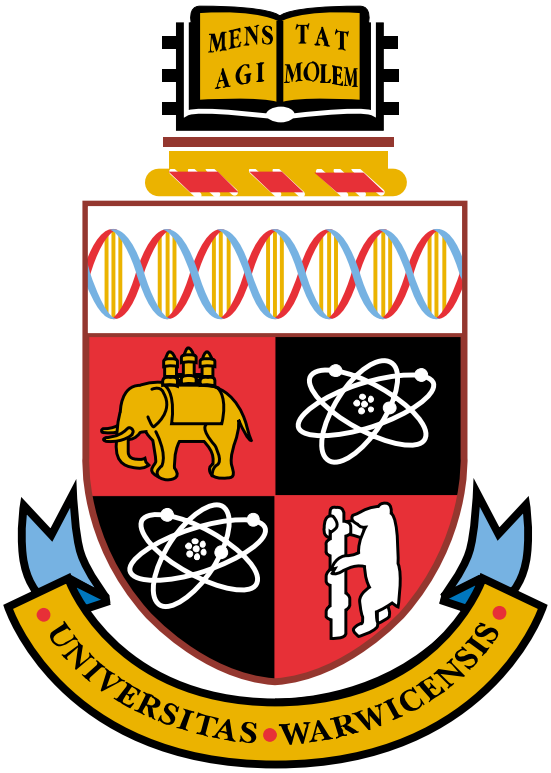


Angular analysis of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode at LHCb

by

Jake Alexander Ward

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Physics

University of Warwick, Department of Physics
Monash University, School of Physics & Astronomy

June 2025

Contents

List of Figures

List of Tables

Copyright notice

Abstract

Flavour changing neutral current processes, such as the $b \to d\ell^+\ell^-$ quark transitions, occur only at loop level within the Standard Model (SM), making them highly sensitive probes for potential New Physics (NP) contributions. This thesis presents work on the first angular analysis of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, where muon pairs do not originate from a resonance. The analysis utilises the datasets collected by the LHCb experiment during the years 2011-2012 and 2015-2018, in proton-proton collisions at the Large Hadron Collider. The combined dataset corresponds to an integrated luminosity of $9\,\text{fb}^{-1}$. The angular observables F_L, S_3, A_6, and A_9 are measured within two distinct bins of dimuon invariant mass squared (q^2) and within the dipion range of $0.62 - 0.92\,\text{GeV}$. Results are kept blind in this work pending review within the LHCb collaboration. Consistency checks using the $B^0 \to J/\psi\pi^+\pi^-$ and $B^0 \to \psi(2S)\pi^+\pi^-$ decay modes show agreement with existing published results, supporting the robustness of the analysis.

Acknowledgements

I owe many people a massive thank you for their contribution to my PhD. I have been lucky enough to have produced this work at both Warwick and Monash University and have appreciated the kindness of many people.

To begin my many thank yous I would like to highlight my supervisors at Warwick, Michal and Tom. For many years they have guided me and been generous with their time. I am extremely grateful to Ulrik for welcoming me into the group at Monash, I appreciated every piece of advice he bestowed.

I am deeply thankful to Lars for his support and patience. His guidance inspired me to pursue the PhD and has continued to push me forward as I complete the work I embarked upon at Glasgow five years ago. I would also like to extend this thanks to Michael, for his patience as I slowly learned how to code.

The many many people I have known at Warwick have been approachable and a delight to be around, the lunches eating poutine were a highlight. I'd like to single out Anja, Aidan, Lorenzo, Mathew, and Tom. Their company has been a blessing, both in Warwick and Geneva. My final wish is that the wall-of-bad-plots be maintained in my honour.

My year in Australia was a wonderful chapter in my life. I had the pleasure to share a space with a fantastic group of people. Among many others I would like to thank Tom, Sam, Mathew, and Rongrong for making my year down under a delight. Thank you to Eliot and Jav, I've really enjoyed our conversations over the years, thank you both for your friendship.

My time at CERN was an eye-opening experience and full of many wonderful people. I am especially grateful to have shared the experience with Aidan (again), Dan and Niall. I hope we continue to cross paths for many years to come.

Outside of the world of CERN, I am incredibly appreciative of the support of my friends and family. Freya, thank you for being my guinea pig and for taking this journey first.

Finally, I would like to offer my unwavering apology to Fiona for my four year jaunt around the globe. You have been a rock throughout this journey and i have always felt that your support is never ending. I look forward to our future together, I have no doubt it will involve penguins.

Declaration

This thesis is submitted to the University of Warwick and Monash University in support of my application for the degree of Doctor of Philosophy in Physics. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented in this thesis was carried out by the author, except in the cases outlined below:

Chapter 1 introduces the theoretical and experimental status of the field of flavour physics. This chapter is composed of material from textbooks and scientific journals. Chapter 2 provides an overview of the LHC accelerator complex and the LHCb experiment, with information derived from technical design reports and published literature.

Chapters 3 to 6 describe the physics analysis presented in the thesis and is composed entirely of the authors own work. The data and software from the LHCb collaboration is utilised in this analysis, and it is clearly stated when it is used.

Jake Alexander Ward
June 2025

Introduction

The Standard Model (SM) of particle physics is a theory which has thus far proved to be robust in describing the nature of elementary particles and three of the four fundamental forces. The predictions of the SM have explained many experimental particle physics phenomena and the behaviour of subatomic particles. Therefore, it represents our best understanding of the subject matter, however it is not a complete theory of fundamental physics.

The theory was developed gradually throughout the 1950s to the early 1970s and has since had success in having many of its predictions confirmed, culminating with the discovery of the Higgs Boson in 2012 [2, 3]. While the SM has been incredibly successful in its predictions over the last 60 years, there are various shortcomings which need to be resolved by experiments, requiring an expansion of the theory or a completely new approach. The SM has thus far been unable to explain the existence of dark matter, the asymmetry of matter and antimatter in the universe, or the origin of finite neutrino mass. These phenomena, which are not explained by the SM, indicate that the SM is an approximation of a more complete theory. Tests probing the limits of the SM predictions take the form of searches for new particles and performing precision measurements of SM properties, the subject of this thesis is the latter. Precision measurements could indicate New Physics (NP) contributions which give rise to deviations from SM predictions. The LHCb experiment specialises in indirect searches for NP whereby measurements are performed to test the SM and seek deviations from the expected results.

One avenue physicists are exploring for signs of NP is the study of the rare decays of beauty hadrons, such as the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode. This decay is an example of the $b \to d\ell^+\ell^-$ quark transition, which is a flavour changing neutral current (FCNC) process. These processes are forbidden at tree level in the SM and are sensitive to the effects of NP, which can alter amplitudes or introduce new particles. Measuring observables related to these rare decays, such as those associated with the angular topology of the decay, provide insights into the the SM and its limitations.

This thesis presents an analysis focusing on the measurement of angular observables in the rare $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, using data collected by the LHCb experiment at CERN. Chapter 1 introduces the theoretical background, from the SM through to heavy flavour physics and an overview of the experimental status of $b \to d\ell^+\ell^-$ processes. Chapter 2 details the LHCb detector and the reconstruction of beauty hadron decays. Chapters 3 to 6 discuss the angular analysis of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, including the preparation of the data, the measurement of the observables, and the evaluation of systematic uncertainties.

During the course of my PhD, I contributed to an additional physics analysis that

involved measuring the lifetime of the Ω_c^0 and Ξ_c^0 baryons. The measurement was performed using the fully-hadronic $\Omega_b^- \to \Omega_c^0\pi^-$ and $\Xi_b^- \to \Xi_c^0\pi^-$ decay modes, respectively, and was made relative to the lifetime of the D^0 from $B^- \to D^0\pi^-$ decays. The theoretical predictions for the hierarchy of the singly-charmed baryon lifetimes are based on the Heavy Quark Effective Theory (HQET) [4], which predicted the hierarchy of:

$$\tau(\Omega_c^0) < \tau(\Xi_c^0) < \tau(\Lambda_c^+) < \tau(\Xi_c^+). \tag{1}$$

This hierarchy remained in agreement with experimental results until 2018, when the LHCb collaboration published a new measurement of the Ω_c^0 lifetime [5], followed by the lifetimes of the other baryons in 2019 [6]. These results altered the established hierarchy to:

$$\tau(\Xi_c^0) < \tau(\Lambda_c^+) < \tau(\Omega_c^0) < \tau(\Xi_c^+). \tag{2}$$

Notably, the measured lifetime of the Ω_c^0 was approximately four times larger than the previous world average [7], representing a discrepancy of around 7σ. Similarly, the Ξ_c^0 lifetime was determined to be larger by approximately 3σ. These significant discrepancies highlight the need for further measurements to clarify our understanding of heavy-flavour hadrons.

The analysis has recently been concluded, with a paper in preparation at the time of writing. As the initial work on this analysis formed the core of my Masters by Research thesis, it will not be discussed further in this thesis, and is mentioned here for completeness.

Chapter 1

The Standard Model and $b \to q\ell^+\ell^-$ processes

In this chapter an overview of the Standard Model (SM) of particle physics is given prior to a focused discussion on the specifics of $b \to q\ell^+\ell^-$ processes. Unless otherwise cited the discussion follows the structure of Ref. [8], with additional information from Ref. [9]. The chapter then progresses to explaining the theoretical and experimental motivations behind the angular analysis of $B^0 \to \rho^0\mu^+\mu^-$ decays which is the focus of this thesis. The structure of this discussion follows predominantly the reviews of the field in Refs. [10] and [11], with others referenced throughout.

1.1 The Standard Model of Particle Physics

Elementary particles are the building blocks of matter and their physical interactions. We categorise particles into two types, half integer spin particles, named *fermions* and integer spin particles, named *bosons*. Fermions are the constituents of matter, while bosons are the mediators of force. Additionally, the Higgs boson, although not a force mediator, acts by giving mass to particles through the Brout-Englert-Higgs mechanism mechanism.

There are four fundamental forces which govern the known universe: *gravity*, the *weak* force, the *strong* force, and the *electromagnetic* force. The SM explains how the weak, strong, and electromagnetic forces are mediated by bosons, which are summarised in Table 1.1. However, gravity, which is many orders of magnitude weaker than the other forces, is not included in the SM. Interactions involving the electromagnetic force are mediated by exchanges of *photons* (γ) between charged particles. The strong force is governed by *gluon* (g) exchange, binding the quarks into nucleons and nucleons into nuclei. Finally the weak force is mediated by three vector bosons known as the $W^{\pm}$, which are responsible for the flavour changing of particles, and the Z boson. The *Higgs* (H^0) boson has spin of zero and is responsible for the mass of the fermions and the $W^{\pm}$ and Z bosons.

Force	**Mediator(s)**	**Range** [m]	**Mass** [GeV/c^2]
Electromagnetic	γ	∞	0
Weak	$W^{\pm}, Z^0$	10^{-18}	80.4, 91.2
Strong	g	10^{-15}	0

Table 1.1: The fundamental forces described by the SM, including their mediating bosons, range, and corresponding boson mass.

All fermions exist as one of two types, *quarks* and *leptons*. Both of these groups contain six distinct *flavours* of particles, split into three pairs, or *generations*, with the first generation being the lightest and most stable. The up (u) and down (d) quarks make up the first generations, followed by the second generation charm (c) and strange (s), and the third generation of top (t) and bottom (b). The quarks are electrically charged and also have a *colour* charge, with three possible colour charges defined as red, green, and blue, and corresponding anti-colour charge. The bound-states of quarks are required to be colourless, a concept introduced by Gell-Mann in the quark model [12], which ensures that quark combinations are stable. Due to these charges the quarks experience interactions mediated by the weak, electromagnetic, and strong forces. The leptons are arranged similarly, with the first generation consisting of the well-known electron (e) and the neutral electron neutrino (ν_e), followed by the muon (μ) and is neutrino (ν_μ), and finally the tauon (τ) and its neutrino (ν_τ). The neutrinos have very small mass and are electrically neutral, while the e, μ, and τ are electrically charged, all leptons have no colour charge. Due to their lack of electric and colour charge, neutrinos only experience the effects of the weak force. The electric charges and masses of the leptons and quarks are given in Table 1.2.

Each spin-half particle has a corresponding anti-particle with opposite charge-related quantum numbers and the same mass. Anti-matter particles can often be distinguished experimentally from their matter counterparts by their opposite deflection in a magnetic field due to their opposite charge, though neutral anti-particles require other identification methods, such as their interaction behaviour.

Quarks are not to be found individually, they are observed in bound-states named hadrons. The most commonly observed hadrons are bound states of a quark and anti-quark pair, a type of *meson*, and three quarks, a type of *baryon*. Several measurements at the LHCb detector and B-factories [13–21] have given confirmation that there exists further exotic bound states with four (*tetra-*) and five (*penta-*) quarks.

