative decay $\sigma^+ \to p\gamma$ in quantum chromodynamics*, Nuclear Physics B **312** (1989) 509.

[69] Fermilab Lattice and MILC collaborations, J. A. Bailey *et al.*, *Update of $|V_{cb}|$ from the $\bar{B} \to D^* \ell\bar{\nu}$ form factor at zero recoil with three-flavor lattice QCD*, Phys. Rev. D **89** (2014) 114504, `arXiv:1403.0635`.

[70] G. Martinelli, S. Simula, and L. Vittorio, *Constraints for the semileptonic B→D(*) form factors from lattice QCD simulations of two-point correlation functions*, Phys. Rev. D **104** (2021) 094512, `arXiv:2105.07851`.

[71] Fermilab Lattice and MILC collaborations, A. Bazavov *et al.*, *Semileptonic form factors for $B \to D^* \ell\nu$ at nonzero recoil from $2+1$-flavor lattice QCD: Fermilab Lattice and MILC Collaborations*, Eur. Phys. J. C **82** (2022) 1141, `arXiv:2105.14019`, [Erratum: Eur.Phys.J.C 83, 21 (2023)].

[72] E. Eichten and B. Hill, *An effective field theory for the calculation of matrix elements involving heavy quarks*, Physics Letters B **234** (1990) 511 .

[73] J. G. Korner and G. A. Schuler, *Exclusive Semileptonic Heavy Meson Decays Including Lepton Mass Effects*, Z. Phys. C **46** (1990) 93.

[74] A. Sirlin, *Large m(W), m(Z) Behavior of the O(alpha) Corrections to Semileptonic Processes Mediated by W*, Nucl. Phys. B **196** (1982) 83.

[75] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[76] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Precision corrections to dispersive bounds on form-factors*, Phys. Rev. D **56** (1997) 6895, arXiv:hep-ph/9705252.

[77] B. Grinstein and A. Kobach, *Model-Independent Extraction of $|V_{cb}|$ from $\bar{B} \to D^* \ell \overline{\nu}$*, Phys. Lett. B **771** (2017) 359, arXiv:1703.08170.

[78] D. Bigi, P. Gambino, and S. Schacht, *$R(D^*)$, $|V_{cb}|$, and the Heavy Quark Symmetry relations between form factors*, JHEP **11** (2017) 061, arXiv:1707.09509.

[79] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, *Tensions and correlations in $|V_{cb}|$ determinations*, Phys. Rev. D **96** (2017) 091503, arXiv:1708.07134.

[80] D. Bigi, P. Gambino, and S. Schacht, *A fresh look at the determination of $|V_{cb}|$ from $B \to D^* \ell \nu$*, Phys. Lett. B **769** (2017) 441, arXiv:1703.06124.

[81] D. Becirevic, M. Fedele, I. Nisandzic, and A. Tayduganov, *Lepton Flavor Universality tests through angular observables of $\overline{B} \to D^{(*)} \ell \overline{\nu}$ decay modes*, arXiv:1907.02257.

[82] D. Becirevic, S. Fajfer, I. Nisandzic, and A. Tayduganov, *Angular distributions of $\bar{B} \to D^{(*)} \ell \bar{\nu}_\ell$ decays and search of New Physics*, Nucl. Phys. B **946** (2019) 114707, arXiv:1602.03030.

[83] D. van Dyk *et al.*, *EOS - A Software for Flavor Physics Phenomenology*, arXiv:2111.15428.

[84] M. Bordone, M. Jung, and D. van Dyk, *Theory determination of $\bar{B} \to D^{(*)} \ell^- \bar{\nu}$ form factors at $\mathcal{O}(1/m_c^2)$*, Eur. Phys. J. C **80** (2020) 74, arXiv:1908.09398.

[85] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0} \mu^+ \mu^-$*, JHEP **08** (2013) 131, arXiv:1304.6325.

[86] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, JHEP **01** (2009) 019, arXiv:0811.1214.

[87] S. Dambach, U. Langenegger, and A. Starodumov, *Neutrino reconstruction with topological information*, Nucl. Instrum. Meth. A **569** (2006) 824, arXiv:hep-ph/0607294.

[88] Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020) 083C01.

[89] D. Hill, M. John, W. Ke, and A. Poluektov, *Model-independent method for measuring the angular coefficients of $B^0 \to D^{*-}\tau^+\nu_\tau$ decays*, JHEP **11** (2019) 133, arXiv:1908.04643.