Flavour	Charge	Mass [MeV/c^2]
Leptons		
e	-1	0.511 ± 0.000
ν_e	0	$< 0.8 \cdot 10^{-6}$
μ	-1	105.7 ± 0.000
ν_μ	0	< 0.19
τ	-1	1776.9 ± 0.1
ν_τ	0	< 18.2
Quarks		
u	$+2/3$	2.16 ± 0.07
d	$-1/3$	4.70 ± 0.07
c	$+2/3$	$(1.27 \pm 0.000) \cdot 10^3$
s	$-1/3$	93.5 ± 0.8
t	$+2/3$	$(172.52 \pm 0.33) \cdot 10^3$
b	$-1/3$	$(4.18 \pm 0.01) \cdot 10^3$

Table 1.2: Electric charges and messes of leptons and quarks in the Standard Model. The masses of quarks are approximate, as quarks are always observed in bound states [22].

1.1.1 Symmetries

The SM is a Quantum Field Theory (QFT) that describes particles as excitations of quantum fields. It is a gauge theory associated with the local gauge symmetry group

$$SU(3)_C \times SU(2)_I \times U(1)_Y, \tag{1.1}$$

under which it is invariant. The symmetry group $SU(3)_C$ corresponds to the colour charge (C) in Quantum Chromodynamics (QCD), which governs the strong interactions between quarks. The electroweak interactions, which unify the electromagnetic and weak forces, were developed by Glashow, Weinberg, and Salam [23–25]. This unification is described by the symmetry $SU(2)_I \times U(1)_Y$, which defines the electroweak (EW) force. The electric charge Q is related to the weak isospin I and weak hypercharge Y via the Gell-Mann–Nishijima relation

$$Q = I_3 + \frac{Y}{2}. \tag{1.2}$$

where I_3 is the third component of the weak isospin.

The SM requires that the transformations of fields be invariant under local gauge transformations associated with its symmetry groups, ensuring local gauge invariance. The electroweak symmetry is spontaneously broken at low energies. This breaking leads to the emergence of the electromagnetic force, described by the $U(1)$ symmetry, which is a part of Quantum Electrodynamics (QED). The Higgs mechanism is responsible for this

symmetry breaking, resulting in the generation of the Higgs boson and fermion masses.

The interactions of fields and particles within the SM can be described by Lagrangian density, which dictates the behaviours of each component of the SM:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{1.3}$$

The components of the SM Lagrangian separately describe interactions with the strong force via the QCD Lagrangian, $\mathcal{L}_{\text{QCD}}$; interactions with the unified weak and electromagnetic forces via the electroweak Lagrangian, $\mathcal{L}_{\text{EW}}$; the mechanism for giving the mass to the $W^{\pm}$ and Z bosons via the Higgs Lagrangian, $\mathcal{L}_{\text{Higgs}}$; and the generation of fermion masses through interactions with the Higgs field, $\mathcal{L}_{\text{Yukawa}}$. Within each of these Lagrangians, there are also terms that describe the behaviour of free particles, representing their dynamics in the absence of interactions.

1.1.1.1 Quantum Chromodynamics

The QFT associated with the strong interaction is QCD. The QCD Lagrangian is invariant under $SU(3)$ phase transformations defined by

$$\psi(x) \rightarrow \psi'(x) = e^{ig_s\alpha^a(x)T_a}\psi(x), \tag{1.4}$$

where T_a are the eight generators of the $SU(3)$ symmetry group, $\alpha^a(x)$ are the eight corresponding functions in space-time, and g_s is the QCD coupling constant. The generators are 3×3 matrices and thus an additional three degrees of freedom, known as colour, are required, such that

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix} \quad \text{and} \quad \psi^{\dagger} = (\psi_r^*,\, \psi_g^*,\, \psi_b^*). \tag{1.5}$$

where ψ_i are the quark fields associated with the three colour charges: red, green, and blue, respectively.

To retain the required local gauge invariance eight new tensor fields, $F^a_{\mu\nu}$, are introduced. These field tensors are given by

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f_{abc} A^b_\mu A^c_\nu. \tag{1.6}$$

The parameters f_{abc} are the commutation structure constants of the $SU(3)$ symmetry group. The index a of the gluon fields, A_μ, runs from 1-8, representing the eight kinds of gluons. The Lagrangian for the QCD processes within the SM is expressed as

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_{q,i}(i\gamma^\mu \delta_\mu \delta_{i,j} - g_s \gamma^\mu T^c_{i,j} A^c_\mu)\psi_{q,j} - \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a, \tag{1.7}$$

where the repeated indices are summed over. The quark-field spinors $\psi_{q,i}$ are related to a quark with flavour q. The indices i and j are colour indices which run from 1-3, representing the 3 colours. The $T^c_{i,j}$ are the Gell-Mann matrices which are generators of the $SU(3)$ group. The terms within the sum concern the kinematics of the quarks and the colour-changing interactions between two quarks and a gluon, the final term concerns gluon self-interactions.

1.1.1.2 Electroweak interactions

The electroweak theory unifies the interactions of the electromagnetic and weak forces under a single framework, described by the $SU(2)_I \times U(1)_Y$ gauge group. The forces were unified by Glashow, Weinberg, and Salam [23–25]. They showed that the two forces are aspects of the same unified electroweak framework.

Quantum Electrodynamics

The interactions between electrically charged particles are mediated by the massless photon and are described in the QED Lagrangian,

$$\mathcal{L}_{\text{QED}} = i\bar{\psi}\gamma^\mu \partial_\mu \psi + q\bar{\psi} A_\mu \psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \tag{1.8}$$

The first term describes the kinematics of the charged particle, the second term is related to the interactions between the charged particle and the photon, and the final term describes the kinematics of the photon. The interactions are described by the gauge group $U(1)$ with a single gauge field, A_μ. This electromagnetic field corresponds to the photon.

The weak force

The weak force, or weak nuclear force, is responsible for processes that change quark and lepton flavour, as well as other interactions, and is mediated by the $W^\pm$ and Z bosons. These interactions are described by an $SU(2)_L$ gauge group. $SU(2)_L$ refers to the local gauge symmetry acting on left-handed fermions (L). This symmetry is associated with three massless gauge fields which correspond to the weak isospin generators, $W^{1,2,3}$. The linear combination of W^1 and W^2 forms the weak charged current $W^\pm$ bosons,

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp i W^2_\mu). \tag{1.9}$$

The third field, W^3_μ, is neutral.

Electroweak unification

The interactions described by the electroweak symmetry group are governed by two coupling constants: g for the $SU(2)$ gauge group, and g' for the $U(1)_Y$ gauge group. The $SU(2)$ generators correspond to the weak isospin, T, and mediate the interactions involving the three gauge bosons, $W^{\pm}$ and Z, with coupling constant g. The generator of the $U(1)_Y$ group is the hypercharge, Y, has coupling constant g'. $W^{\pm}$ bosons only interact with left-handed particles, the field of which are doublets written as

$$\psi_L^{quark} = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad \psi_L^{lepton} = \begin{pmatrix} \nu_L \\ l_L \end{pmatrix}. \tag{1.10}$$

Here u_L and d_L represent the left-handed components of up-type and down-type quarks, respectively. In the lepton doublet, ν_L represents the left-handed neutrinos and l_L the left-handed charged leptons. The upper element of the vectors have weak isospin of $I_3 = \frac{1}{2}$, where I_3 is the third component of weak isospin, and the lower elements having a value of $I_3 = -\frac{1}{2}$. Right-handed fields form weak isospin singlets, where $I = I_3 = 0$,

$$\begin{aligned} &u_R, \quad c_R, \quad t_R \\ &d_R, \quad s_R, \quad b_R \\ &e_R, \quad \mu_R, \quad \tau_R, \end{aligned} \tag{1.11}$$

which are unaffected by $SU(2)$ transformations, and therefore do not couple to the gauge bosons. The $U(1)_Y$ symmetry has a gauge field B_μ. The gauge fields associated with $SU(2)$ and $U(1)_Y$, namely B_μ and W_μ^3, mix to give rise to the photon (A_μ) and the Z-boson fields. The fields are expressed as linear combinations of the gauge fields:

$$\begin{aligned} Z_\mu &= W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \\ A_\mu &= W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W, \end{aligned} \tag{1.12}$$

where θ_W is the weak mixing angle (or the Weinberg angle). The weak mixing angle is usually expressed as $\sin^2\theta_W$, which is determined to be 0.2235 ± 0.0001 [22]. The angle has been measured using different processes, enabling the determination of several properties of the Z-boson. The mass of the Z boson is related to the mass of the $W^{\pm}$ bosons via the relation

$$m_Z = \frac{m_W}{\cos\theta_W}. \tag{1.13}$$

The weak couplings of the bosons, g and g', are related to the definition of the electric charge by

$$e = g \sin\theta_W = g' \cos\theta_W. \tag{1.14}$$

The electroweak Lagrangian is expressed as

$$\mathcal{L}_{\text{EW}} = (D_\mu \psi)^\dagger (D^\mu \psi) - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu}, \tag{1.15}$$

where the term $-\frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu}$, describes the self-interactions of the W bosons, while $-\frac{1}{4} B_{\mu\nu} B^{\mu\nu}$ describes the dynamics of the B boson, which is not self-interacting. The index a runs over 1, 2, and 3. The first term represents the kinetic term for the fermions and their interactions with the gauge bosons via the covariant derivative,

$$D_\mu = \partial_\mu - i \frac{g'}{2} Y B_\mu - i \frac{g}{2} T_j W^j_\mu, \tag{1.16}$$

where T_j are the $SU(2)$ generators and j runs over 1, 2, and 3. The covariant derivative D_μ ensures the Lagrangian is gauge invariant and can be split into left- and right-handed terms, as the $SU(2)$ gauge bosons only couple to left-handed fermions,

$$D_\mu \psi_L = \partial_\mu - i \frac{g'}{2} Y B_\mu - i \frac{g}{2} T_j W^j_\mu, \tag{1.17}$$

$$D_\mu \psi_R = \partial_\mu - i \frac{g'}{2} Y B_\mu. \tag{1.18}$$

1.1.1.3 The Brout-Englert-Higgs mechanism

The Brout-Englert-Higgs mechanism explains how fundamental particles, particularly the $W^\pm$ and Z bosons, acquire mass in the SM. In the absence of the mechanism, unitarity is found to be violated in $W^+W^- \to W^+W^-$ scattering at centre-of-mass energies of approximately 1 TeV [26]. This violation is associated with the gauge bosons needing to be massive, if they were massless the probabilities of certain processes would exceed the limit of 1. To address this, the Brout-Englert-Higgs mechanism [27, 28] introduces the scalar Higgs boson, H^0, which restores unitarity by giving mass to the $W^\pm$ and Z bosons through spontaneous symmetry breaking.

The Higgs model consists of two complex scalar fields, ϕ. These fields represent the physical degrees of freedom in the Higgs field that appear after symmetry breaking, including the Higgs boson. The fields are arranged in a weak isospin doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.19}$$

with a corresponding Lagrangian of

$$\mathcal{L}_{H^0} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi). \tag{1.20}$$

The first term is the kinetic term which describes the kinetic energy of the Higgs field and its interactions with the gauge bosons. The covariant derivative, D_μ, introduces interactions between the Higgs field and the gauge bosons. The derivative is given by

$$D_\mu = \partial_\mu - i\frac{g}{2}\tau^a W_\mu^a - i\frac{g'}{2}B_\mu, \tag{1.21}$$

where τ^a are the Pauli matrices. The previously introduced gauge fields, W_μ^a and B_μ, are related to the $SU(2)$ and $U(1)_Y$, respectively. The second term of the Lagrangian, $V(\phi)$, describes the self-interactions of the Higgs boson which is responsible for spontaneous symmetry breaking. The potential $V(\phi)$ is the Higgs potential, given by

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{1.22}$$

In this expression, μ is a real parameter, and for spontaneous symmetry breaking to occur, μ^2 must be negative. This leads to a non-zero vacuum expectation value (VEV). The coupling constant λ is real and is required to be positive. For negative μ^2 there are an infinite set of minima satisfying the relation:

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda}. \tag{1.23}$$

A particular minimum is chosen, around which the field is expanded. This is expressed in the unitary gauge, which removes unphysical degrees of freedom. The Higgs doublet, including the Higgs scalar field $h(x)$, is:

$$\phi_{min}(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix}. \tag{1.24}$$

where ν is the VEV. The VEV is related to μ and λ by

$$\nu = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{1.25}$$

Inserting this field into the Lagrangian in Equation 1.20 yields the masses of the electroweak bosons from the Salam-Weinberg model. These masses are determined to be

$$m_W = \frac{gv}{2} \quad \text{and} \quad m_Z = \frac{\sqrt{g^2 + g'^2}\,\nu}{2}, \tag{1.26}$$

The photon remains massless as the $U(1)_Y$ gauge symmetry is unbroken. The Higgs boson is determined from $h(x)$ as an excitation with mass

$$m_H = \sqrt{2\lambda}\, v. \tag{1.27}$$

Therefore the Brout-Englert-Higgs mechanism ensures unitarity of electroweak interactions by providing mass to the $W^{\pm}$ and Z bosons, which is essential for the stability of the theory at high energies. The mechanism also predicts the existence of a scalar boson, discovered in 2012 by ATLAS and CMS [2, 3].

1.1.1.4 Yukawa interactions and fermion mass

Fermions acquire mass through the mechanism of Yukawa interactions with the Higgs field following electroweak symmetry breaking.

In the SM, left-handed fermions are placed in an $SU(2)$ doublet, ψ_L^{quark}, which was defined in Equation 1.10. Right-handed fermions are placed in $SU(2)$ singlets, q_r, where q represents the quark flavour, as defined in Equation 1.11. The combination $\overline{\psi}_L^{quark}\phi$ is invariant under the $SU(2)$ gauge transformations. When adding the right-handed fermion, q_r, the expression $\overline{\psi}_L^{quark}\phi q_r$ remains invariant under both $SU(2)$ and $U(1)_Y$, as does its Hermitian conjugate.

Fermion masses are generated from coupling to the Higgs field, requiring the Higgs doublet, ϕ, from Equation 1.19, and its conjugate form:

$$\phi_c = -i\sigma_2\phi^* = \begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -\nu + h(x) \\ 0 \end{pmatrix}. \tag{1.28}$$

Both forms are used in the associated Lagrangian. The Yukawa Lagrangian is thus defined for all fermions as:

$$\mathcal{L}_{\text{Yukawa}} = -g_f \left[\overline{\psi}_L^{quark} \phi q_r + (\overline{\psi}_L^{quark} \phi q_r)^\dagger \right] + g_f \left[\overline{\psi}_L^{quark} \phi_c q_r + (\overline{\psi}_L^{quark} \phi_c q_r)^\dagger \right]. \tag{1.29}$$

Including the Hermitian conjugate term ensures that the Yukawa Lagrangian is Hermitian, thus the physical requirement that the mass is real is maintained. The parameters g_f are the Yukawa couplings of fermions to the Higgs field:

$$g_f = \sqrt{2}\frac{m_f}{\nu}. \tag{1.30}$$

The parameter ν is the Higgs VEV, with a value of 246 GeV. This mechanism incorporates the masses of the charged fermions into the SM Lagrangian by substituting m_f with g_f. However, it does not provide a mechanism within the SM framework for providing neutrino mass, which remains a known limitation in the SM.

1.1.2 Flavour Physics

Flavour physics concerns the properties of quarks and leptons and their interactions with the $W^{\pm}$ and Z bosons, which mediate charged and neutral current interactions respectively. Understanding how the flavour of particles transforms through weak force interactions is crucial in understanding the SM. Quark flavour changing in weak interactions is parameterised by the Cabibbo, Kobayashi, and Maskawa [29, 30] (CKM) unitary matrix.

In Equation 1.30, it is shown that the mass of a fermion is directly proportional to it's Yukawa coupling, g_f. To account for mixing between generations of quarks, g_f generalises to 3×3 complex Yukawa matrices Y_u and Y_d, which describe the interactions of the left-handed doublets and right-handed quarks, respectively. Substituting these matrices into Equation 1.30 gives

$$\mathcal{L}_{\text{Yukawa}} = -\bar{Q}_L Y_u \phi_c u_R - \bar{Q}_L Y_d \phi d_R + \text{h.c.} \,, \tag{1.31}$$

where u_R and d_R are the right-handed up-type and down-type quark singlets, respectively, and Q_L is the left-handed quark doublet,

$$Q_L = \begin{pmatrix} u_l \\ d_l \end{pmatrix}. \tag{1.32}$$

Each element in the Yukawa matrices corresponds to an interaction between a left-handed quark from one generation and a right-handed from another. These interactions result in the mass terms

$$M_u = \frac{\nu}{\sqrt{2}} Y_u \quad \text{and} \quad M_d = \frac{\nu}{\sqrt{2}} Y_d, \tag{1.33}$$

where M_u and M_d are the mass matrices for up-type and down-type quarks, respectively. However, these matrices are in the flavour basis must be diagonalised to the mass basis, where the flavour basis refers to the representation in which quark weak interactions are defined, while the mass basis corresponds to the physical states with well-defined masses. This diagonalisation of M_u and M_d can be achieved by performing unitary transformations on the left- and right-handed components of the quarks:

$$Y_u = V_u^{\dagger} \hat{Y}_u U_u \quad \text{and} \quad Y_d = V_d^{\dagger} \hat{Y}_d U_d, \tag{1.34}$$

where, $\hat{Y}_u$ and $\hat{Y}_d$ are diagonal matrices containing the masses of the up-type and down-type quarks, respectively. The matrices V_u, U_u, V_d, and U_d are unitary transformations that diagonalise the Yukawa matrices, where V_u and V_d act on the left-handed quarks and U_u and U_d act on the right-handed quarks.

As the unitary transformations for up-type and down-type quarks act independently, the quark flavour basis does not align with the mass basis. This misalignment gives rise to the CKM matrix,

$$V_{\text{CKM}} = V_u^\dagger V_d, \tag{1.35}$$

which describes the probabilities of transition between the different quark flavours during weak decays mediated by the $W^\pm$ bosons. The CKM matrix is defined as:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.36}$$

This matrix encapsulates how quarks transform from one flavour to another. Each element, V_{ij}, represents the strength of the transition between the ith and jth quark flavours.

The standard parametrisation of the CKM matrix in terms of three angles and complex *CP*-violating phase, δ_{13}, is:

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{12}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & -c_{12}c_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.37}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$. The CKM matrix can also be expressed in the Wolfenstein parameterisation [31]:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.38}$$

The CKM matrix has a hierarchical structure, the transitions between quarks within the same generation are much stronger than those between different generations. The Wolfenstein parameterisation is an approximation of the CKM matrix which explicitly captures this nature using the expansion parameter λ, which has a value approximately equal to the sine of the Cabibbo angle, $\lambda \sim 0.23$ [32]. The remaining parameters are determined to be $A \sim 0.83$, $\rho \sim 0.16$, and $\eta \sim 0.35$.

Transitions within the same generations are determined by the diagonal elements, V_{ud}, V_{cs}, and V_{tb}. The hierarchical nature of the CKM matrix means that these transitions are strong, with magnitudes approximate to 1. Those transitions across generations are increasingly suppressed by powers of λ. Such transitions are Cabibbo-suppressed (CS).

The CKM matrix elements are fundamental parameters of the SM, i.e. they are not predicted by it. Their precise determination is therefore important and through a range of experiments the matrix elements have been determined. For example, the $|V_{ud}|$ element was determined by comparing nuclear β-decay and muon β-decay rates, and $|V_{us}|$ using semi-leptonic decays of kaons, $K_L^0 \to \pi^{\pm} e^{\mp} \nu_e$.

The final CKM element, V_{tb}, was first determined in 2005 by the CDF collaboration at FERMILAB [33], using a direct measurement of the cross section for single top quark production. As of May 2024 the fit results for the magnitudes of all nine CKM elements [34] are

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97435 \pm 0.00016 & 0.22501 \pm 0.00068 & 0.003732^{+0.00090}_{-0.00085} \\ 0.22487 \pm 0.00068 & 0.97349 \pm 0.00016 & 0.04183^{+0.00079}_{-0.00069} \\ 0.00858^{+0.00019}_{-0.00017} & 0.04111^{+0.00077}_{-0.00068} & 0.999118^{+0.000029}_{-0.000034} \end{pmatrix} \tag{1.39}$$

The unitarity of the CKM matrix can be represented as a triangle in the complex plane, with an area that measures *CP* violation in the Standard Model. These are known as unitarity triangles. One commonly used triangle is shown in Figure 1.1.1 and is derived from the relation

$$V_{ud}{V_{ub}}^* + V_{cd}{V_{cb}}^* + V_{td}{V_{tb}}^* = 0. \tag{1.40}$$

Dividing this relation by the term $V_{cd}{V_{cb}}^*$ normalises one of the triangle's sides to unity. The angles within the triangle, α, β, and γ (also referred to as ϕ_2, ϕ_1, and ϕ_3, respectively, in some conventions), are fundamental to the SM, and great effort is taken to measure them precisely. The angles are defined as

$$\alpha = \arg\left(-\frac{V_{td}{V_{tb}}^*}{V_{ud}{V_{ub}}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}{V_{cb}}^*}{V_{td}{V_{tb}}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}{V_{ub}}^*}{V_{cd}{V_{cb}}^*}\right). \tag{1.41}$$

This unitarity triangle directly involves CKM elements that govern B-meson decay, with the angles α, β, and γ corresponding to measurable *CP*-violating asymmetries. These angles can be extracted from time-dependent *CP* asymmetries in B^0 decays. In an angular analysis of certain decay processes, the interference between different helicity amplitudes provides information into *CP* violation, which will be explored in the next section.

A global fit to multiple measurements tests for consistency across each of the measurements and provides a unified test of the SM predictions. These fits are frequently updated with new results. An example of a triangle produced by the CKMfitter global fitting group [32] is shown in Figure 1.1.1.

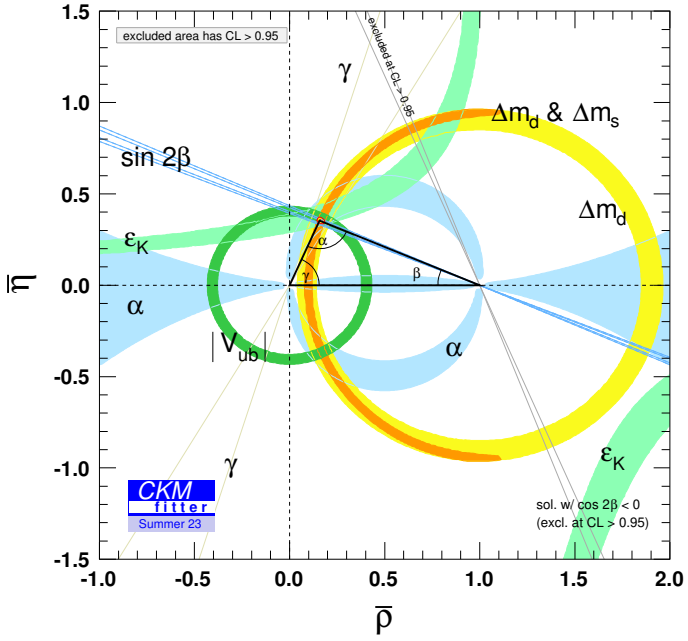


Figure 1.1.1: Constraints on CKM parameters in the complex $(\bar{\rho}, \bar{\eta})$ plane. The orange region of the global combination corresponds to 68% Confidence Level [32]

1.2 Examining $b \to q\ell^+\ell^-$ processes

This section introduces the theoretical concepts relevant to the quark transitions in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode. The angular observables, which are to be measured later in the thesis, are introduced. In Section 1.3, an overview of the status of experimental measurements will be given, which will build upon the theoretical foundation established here.

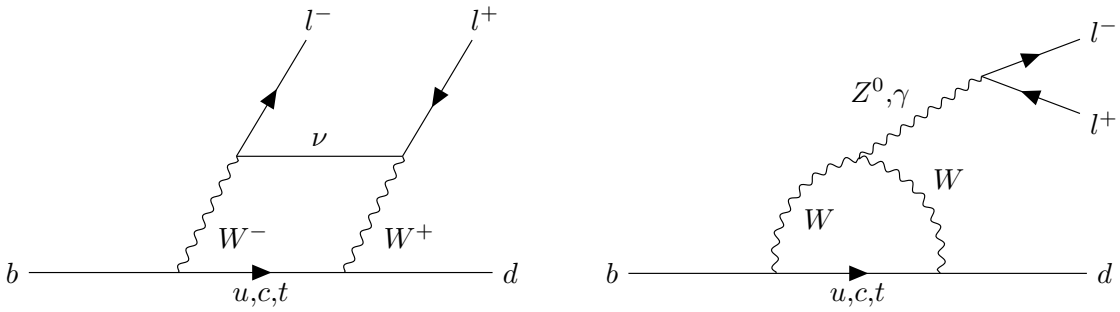
1.2.1 Flavour Changing Neutral Currents

Flavour-changing-neutral-current ($FCNC$) processes involve the transition of a quark to another flavour of the same type, either a down (d,s,b) to another down-type quark, or an up (u,c,t) to another up-type quark. Examples of $FCNC$ processes are $b \to q\ell^+\ell^-$ transitions, where a b quark changes flavour in a decay which produces a q-quark (where q represents either a d or an s quark) and an $\ell^+\ell^-$ pair in the final state. The focus of this thesis is the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, which is an example of a $b \to d\ell^+\ell^-$ quark transition. These $FCNC$ transitions, are forbidden at tree-level in the SM and are suppressed by the Glashow-Iliopoulos-Maiani (GIM) mechanism [35]. In the SM formalism, these transitions are allowed through loop level, occurring via electroweak penguin and box

Feynman diagrams [36]. Additional suppression arises due to these being cross-generation transitions involving at least one off-diagonal element of the CKM matrix. Consequently these decays experience suppression from the off-diagonal CKM matrix elements $\left|V_{qd}^{*}V_{qb}\right|$, leading to very rare $b\to d\ell^{+}\ell^{-}$ decays with branching fractions of $\mathcal{O}(10^{-8})$ [37], due to the small size of $|V_{td}|$. The Feynman diagrams of $FCNC$ $b\to d\ell^{+}\ell^{-}$ transitions are illustrated in Figure 1.2.1 for leading-order transitions.