[90] G. Ciezarek, A. Lupato, M. Rotondo, and M. Vesterinen, *Reconstruction of semileptonically decaying beauty hadrons produced in high energy pp collisions*, JHEP **02** (2017) 021, arXiv:1611.08522.

[91] SLD collaboration, K. Abe *et al.*, *A Measurement of R(b) using a vertex mass tag*, Phys. Rev. Lett. **80** (1998) 660, arXiv:hep-ex/9708015.

[92] C. Bozzi, *LHCb Computing Resource usage in 2021*, LHCb-PUB-2022-011, 2022.

[93] M. P. Whitehead, *A Palette of Fast Simulations in LHCb*, LHCb-TALK-2018-302, 2018.

[94] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^- \to D^+\pi^-\pi^-$ decays*, Phys. Rev. D **94** (2016) 072001, arXiv:1608.01289.

[95] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, arXiv:1506.08614, [Erratum: Phys.Rev.Lett. 115, 159901 (2015)].

[96] M. Pivk and F. R. L. Diberder, *sPlot: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005) 356.

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

[98] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[99] B. P. Roe *et al.*, *Boosted decision trees, an alternative to artificial neural networks*, Nucl. Instrum. Meth. A **543** (2005) 577, arXiv:physics/0408124.

[100] D0 collaboration, V. M. Abazov *et al.*, *Evidence for production of single top quarks*, Phys. Rev. D **78** (2008) 012005, arXiv:0803.0739.

[101] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, doi: 10.1145/2939672.2939785 arXiv:1603.02754.

[102] J. A. Hanley and B. J. McNeil, *The meaning and use of the area under a receiver operating characteristic (ROC) curve*, Radiology **143** (1982) 29.

[103] A. N. Shiryayev, *Selected works of AN Kolmogorov: Volume II probability theory and mathematical statistics*, vol. 26, Springer Science & Business Media, 1992.

[104] S. A. Braun, *Measurement of the CKM matrix elements $|V_{ub}|/|V_{cb}|$ from semileptonic B_s decays.*, PhD thesis, U. Heidelberg (main), 2020, doi: 10.11588/heidok.00028524.

[105] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[106] J. Eschle, A. Puig Navarro, R. Silva Coutinho, and N. Serra, *zfit: Scalable pythonic fitting*, SoftwareX **11** (2020) 100508.

[107] F. Beaujean, M. Chrzaszcz, N. Serra, and D. van Dyk, *Extracting Angular Observables without a Likelihood and Applications to Rare Decays*, Phys. Rev. D **91** (2015) 114012, arXiv:1503.04100.

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

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

[110] M. L. Waskom, *seaborn: statistical data visualization*, Journal of Open Source Software **6** (2021) 3021.

[111] F. U. Bernlochner *et al.*, *Das ist der HAMMER: Consistent new physics interpretations of semileptonic decays*, Eur. Phys. J. C **80** (2020) 883, arXiv:2002.00020.

[112] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[113] V. Coco *et al.*, *Velo Upgrade Module Nomenclature*, LHCb-PUB-2019-008, 2019.

[114] K. Hennessy, *LHCb VELO Upgrade*, Nucl. Instrum. Meth. A **845** (2017) 97, arXiv:1604.05045.

[115] T. Poikela *et al.*, *VeloPix: the pixel ASIC for the LHCb upgrade*, JINST **10** (2015) C01057.

[116] E. Frojdh *et al.*, *Timepix3: first measurements and characterization of a hybrid-pixel detector working in event driven mode*, JINST **10** (2015) C01039.

[117] P. Collins and P. Charitos, *A new VELO for the LHCb experiment*, https://ep-news.web.cern.ch/content/new-velo-lhcb-experiment, 2021. [Accessed: 2023-03-03].

[118] M. Kraan, *VELO's voyage into the unknown*, https://cerncourier.com/a/velos-voyage-into-the-unknown/, 2022. [Accessed: 2023-03-03].

[119] LHCb collaboration, *LHCb SMOG Upgrade*, CERN-LHCC-2019-005, LHCB-TDR-020, 2019.

[120] C. Gaspar, M. Dönszelmann, and P. Charpentier, *DIM, a portable, light weight package for information publishing, data transfer and inter-process communication*, Comput. Phys. Commun. **140** (2001) 102.

[121] M. Adinolfi *et al.*, *LHCb data quality monitoring*, J. Phys. Conf. Ser. **898** (2017) 092027.

[122] F. Reiss, *LHCb status report*, https://indico.cern.ch/event/1254283/#143-lhcb-status-report. 153rd LHCC meeting - Open session. [Accessed: 2024-01-16].

[123] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.