Several examples of experimental measurements deviate from SM predictions in $b\to s\ell^{+}\ell^{-}$ transitions. These include inconsistencies observed in the angular observables of rare $B^{0}\to K^{*0}\mu^{+}\mu^{-}$ decays [1, 38], as well as branching fraction measurements in low q^{2} bins, which systematically fall below the SM predictions across various decay modes [39–42]. Numerous NP models have been proposed to explain these deviations These models can include new particles at the TeV scale such as the Z' boson and leptoquarks [43, 44], or formulated within the Higgs doublet model [45]. Feynman diagrams with some potential NP particles are shown in Figure 1.2.2. The existence of these models could account for observed branching fractions and angular observables which deviate from SM predictions.

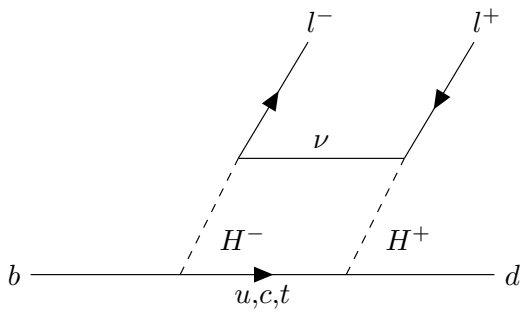
Rare $b\to q\ell^{+}\ell^{-}$ penguin decays are accessible using the datasets collected by the LHCb experiment, which has led to several publications on $b\to s$ transitions. The $b\to d$ process is less well studied, although both transitions involve the same concepts. The experimental results are discussed further in Section 1.3. Since the box-diagram is dominated by t-quark contributions, the relative abundance of each process can be predicted based on the sizes of the CKM elements that govern the $q\to t$ transitions, V_{td} and V_{ts}. The theoretical ratio of the relevant CKM elements is $|V_{td}/V_{ts}|^{2}\sim\frac{1}{25}$ [22], indicating that $b\to s$ decays are expected to be approximately 25 times more abundant than $b\to d$ decays.



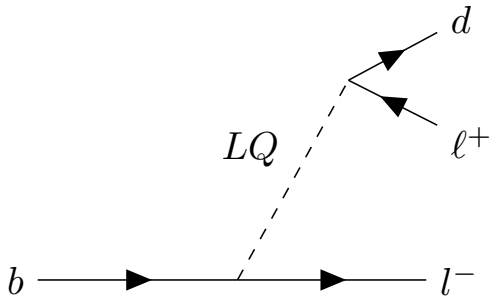
(a) Box diagram

(b) Electroweak penguin diagram

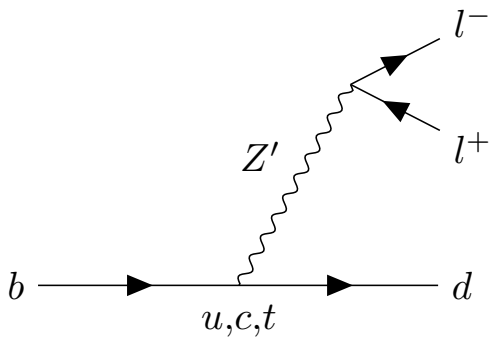
Figure 1.2.1: Feynman diagrams of leading-order $b\to d\ell^{+}\ell^{-}$ quark transitions allowed in the SM. Figure (a) is a box diagram and (b) is an electroweak penguin diagram containing either a Z boson or photon.



(a) Higgs Doublet



(b) Leptoquark (LQ)



(c) Z'

Figure 1.2.2: Feynman diagrams of $b \to d\ell^+\ell^-$ quark transitions involving potential NP particles. Figure (a) involves a Higgs doublet box, (b) and (c) show a penguin loop replaced with the Z' and Leptoquark particles respectively.

1.2.2 Effective Field Theory

Effective theories are useful for application to a specific problem at a specific energy scale. There are two energy scales one must consider when discussing, for example, the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay: the electroweak force, which is governed by the $W^\pm$ boson mass of approximately 80 GeV, and the QCD effects, which are governed by Λ_{QCD}, at approximately 0.2 GeV. The difference in these energy scales leads to a decoupling between the high-energy electroweak interactions and the low-energy QCD physics.

The starting point for the phenomenology of $b \to d\ell^+\ell^-$ transitions is the effective weak Hamiltonian [46],

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{td}^* \sum_i C_i(\mu)O_i(\mu), \tag{1.42}$$

where G_F is the Fermi constant, V_{tq} the CKM matrix element, C_i the Wilson coefficients [47,

48], and O_i a set of local operators involving four-fermion interactions. This Hamiltonian holds for energy scales of $\mu \ll m(W)$. The elements V_{tq} and coefficients C_i describe the strength of the operators within the Hamiltonian.

In the SM, one typically shows Feynman diagrams governed by $W^{\pm}$, Z, and t quark exchanges rather than effective operators. Such diagrams correspond to short-distance scales of $\mathcal{O}(M_{W,Z,t})$. However, for hadrons with masses of $\mathcal{O}(M_{b,c,K})$ undergoing decay, point-like effective vertices are a good approximation. These vertices are represented by the operators O_i, with the associated Wilson coefficients C_i acting as the effective coupling constants [10].

For the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay, the local operators we consider are:

$$\begin{aligned} \mathcal{O}_7 &= \frac{e}{16\pi^2} m_b (\overline{q}\sigma^{\mu\nu} b_R) F_{\mu\nu}, & \mathcal{O}'_7 &= \frac{e}{16\pi^2} m_b (\overline{q}\sigma^{\mu\nu} b_L) F_{\mu\nu}, \\ \mathcal{O}_9 &= \frac{e}{16\pi^2} (\overline{q}\gamma_\mu b_L)(\bar{l}\gamma^\mu l), & \mathcal{O}'_9 &= \frac{e}{16\pi^2} (\overline{q}\gamma_\mu b_R)(\bar{l}\gamma^\mu l), \\ \mathcal{O}_{10} &= \frac{e}{16\pi^2} (\overline{q}\gamma_\mu b_L)(\bar{l}\gamma^\mu\gamma^5 l), & \mathcal{O}'_{10} &= \frac{e}{16\pi^2} (\overline{q}\gamma_\mu b_R)(\bar{l}\gamma^\mu\gamma^5 l). \end{aligned} \tag{1.43}$$

Here, the operators $\mathcal{O}$ are left-handed and $\mathcal{O}'$ are right-handed operators. In $\mathcal{O}_7^{(\prime)}$, we see the mass of the b quark, m_b, and $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The parameter $\overline{q}$ represents light quarks, and in the case of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay, this corresponds to a d. The $\mathcal{O}_7^{(\prime)}$ operators correspond to the electroweak penguin diagrams containing a photon, while the semi-leptonic $\mathcal{O}_9^{(\prime)}$ and $\mathcal{O}_{10}^{(\prime)}$ operators respectively correspond to vector and axial-vector operators. The terms b_L and b_R represent the left- and right-handed components of the b-quark field b, defined as

$$b_L = \frac{1}{2}(1-\gamma^5)b \quad \text{and} \quad b_R = \frac{1}{2}(1+\gamma^5)b, \tag{1.44}$$

where γ^5 is the chirality matrix which projects the left-handed and right-handed components of a field.

When considering the SM the C_i for the $\mathcal{O}'$ operators are suppressed, as the W bosons only couple to left-handed fermions, by a factor proportional to $\mathcal{O}(\frac{m_d}{m_b})$. This means that the contributions from the right-handed operators is small in the SM.

The energy scale of $\mu = m_b$ is the renormalisation scale of the effective Hamiltonian. The scale approximately corresponds to the energy at which the $b \to d\ell^+\ell^-$ processes occur, and at this scale, the heavier particles, specifically the $W^{\pm}$ and Z bosons, have been integrated out. At this renormalisation scale, the corresponding C_i in the SM [49] are given by

$$C_7^{SM} = 0.01, \qquad C_9^{SM} = -1.37, \qquad C_{10}^{SM} = 0.96, \tag{1.45}$$

which are the same for all $b \to d\ell^+\ell^-$ decays as the short-distance physics is the same for each decay mode. When one introduces a NP particle, modifications can appear in the definitions of the coefficients, $C_i^{(\prime)} = C_i^{(\prime)SM} + C_i^{(\prime)NP}$, or in additional operators. Constraints on the values of the coefficients, derived from experimental measurements, are discussed in Section 1.3.5.

1.2.3 Angular basis

The kinematics of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay can be fully expressed using three angular variables, $\vec{\Omega} = (\cos\theta_h, \cos\theta_l, \phi)$, and both the invariant mass of the dimuon and the dipion. Throughout the thesis the square of the dilepton mass is referred as q^2. The angle, θ_h, is defined as the angle between the π^+ (π^-) and the direction of the B^0 ($\overline{B}^0$) in the dipion rest frame. Similarly, θ_l is the angle between the μ^+ (μ^-) and the direction opposite to the B^0 ($\overline{B}^0$) in the dimuon rest frame. Lastly, ϕ is the angle between the planes of the dimuon pair and the dipion pair in the B^0 rest frame. This definition ensures that the angular observables transform appropriately under CP conjugation, relating the B^0 and $\overline{B}^0$ decay distributions. Experimentally, however, the self-conjugation of the final state prevents direct determination of whether the meson was initially a B^0 or $\overline{B}^0$ at the time of decay. Consequently, the angular definitions are inverted for half of the dataset. The topology of the decay, including the definitions of the angles, is illustrated in Figure 1.2.3, the mathematical definitions are given in Appendix A.

It is important to consider the time evolution of the B^0-$\overline{B}^0$ meson system. The B^0-$\overline{B}^0$ oscillations introduce time dependence in the decay amplitudes [50]. However, the relative difference between the decay widths of B^0 and $\overline{B}^0$ is sufficiently small, with $\Delta\Gamma_d/\Gamma_d = 0.1 \pm 1.0\,\text{ps}^{-1}$ [51], that the time-dependence can be considered negligible in the analysis presented in this thesis. Therefore, a time-integrated analysis is performed, combining the decay rates of B^0 and $\overline{B}^0$ without considering the time-dependent oscillations.

1.2.4 Differential Decay Rate

In the q^2 and $\vec{\Omega}$ paramtrisation, the differential decay rates of $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and its charge-conjugate decay are given by

$$\frac{d^4\Gamma}{dq^2 d\vec{\Omega}} = \sum_i I_i(q^2) f_i(\vec{\Omega}), \qquad \frac{d^4\bar{\Gamma}}{dq^2 d\vec{\Omega}} = \sum_i \bar{I}_i(q^2) f_i(\vec{\Omega}), \tag{1.46}$$

where I_i and $\bar{I}_i$ are angular coefficients dependent on q^2, and $f_i(\vec{\Omega})$ are combinations of spherical harmonics that describe the angular distribution of the final-state particles.

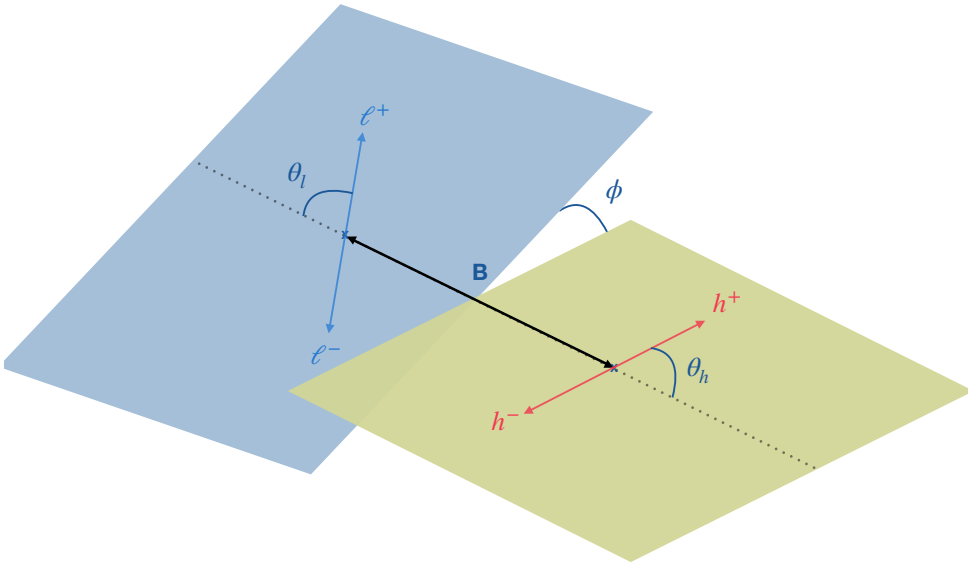


Figure 1.2.3: Angular topology of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays. The topology is described by three angles: θ_h, θ_l, and ϕ. The angle θ_h (θ_l) is defined as the angle between the hadrons (leptons) and the direction of (opposite to) the B^0 in the resonant rest frame. The angle ϕ is the angle between the planes of the dihadron and dilepton pairs in the B^0 rest frame.

These coefficients can be expressed as bilinear combinations of helicity amplitudes ${H_0}^{(L,R)}$, ${H_+}^{(L,R)}$, and ${H_-}^{(L,R)}$, as they originate from the interference of different helicity states in the decay. The bilinear structure arises because the decay rate is proportional to the modulus squared of the total amplitude, leading to terms that involve products of helicity amplitudes with different chiralities, where L and R refer to the left- and right-handed components of the dimuon system.

Alternatively, the angular coefficients can be represented in terms of six complex transversity amplitudes, $\mathcal{A}_{0,\parallel,\perp}^{(L,R)}$, corresponding to the transversity states of the dipion system $(0, \parallel, \perp)$. These transversity amplitudes have definite CP properties; under a CP transformation, $\mathcal{A}_0^{(L,R)}$ and $\mathcal{A}_\parallel^{(L,R)}$ are CP-even, while $\mathcal{A}_\perp^{(L,R)}$ is CP-odd. This means that $\mathcal{A}_0^{(L,R)}$ and $\mathcal{A}_\parallel^{(L,R)}$ correspond to final states that are symmetric under a CP-transformation, whereas $\mathcal{A}_\perp^{(L,R)}$ corresponds to a final state which is asymmetric under a CP-transformation. The helicity amplitudes are related to the transversity amplitudes as follows [52]:

$$\mathcal{A}_{\perp,\parallel}^{(L,R)} = \frac{{H_+}^{(L,R)} \mp {H_-}^{(L,R)}}{\sqrt{2}}, \quad \mathcal{A}_0{}^{(L,R)} \equiv {H_0}^{(L,R)}. \tag{1.47}$$

The angular coefficient are bilinear combinations of these transversity amplitudes. Their definitions, along with the corresponding angular functions, are provided in Table 1.3.

The first half of the table contains the coefficient relating to the P-wave states, which have orbital angular momentum L=1, and couple to vector mesons. Specifically, the predominant resonance in the dipion system of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays is the ρ^0 vector meson, which is an example of a P-wave state. The second half of the table lists coefficients for the S-wave states, which correspond to orbital angular momentum L=0. While the S-wave states exist in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays, they are treated as background contributions in this analysis. The terms in the table are approximations as they neglect the muon mass.

i	I_i	$f_i(\vec{\Omega})$
1s	$\frac{3}{4}(\lvert\mathcal{A}_\parallel^L\rvert^2 + \lvert\mathcal{A}_\perp^L\rvert^2 + \lvert\mathcal{A}_\parallel^R\rvert^2 + \lvert\mathcal{A}_\perp^R\rvert^2)$	$\sin^2\theta_h$
1c	$\lvert\mathcal{A}_0^L\rvert^2 + \lvert\mathcal{A}_0^R\rvert^2$	$\cos^2\theta_h$
2s	$\frac{1}{4}(\lvert\mathcal{A}_\parallel^L\rvert^2 - \lvert\mathcal{A}_\perp^L\rvert^2 + \lvert\mathcal{A}_\parallel^R\rvert^2 - \lvert\mathcal{A}_\perp^R\rvert^2)$	$\sin^2\theta_h\cos 2\theta_l$
2c	$-(\lvert\mathcal{A}_0^L\rvert^2 + \lvert\mathcal{A}_0^R\rvert^2)$	$\cos^2\theta_h\cos 2\theta_l$
3	$\frac{1}{2}(\lvert\mathcal{A}_\perp^L\rvert^2 - \lvert\mathcal{A}_\parallel^L\rvert^2 + \lvert\mathcal{A}_\perp^R\rvert^2 - \lvert\mathcal{A}_\parallel^R\rvert^2)$	$\sin^2\theta_h\sin^2\theta_l\cos 2\phi$
4	$1\,\mathrm{Re}(\mathcal{A}_0^L\mathcal{A}_\parallel^{L*} + \mathcal{A}_0^R\mathcal{A}_\parallel^{R*})$	$\sin 2\theta\sin 2\theta\cos\phi$
5	$\sqrt{2}\,\mathrm{Re}(\mathcal{A}_0^L\mathcal{A}_\perp^{L*} - \mathcal{A}_0^R\mathcal{A}_\perp^{R*})$	$\sin^2\theta\sin\theta\cos\phi$
6s	$2\,\mathrm{Re}(\mathcal{A}_\parallel^L\mathcal{A}_\perp^{L*} - \mathcal{A}_\parallel^R\mathcal{A}_\perp^{R*})$	$\sin^2\theta_h\cos\theta_l$
7	$\sqrt{2}\,\mathrm{Im}(\mathcal{A}_0^L\mathcal{A}_\parallel^{L*} - \mathcal{A}_0^R\mathcal{A}_\parallel^{R*})$	$\sin^2\theta\sin\theta\sin\phi$
8	$-1\,\mathrm{Im}(\mathcal{A}_0^L\mathcal{A}_\perp^{L*} + \mathcal{A}_0^R\mathcal{A}_\perp^{R*})$	$\sin 2\theta\sin 2\theta\sin\phi$
9	$\mathrm{Im}(\mathcal{A}_\parallel^{L*}\mathcal{A}_\perp^L + \mathcal{A}_\parallel^{R*}\mathcal{A}_\perp^R)$	$\sin^2\theta_h\sin^2\theta_l\sin 2\phi$
S1	$\frac{4}{\sqrt{3}}\mathrm{Re}(\mathcal{A}_0^{L*}\mathcal{A}_s^L + \mathcal{A}_0^{R*}\mathcal{A}_s^R)$	$\cos\theta_h\sin^2\theta_l$
S2	$\sqrt{\frac{2}{3}}\mathrm{Re}(\mathcal{A}_s^L\mathcal{A}_\parallel^{L*} + \mathcal{A}_s^R\mathcal{A}_\parallel^{R*})$	$\sin\theta_h\sin 2\theta_l\cos\phi$
S3	$2\sqrt{\frac{2}{3}}\mathrm{Re}(\mathcal{A}_s^L\mathcal{A}_\perp^{L*} - \mathcal{A}_s^R\mathcal{A}_\perp^{R*})$	$\sin\theta_h\sin\theta_l\cos\phi$
S4	$2\sqrt{\frac{2}{3}}\mathrm{Im}(\mathcal{A}_s^L\mathcal{A}_\parallel^{L*} - \mathcal{A}_s^R\mathcal{A}_\parallel^{R*})$	$\sin\theta_h\sin\theta_l\sin\phi$
S5	$\sqrt{\frac{2}{3}}\mathrm{Im}(\mathcal{A}_s^L\mathcal{A}_\perp^{L*} + \mathcal{A}_s^R\mathcal{A}_\perp^{R*})$	$\sin\theta_h\sin 2\theta_l\sin\phi$

Table 1.3: Angular observables I_i and their corresponding angular terms in the limit of zero $m(\mu)$. The terms in the second half of the table arise from the $\pi^+\pi^-$ S-wave contribution to the final state. The $\bar{I}_i$ coefficients are obtained by making the substitution $A \to \bar{A}$. Table taken from Ref. [1].

One further defines the observables as *CP* averages (S_i) and asymmetries (A_i),

$$S_i = \frac{I_i + \bar{I}_i}{(d\Gamma + d\bar{\Gamma})/dq^2}, \qquad A_i = \frac{I_i - \bar{I}_i}{(d\Gamma + d\bar{\Gamma})/dq^2}. \tag{1.48}$$

The number of P-wave observables are reduced due to the relationships of $I_{1s} = 3I_{2s}$ and $I_{1c} = -I_{2c}$. Another relation can be imposed, $\frac{3}{4}(2I_{1s} + I_{1c}) - \frac{1}{4}(2I_{2s} + I_{2c}) = 1$, which serves as a normalisation condition, ensuring the total differential decay rate is normalised to 1. Imposing these relations leaves a total of eight observables, down from eleven. The relationship of the observables to the longitudinal polarisation of the dipion system is used to define F_L, and the forward-backward asymmetry of the dimuon system A_{FB},

$$F_L = S_{1c} = 1 - \frac{4}{3}S_{1s}, \tag{1.49}$$

$$A_{FB} = \frac{3}{4}S_{6s}. \tag{1.50}$$

As discussed in Section 1.2.3, the time-dependent effects in the B^0-$\overline{B}^0$ system arise due to oscillations and the non-zero decay width difference $\Delta\Gamma_d$. However, since $\Delta\Gamma_d$ is consistent with zero within uncertainties, these effects are negligible, justifying the use of a time-integrated analysis in this thesis.

A decay is said to be self-conjugating when its final state is the same for both the particle and its antiparticle, which is the case for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, where the decay originates from either a B^0 or a $\overline{B}^0$. This symmetry prevents experimental determination of the initial flavour of the b-hadron, as both possibilities lead to the same observed final state. Because the decay is self-conjugating, the angular definitions inherently treat $\overline{B}^0$ decays as if they were B^0 decays, effectively inverting the angles for these events. Under a *CP* transformation, the angular variables transform as $\cos\theta_h \to -\cos\theta_h$, $\cos\theta_l \to -\cos\theta_l$, and $\phi \to \phi + \pi$.

By performing this inversion, we effectively average over both decay possibilities (B^0 and $\overline{B}^0$), meaning we access only one set of angular observables in the *CP*-averaged basis. Specifically, for each angular observable I_i, either its *CP*-averaged (S_i) or *CP*-asymmetric (A_i) term cancels out. This is a direct consequence of summing over both flavour possibilities in the dataset. Combining this cancellation with the definitions discussed above the *CP*-averaged angular distribution can therefore be written as

$$\begin{aligned}\frac{1}{d\Gamma/dq^2}\frac{d^4(\Gamma+\bar{\Gamma})}{dq^2 d\vec{\Omega}}\bigg|_P = \frac{9}{32\pi}\Bigg[&\frac{3}{4}(1-F_L)\sin^2\theta_h + F_L\cos^2\theta_h \\ &+ \frac{1}{4}(1-F_L)\sin^2\theta_h\cos 2\theta_l + F_L\cos^2\theta_h\cos 2\theta_l \\ &+ S_3\sin^2\theta_h\sin^2\theta_l\cos 2\phi + S_4\sin^2\theta_h\sin^2\theta_l\cos\phi \\ &+ A_5\sin^2\theta_h\sin\theta_l\cos\phi + A_6\sin^2\theta_h\cos\theta_l \\ &+ S_7\sin^2\theta_h\sin\theta_l\sin\phi + A_8\sin^2\theta_h\sin^2\theta_l\sin\phi \\ &+ A_9\sin^2\theta_h\sin^2\theta_l\sin 2\phi\Bigg].\end{aligned} \tag{1.51}$$

1.2.5 S-wave interference

The $\pi^+\pi^-$ final state decays predominantly via the $\rho^0 \to \pi^+\pi^-$ decay, which follows a P-wave configuration. However, additional dipion resonances contribute via S-wave configurations, the most significant of which in this analysis is the $f_0(500) \to \pi^+\pi^-$. These S-wave contributions introduce two complex amplitudes, $\mathcal{A}_s^L$ and $\mathcal{A}_s^R$, which lead to six additional angular terms, as listed in the bottom half of Table 1.3. The presence of S-wave interference modifies the P-wave differential decay rate from Equation 1.51, introducing a scaling factor for the S-wave fraction in the dataset (F_S) along with additional terms, written as

$$\begin{aligned}\frac{1}{d\Gamma/dq^2}\frac{d^4(\Gamma+\bar{\Gamma})}{dq^2 d\vec{\Omega}}\bigg|_{S+P} = (1-F_S)&\frac{1}{d\Gamma/dq^2}\frac{d^4\Gamma}{dq^2 d\vec{\Omega}}\bigg|_P \\ &+ \frac{3}{16\pi}F_S\sin^2\theta_l \\ &+ \frac{9}{32\pi}A_{S1}\sin^2\theta_l\cos\theta_h \\ &+ \frac{9}{32\pi}(S_{S2}\cos\phi + S_{S5}\sin\phi)\sin 2\theta_l\sin\theta_h \\ &+ \frac{9}{32\pi}(S_{S3}\cos\phi + S_{S4}\sin\phi)\sin\theta_l\sin\theta_h.\end{aligned} \tag{1.52}$$

1.2.6 Composition of $\pi^+\pi^-$ states

There is no existing measurement of the amplitudes of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode in the rare q^2 bins, however the amplitudes have been measured in the J/ψ q^2 bin [53]. The composition of the $\pi^+\pi^-$ distribution is determined, as are the relative proportions of the resonances belonging to different spin states. The different components in the $\pi^+\pi^-$ distribution are shown in Figure 1.2.4.

There are a number of resonances in the P-wave configuration, comprising approximately 73% of the distribution. The S-wave component comes from the $f_0(500)$ resonance, which accounts for about 20% of the region. The remaining contributions come from higher spin states. There is also a component coming from $K^0_s \to \pi^+\pi^-$ decays. This is an example of a $b \to s$ transition, and is therefore a background in the angular analysis. The K^0_s has a long lifetime, $(89.54 \pm 0.04)ps$, when compared to the lifetime of the B^0 mesons, of $(1.517 \pm 0.004)ps$. The majority of the K^0_s decays are therefore not going to be mistaken for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal, however a small amount will decay quickly and appear in the data.

Due to the proximity of the ρ^0 (770) and $\omega(782)$ P-wave states in mass, and their shared angular momentum quantum numbers, there is interference between the two resonances. The interference modifies the decay rates of the interfering components. The measurements in this thesis relate to the P-wave observables in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode as a whole, with no discrimination between the different P-wave states. Due to the limited sample size and low statistics, a detailed determination of the effects of this interference was not possible, and as such, these effects were not included in the analysis. Consequently, the effective statistics may be slightly reduced, but the contribution of the higher-spin states is negligible within the current sample size, in the fit region of $m(\pi^+\pi^-) \in [620, 920]\,\mathrm{MeV}/c^2$, as discussed further in Section 3.4.

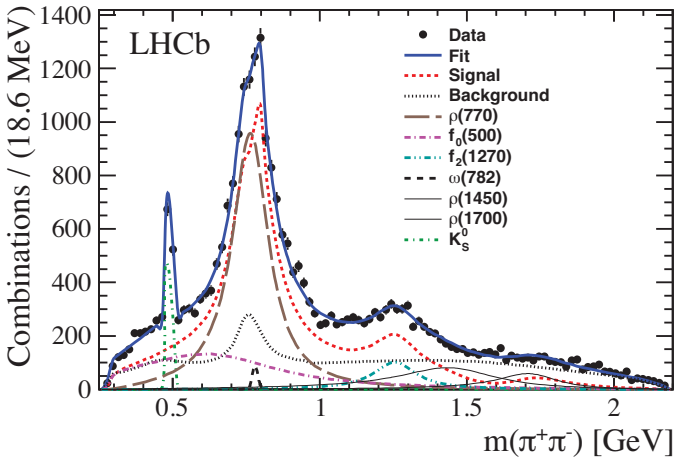


Figure 1.2.4: Composition of the $\pi^+\pi^-$ distribution in $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays [53].

1.3 Status of experimental measurements

There have been several measurements of angular observables in $b \to s\ell^+\ell^-$ processes performed by the LHCb collaboration, and others, in recent years. Some of these measurements have shown tensions with the SM predictions which have prompted further searches with larger datasets and analyses of more decay modes.

Three predominant types of analysis are performed on $b \to q\ell^+\ell^-$ decay processes. There are several different observables which can be measured that provide access to Wilson coefficient values. The measurements of differential branching fractions, angular distributions, and asymmetries in lepton-flavour production are all summarised below. The three analysis types are listed in order of increasing experimental complexity, and decreasing SM prediction uncertainty. Measurements using $b \to s\ell^+\ell^-$ processes are given followed by a discussion on $b \to d\ell^+\ell^-$ results and ending with a global view of the constraints provided by experimental results.

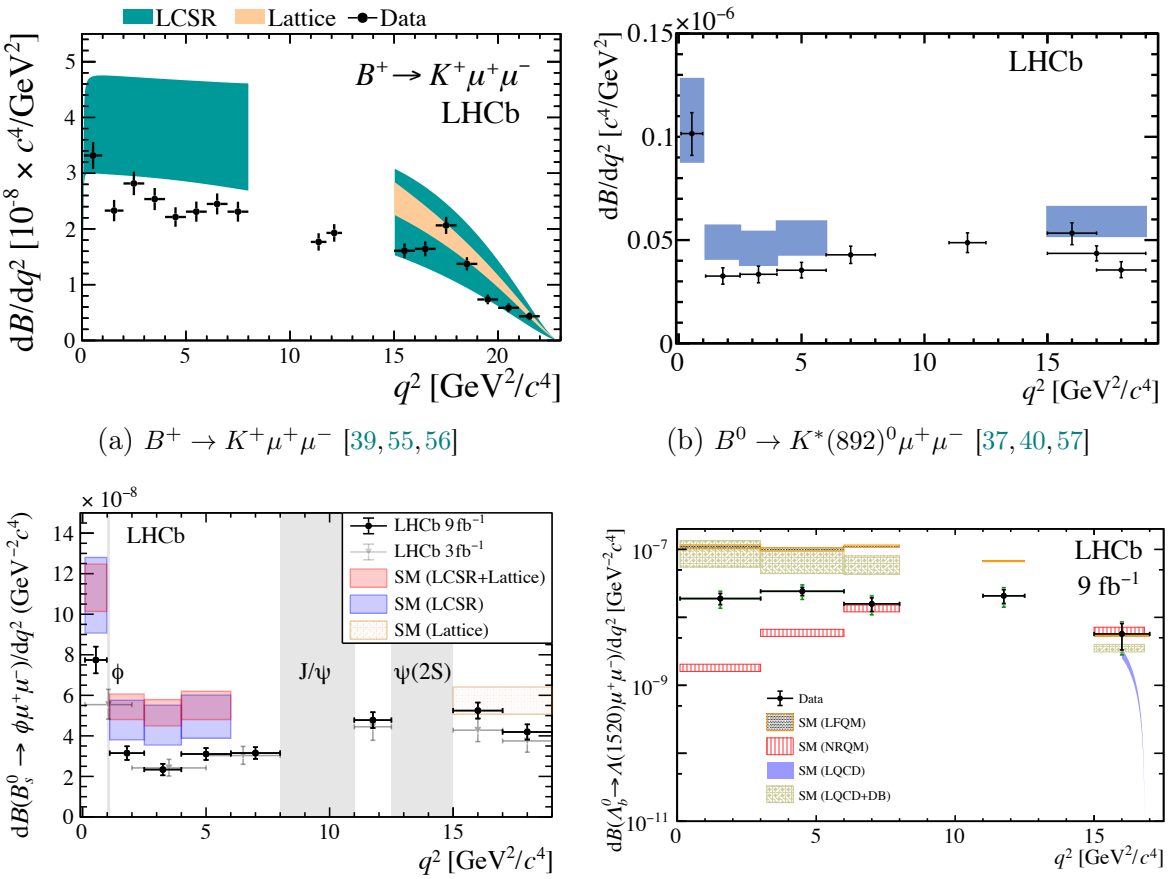
1.3.1 Differential Branching Fractions

Difficulties in the determination of differential branching fractions lead to large uncertainties in the theoretical predictions. Contributions from hadronic uncertainties such as from charm-loops [54] add to this problem. Experimentally these measurements are more straight forward. The current status of the measurements include several $b \to s\ell^+\ell^-$ results from the LHCb collaboration [39–42] is shown in Figure 1.3.1. The measurements are performed in bins of the dilepton mass squared (q^2). There are differences in the low-q^2 bins between the measurements of the $B^+ \to K^+\mu^+\mu^-$ (Figure 1.3.1a) and $B^0 \to K^*(892)^0\mu^+\mu^-$ (Figure 1.3.1b). These differences arise due to the spin of the hadronic system affecting the branching fraction in this low-q^2 region, the contribution is enhanced in the decay with the $K^*(892)^0$ vector meson resonance. Several of these experimental measurements lie systematically lower than the theoretical SM predictions in the low q^2 bins, suggesting the presence of unknown physics.

1.3.2 Angular observables

Angular distributions provide several observables, such as those introduced in Section 1.2.3, that are sensitive to Wilson coefficients. The datasets for various $b \to s\ell^+\ell^-$ decay modes available to the LHCb collaboration have been substantial enough to enable measurements of these observables in recent years. However, to date, no previous $b \to d\ell^+\ell^-$ angular analysis has been published.

For example, several measurements of angular observables have been performed using



(a) $B^+ \to K^+\mu^+\mu^-$ [39, 55, 56]

(b) $B^0 \to K^*(892)^0\mu^+\mu^-$ [37, 40, 57]

(c) $B_s^0 \to \phi\mu^+\mu^-$ [37, 41, 58–60]

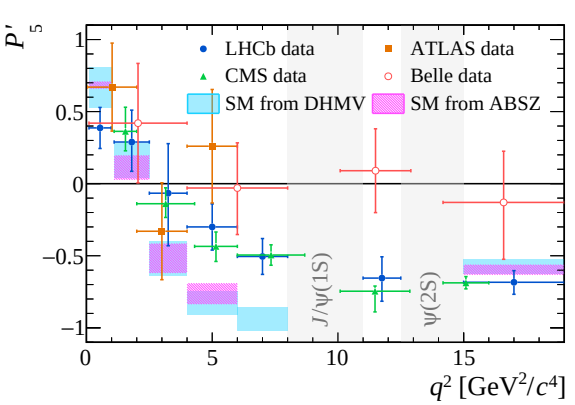
(d) $\Lambda_b^0 \to \Lambda(1520)\mu^+\mu^-$ [42, 61–64]

Figure 1.3.1: Differential branching fraction results and theoretical predictions.

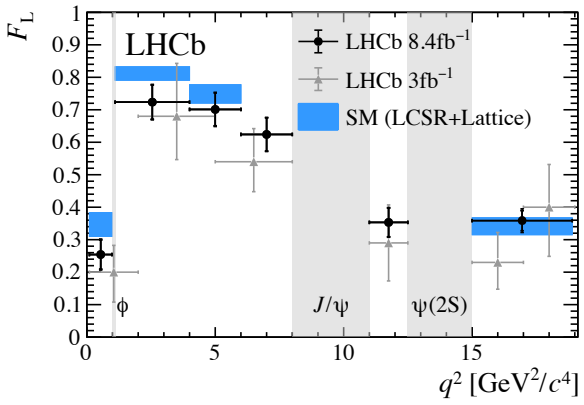
the $B^0 \to K^*(892)^0\mu^+\mu^-$ decay mode by the LHCb collaboration [1,38], ATLAS [65], CMS [66], and Belle [67]. These analyses have included measurements of the forward-backward asymmetry (A_{FB}) and optimised observables [68] that reduce theoretical uncertainties. By taking a ratio of certain observables, these uncertainties can be significantly reduced. For example, taking an appropriate ratio of the S_5 and F_L observables yields the P_5' observable with reduced theoretical uncertainties:

$$P_5' = \frac{S_5}{\sqrt{F_L(1-F_L)}}. \tag{1.53}$$

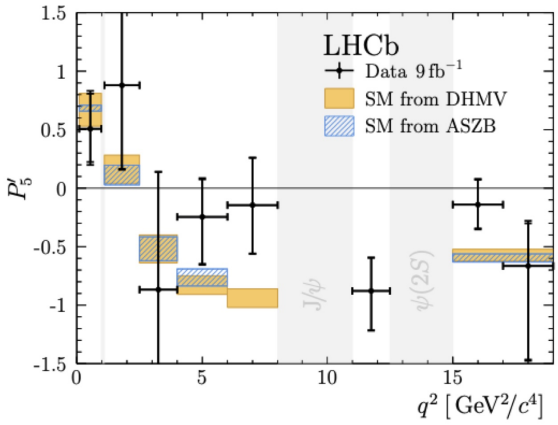
The LHCb collaboration has performed multiple $b \to s\ell^+\ell^-$ angular analyses to date, several of which have shown tensions with SM predictions. Figure 1.3.2 shows a selection of results where tensions are observed. For example, Figure 1.3.2a shows the LHCb measurement of P_5' [1] across multiple q^2 bins, where a tension of 3.4σ was observed over two q^2 bins. Shown in the same figure are the aforementioned results of measurements by ATLAS, Belle,and CMS, all showing similar tensions with the SM predictions [58, 69].



(a) $B^0 \to K^*(892)^0\mu^+\mu^-$ [1, 38, 58, 65–67, 69]



(b) $B_s^0 \to \phi\mu^+\mu^-$ [37, 59, 60, 70]



(c) $B^+ \to K^{*+}\mu^+\mu^-$ [58, 71–73]

Figure 1.3.2: Angular observables measurements and theoretical predictions.

These tensions raise the possibility of NP effects, potentially via modifications to the Wilson coefficients C_9 or C_{10}.

1.3.3 Lepton flavour universality ratios

Lepton flavour universality (LFU) ratios are theoretically cleaner than predictions of branching fractions and angular observables as they benefit from the cancellation of form factors and hadronic uncertainties. When measuring the ratio of electronic to muonic decay modes, the hadronic uncertainties cancel entirely. Any remaining difference is attributable to final state radiation, which can give a maximum deviation of about 1% [74].

In the SM the coupling of leptons to gauge boson is universal. Experimental measurements provide useful tests for NP in $b \to q\ell^+\ell^-$ decays. The LFU ratio for a b-hadron (H_b) decaying into an s-hadron (H_s) and two leptons is defined as:

$$\mathcal{R}_{H_s} = \frac{\int_{q^2_{\min}}^{q^2_{\max}} \frac{d\Gamma(H_b \to H_s \mu^+ \mu^-)}{q^2} dq^2}{\int_{q^2_{\min}}^{q^2_{\max}} \frac{d\Gamma(H_b \to H_s e^+ e^-)}{q^2} dq^2}, \tag{1.54}$$

where q^2 is the squared invariant mass of the dilepton system. At low q^2 values, approaching the kinematic limit of the muon mass at $q^2 = 4m_\mu^2 \sim 0.045\,\mathrm{GeV}^2/c^4$, the LFU ratio is predicted to be slightly smaller than one due to the constrained phase space [75]. Above values of $q^2 = 0.1\,\mathrm{GeV}^2/c^4$, this effect is negligible and the prediction is unity in the SM.

The LHCb collaboration has produced multiple measurements of LFU ratios, notably the most precise measurements of the $\mathcal{R}_K$ and $\mathcal{R}_{K^*}$ [76, 77], using $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decay modes respectively. Tensions with the standard model prediction emerged in LHCb LHCb measurements, but a reevaluation of systematic uncertainties in the 2022 LHCb results [74, 78] led to measurements that are now compatible with the SM to within $\sim 0.2\sigma$, as shown in Fig. 1.3.3.

Other results include the LHCb measurement using the $\Lambda_b^0 \to pK^-\ell^+\ell^-$, in agreement to one standard deviation with the SM prediction, [79] as well as the Belle [80–82] and BaBar [83] results which are also compatible with the SM. Although the tension in $\mathcal{R}_K$ and $\mathcal{R}_{K^*}$ has been resolved, further analyses using varied decay modes and increased dataset sizes remain crucial for testing lepton flavour universality with higher precision and exploring potential deviations in other observables.

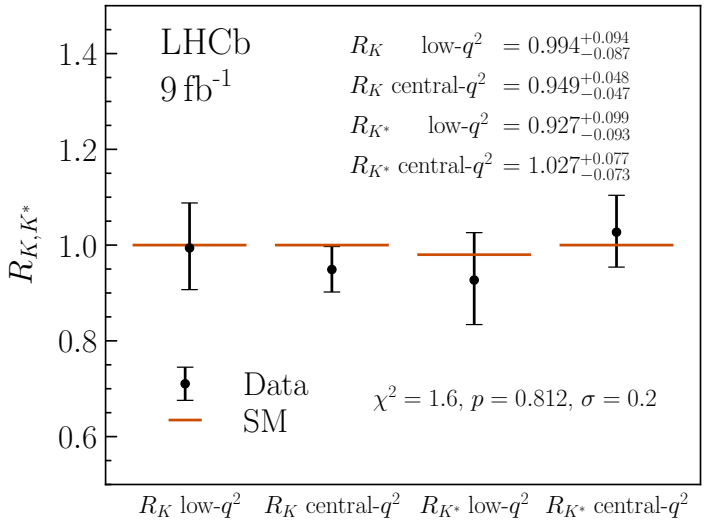


Figure 1.3.3: Measured values of LFU observables in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays and their overall compatibility with the SM [78].

1.3.4 Status of $b \to d\ell^+\ell^-$ measurements

There have been relatively few experimental measurements of $b \to d\ell^+\ell^-$ processes compared to the $b \to s\ell^+\ell^-$. As discussed, the branching ratios for $b \to d\ell^+\ell^-$ processes are suppressed compared to $b \to s\ell^+\ell^-$ transitions due to the additional Cabibbo suppression in $b \to d$ transitions. Consequently, these decays are even rarer, requiring larger datasets for experimental study.

The first measurement of a semileptonic $b \to d$ transition was the $B^+ \to \pi^+\mu^+\mu^-$ branching fraction in 2012 by the LHCb collaboration [84], and the ratio of CKM elements $|V_{td}/V_{ts}| = 0.24^{+0.05}_{-0.04}$. This was followed by the measurement of the differential decay distribution in the same decay mode [85], shown in Figure 1.3.4. Thus far there have been a few other measurements, such as the upper bound on the branching fraction of $B^0 \to \mu^+\mu^-$ decay mode [86], along with the first evidences of the decay modes $B_s^0 \to \overline{K}^{*0}\mu^+\mu^-$ [87] and $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ [88], by the LHCb collaboration.

The branching fraction of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decay modes have been measured using the Run I, $3\,\text{fb}^{-1}$ dataset, from LHCb [89]. The theoretical prediction of the branching fraction is of $\mathcal{O}(10^{-8})$ [90, 91], the analysis measured the branching fraction, at a statistical significance of 4.8σ, to be

$$\mathcal{B}(B^0 \to \pi^+\pi^-\mu^+\mu^-) = (2.11 \pm 0.51(\text{stat}) \pm 0.15(\text{syst}) \pm 0.16(\text{norm})) \times 10^{-8}. \quad (1.55)$$

The analysis also confirmed the first observation of the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, an example of a $b \to s\ell^+\ell^-$ process, which is another important probe for NP due to its sensitivity to FCNCs. The branching fraction measured, with a statistical significance of 7.2σ, to be

$$\mathcal{B}(B_s^0 \to \pi^+\pi^-\mu^+\mu^-) = (8.6 \pm 1.5(\text{stat}) \pm 0.7(\text{syst}) \pm 0.7(\text{norm})) \times 10^{-8}. \quad (1.56)$$

1.3.5 Global fits

Although tensions with the SM have been observed from $b \to s\ell^+\ell^-$ processes, to date none are significant enough to rule out the SM. By using the Operator Product Expansion (OPE) [95] formalism one can convert these results into constraints on the corresponding NP Wilson coefficients,

$$C_i^{\text{NP}} = C_i - C_i^{\text{SM}}. \quad (1.57)$$

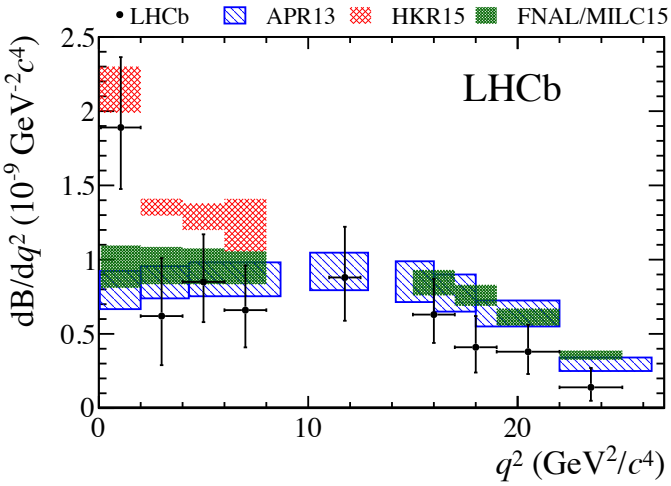


Figure 1.3.4: Experimental measurements of $\mathcal{B}(B^+ \to \pi^+\mu^+\mu^-)$ [85], in bins of q^2. This decay is an example of a $b \to d\ell^+\ell^-$ process. Also shown are SM predictions from APR13 [92], HKR15 [93], and from lattice QCD calculations from FNAL/MILC15 [94]. The predictions from the HKR15 group differ from the others in low q^2 due to the inclusion of light quark resonances, such as contributions from the $B^+ \to \pi^+\rho^0$ decay mode, where ρ^0 decays to either a $\mu^+\mu^-$ or $\pi^+\pi^-$ final state. The branching fraction is determined to be $\mathcal{B}(B^+ \to \pi^+\mu^+\mu^-) = (1.83 \pm 0.24 \pm 0.05) \times 10^{-8}$.

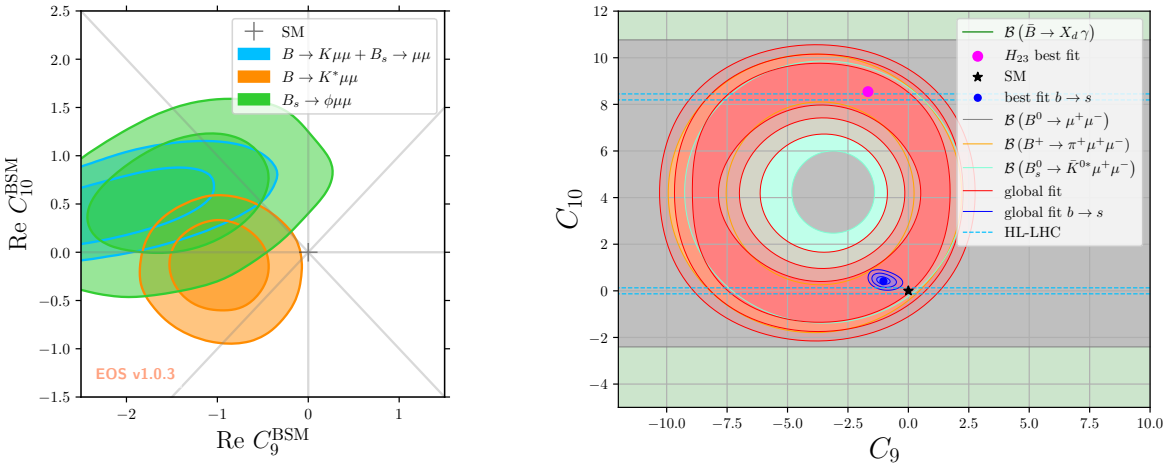
In the $b \to q\ell^+\ell^-$ processes the vector coupling strength (C_9) and the axial vector coupling strength (C_{10}) are relevant. Shifts in the two Wilson coefficients can help to explain the tension between predictions and data. These shifts are determined using a global fit to compare measurements of LFU ratios, angular observables, and branching fractions. A theoretical model uses these measurements as constraints and employs statistical sampling to produce likelihood contours for the parameter phase-space. Each new result refines the constraints and with the increased precision the effect of NP can be determined.

Multiple independent global fits [72,96–98] are performed by different groups, each using their own statistical frameworks and selections of experimental results. One such group has studied the aforementioned deviations in $b \to s\ell^+\ell^-$ differential branching fractions [72], incorporating new theoretical insights. Combined with various angular observables and LFU measurements, tensions between the measurements and SM predictions of 2.7σ and 2.6σ were obtained for the $B^0 \to K^*(892)^0\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ BFs respectively. The fits are performed simultaneously in the C_9 and C_{10} Wilson coefficients. The point $(Re\ C_9^{\mathrm{BSM}}, Re\ C_{10}^{\mathrm{BSM}}) = (-1.0, +0.4)$ is compatible with all decay channels at the $\sim 1\sigma$ level. The global fit presented in this study is shown in Figure 1.3.5a.

For the measured parameters concerning $b \to d\ell^+\ell^-$ processes, Figure 1.3.5b shows that the contours resulting from the global fits to the C_9 vs C_{10} Wilson coefficients [49] are less constrained compared to the $b \to s\ell^+\ell^-$ case. This figure presents contours from

fits to semileptonic four-fermion operators, where the shapes of the contours are affected by the experimental results. for instance, the branching fraction of the $B_s^0 \to \mu^+\mu^-$ decay mode is proportional to $\left|{C_{10}}^2\right|$, while $\mathcal{B}(B^+ \to \pi^+\mu^+\mu^-) \propto \left|{C_9}^2\right| + \left|{C_{10}}^2\right|$. Additionally, results from angular observables are sensitive to C_9-C_{10} interference.

Further results from this angular analysis, along with future studies, will improve the constraints. However, there is a greater deal more results required before the constraints on the Wilson coefficients reaches the precision of the $b \to s\ell^+\ell^-$ processes, shown in blue in Figure 1.3.5b. With more results from LHCb, Belle II, and eventually HL-LHC, the uncertainties on the global fits will reduce, and the nature of potential NP processes will become clear.



(a) (b)

Figure 1.3.5: Global fit in C_9 and C_{10} space for (a) $b \to s\ell^+\ell^-$ [72] and (b) $b \to d\ell^+\ell^-$ processes [49]. The shaded regions represent the parameter space consistent with current experimental data. This global fit aids in the understanding of contributions of NP to rare decays.

Chapter 2

The LHCb Experiment

This chapter introduces the CERN accelerator complex and the LHCb experiment, located on the Large Hadron Collider (LHC) [99]. The LHCb detector provides the data for this analysis and is aimed at the study of b-physics. The detector and its sub-detectors are described in detail, with information primarily drawn from the initial design report and performance report [100, 101].

2.1 CERN accelerator complex

The LHC [99] is a particle accelerator at the European Organisation for Nuclear Research (CERN) accelerator complex located on the border between France and Switzerland, close to Geneva. It is a circular accelerator with a circumference of 27 kilometres. It consists of superconducting magnets kept at an ultrahigh vacuum. The LHC is designed to propel two beams of particles, usually protons, close to the speed of light in opposite directions. They are collided at four locations on the ring, at which are located four major experiments. It is the largest, and most powerful, human-produced accelerator built thus far and the latest in a series of accelerators located at CERN.

CERN has a rich history in accelerator physics, beginning in 1957 with the 600 MeV Synchrocyclotron (SC) which provided the beams for CERN's first particle physics experiments and the discovery of the pion in 1958 [102] via the $\pi^- \to e^+\nu$ decay, mere hours after the first particles were accelerated [103].

The LHC is the latest in the line of CERN's large accelerators, it is fed the particles which it accelerates to the desired energies by a series of accelerators as shown in Fig 2.1.1. Protons are obtained from a bottle of hydrogen gas and, until 2018, accelerated by the Linear accelerator 2 (Linac2) which has been used as the starting accelerator for protons at CERN for 40 years, superseded by the Linac4. At the Linac hydrogen is passed through an electric field where electrons are stripped away. By the end of the accelerator the protons reached an energy of 50 MeV.

The protons are injected into the PS Booster, a series of four superimposed synchrotron rings which accelerate the protons to 2 GeV for injection into the PS. The PS subsequently accelerates the protons to 26 GeV before the SPS accelerates them further to energies of 450 GeV before finally being received by the LHC [104].

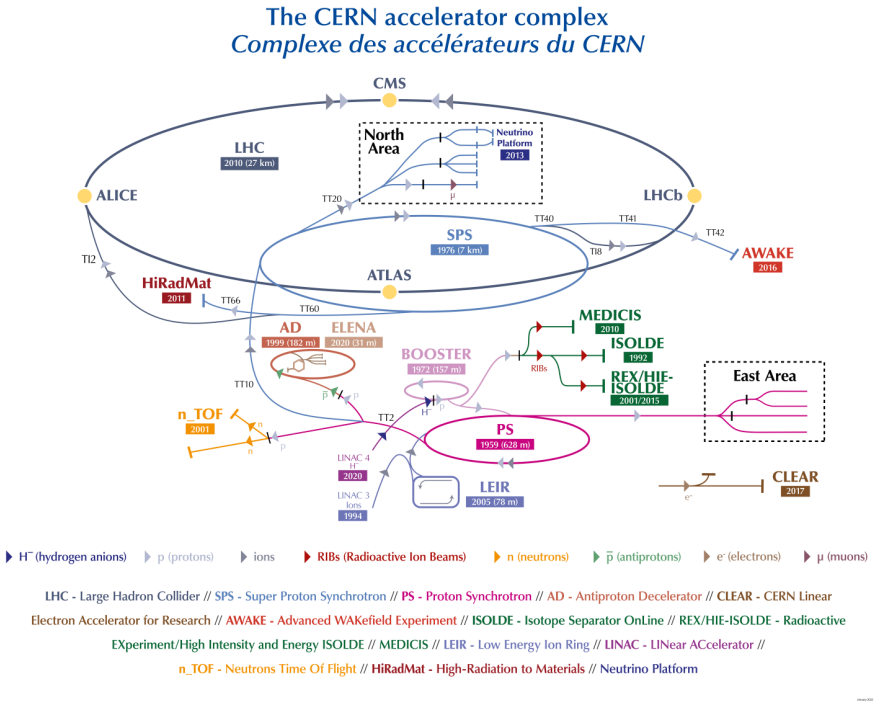


Figure 2.1.1: CERN accelerator complex [105].

The LHC was conceived in the early 1980's during the construction of the Large Electron-Positron (LEP) collider. In 1994, after many years of research and development work the CERN Council approved the construction of the LHC in the same tunnel as LEP.

The accelerator completed commissioning in time for the first data-taking period, relevant to this thesis, which took place between 2011-2012 and is known as Run I. This was followed by a long shutdown period, and then Run II from 2015 to 2018. Data-taking for Run III began in 2022, however this is outside the scope of this thesis.

The LHC ring is split into eight arcs, each containing 154 superconducting dipole magnets (for a total of 1232, at a strength of over 8 T each), used for bending the beams, and eight straight insertion sections, consisting of straight sections and book-ended by transition regions. These transition regions consist of four interaction points and areas for the beam dump and bean cleanup, shown in the LHC schematic in Fig 2.1.2. Additionally to these dipole magnets there are quadrupole, sextupole, octupole and decapole magnets which are used to stabilise the beam.

The charged particles enter the ring and are accelerated using superconducting radiofrequency (RF) cavities, operating at 400 MHz. The RF cavities also have the role of keeping protons bunched tightly, with each proton bunch containing approximately $\mathcal{O}(10^{11})$ protons (in Run I). There is a minimum separation between bunches of protons of 25 ns and the LHC is designed to accommodate at most 2808 bunches within each beam. Keeping the bunches as tight as possible allows for maximising the number of collisions achievable. The RF cavities require cooling to 4.5K which is achieved using super fluid helium at 1.9K.

The maximum energy obtainable by the LHC is primarily related to the size of the ring and the strength of the magnetic fields used to steer the particles. The energy is proportional to the ring radius and the strength of the magnets, which bend the paths of the particles, maintaining their circular trajectory. The design of the LHC constrains the maximum energy per beam to approximately 7 TeV.

When colliding two bunches of n_x particles the two key metrics for assessing the performance of the collisions are: the centre-of-mass energy, which determines the types of particles which can be studied or discovered [106] and the instantaneous luminosity ($\mathcal{L}$), which determines the event rate. For a given process the number of interactions (N) is determined as the luminosity integrated over the lifetime of the machine operation and the cross section (σ), related to the probability for interaction [107]. The yield of a given process is given by:

$$N = \sigma \int \mathcal{L}(t)dt. \tag{2.1}$$

The integrated luminosity therefore needs to be determined to convert the number of events to a cross-section, this is done using the parameters of the accelerated beams. The instantaneous luminosity is defined, presuming Gaussian beam profiles and head-on collisions, as

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}, \tag{2.2}$$

where σ_x and σ_y are the beam widths, n_1 and n_2 the number of particles in each bunch, and f the collision frequency. In the LHC, the collision frequency is 40 MHz, although not all of the bunches are filled, resulting in a visible rate of approximately 30 MHz.

At four of the eight interaction points, there is an experiment with a unique detector dedicated to its own physics program. These experiments gained approval between 1996 and 1998, and construction commenced at the chosen sites for the detectors. At the transition points in octants 1 and 5 are the "A Toroidal LHC ApparatuS" (ATLAS) [108] and "Compact Muon Solenoid" (CMS) [109] experiments, both are general purpose barrel-

shaped detectors, designed with the primary goal of discovering the Higgs Boson. At the interaction point in octant 2 is the "A Lead Ion Collision Experiment" (ALICE) [110] experiment, designed to study the products of collisions between heavy nuclei. Finally at point 8 is the "Large Hadron Collider beauty" (LHCb) experiment, the focus of the following chapter.

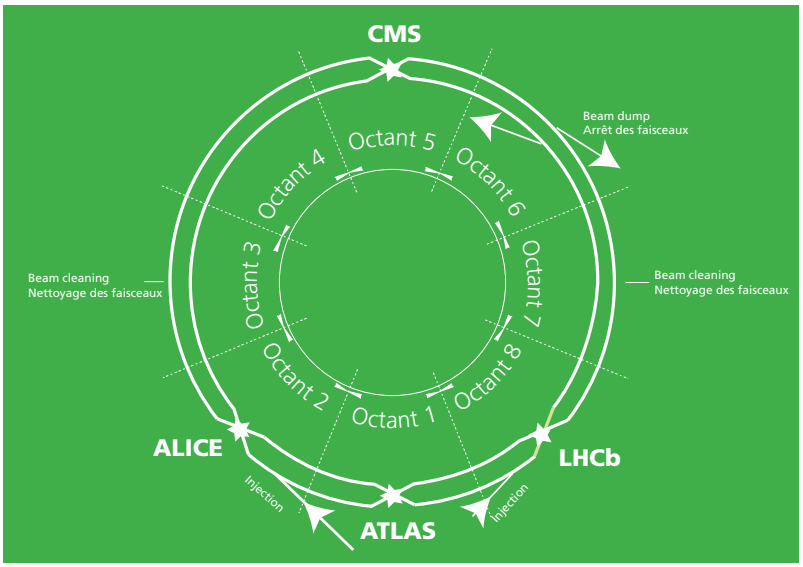


Figure 2.1.2: LHC schematic diagram taken from [104], an adapted version of Fig 1 from the LHC conceptual design report from 1995 [111].

Other experiments are located on the LHC, these include TOTEM [112], LHCf [113], and MoEDAL-MAPP [114]. Located next to ATLAS and CMS respectively, the LHCf and TOTEM experiments focus on the study of forward particles, which are produced at small angles relative to the beamline. Rather than colliding head-on, some particles in the beams undergo peripheral interactions, skimming each other which leads to the production of the forward particles further along the LHC beam line. These particles carry significant fractions of the beam momentum and are important probes for studying phenomena such as diffractive scattering [115]. Both experiments are positioned on either side of their respective collision points.

Near to the LHCb collision point is the MoEDAL-MAPP experiment, searching for a particle called the magnetic monopole and other exotic particles. FASER [116] and SND@LHC [117] are located near to ATLAS and are the newest LHC experiments, searching for light new particles and neutrinos.

2.2 Introduction to LHCb Experiment

The LHCb experiment is specially designed to probe the differences between matter and antimatter via the study of the b and c quarks in proton-proton collisions. The LHCb experiment was built in the cavern excavated for the DELPHI experiment located on the LEP accelerator, constraining the design of the detector to the forward region. Unlike ATLAS and CMS, general purpose detectors which enclose the interaction point, LHCb is a single-arm forward spectrometer, detecting the particles which travel forward from the collision in a single direction covering angles from approximately 15 mrad to 300 (250) mrad in the experiments bending (non-bending) plane. The angular acceptance of LHCb corresponds to the pseudo-rapidity range of $2 \leq \eta \leq 5$. [1] This configuration leverages the fact b-quarks are predominantly produced within a forward cone at the LHC energies.

The LHCb detector contains multiple sub-detectors, detailed throughout the rest of the chapter. Starting with the VELO, which is close to the collision point, it extends for a length of 21 m. A schematic view of the detector configuration for Run I and Run II is shown in Figure 2.2.1. In this diagram, the coordinate system is defined such that the z-axis follows the beamline, with proton-proton collisions occurring at z=0. The y-axis represents the vertical height of detector, while the x-axis points towards the centre of the LHC ring. The region of the detector at positive (negative) z values is referred to as the forward (backward) or downstream (upstream) end.

The detector subsystems are mostly assembled in two halves, located left and right when looking at the detector from the interaction point. These two halves can be slotted in and out of the detector for maintenance and provide access to electronics and the beam-pipe.

The acceptance of the LHCb detector is motivated by the production of the heavy b and c quarks predominantly occurring in the forward or backward cone. The detector acceptance covers 1.8% of the solid angle but covers 27% of the $b\bar{b}$ production [118].

2.3 Tracking and Positioning

The sub-detectors within the LHCb experiment serve one of three primary purposes: tracking the particles, identifying them, or measuring their energy through calorimetry. The set of detectors that track the particles as they traverse the detector are (moving downstream from the interaction point), the Vertex Locator (VELO), Tracker Turicensis (TT), the dipole magnet, and three tracking stations (T1-T3). Each is discussed in the following sections.

[1] Here, pseudorapidity is defined as $\eta = -\ln(\tan(\theta/2))$, where θ denotes the angle between the particle and the beamline.

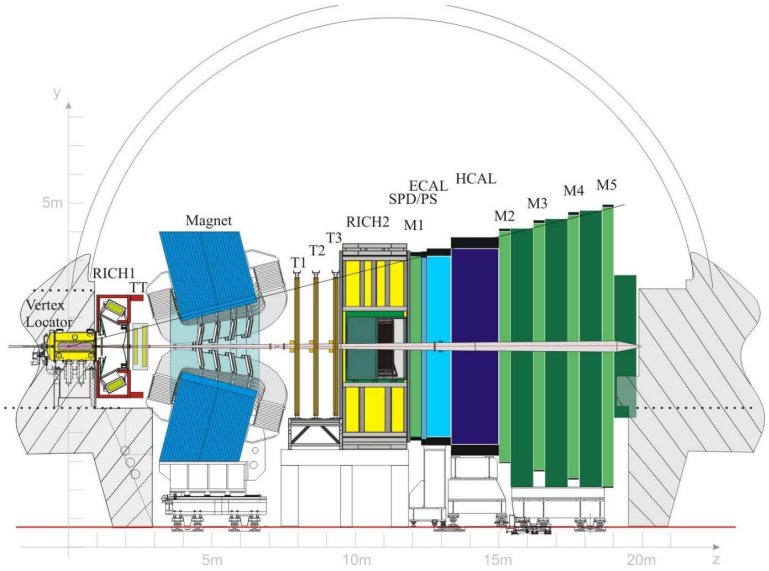


Figure 2.2.1: Sketch of the side-view of the LHCb detector during Run I and Run II [119]. The collisions take place at the interaction point, inside the vertex locator, to the left of the diagram, and proceed right, through each of the sub-detectors.

As each particle is registered by a detector layer, it is recorded as a hit. These hits are used to reconstruct the path travelled by each particle. By reconstructing all of the trajectories from a collision, an event can be partially reconstructed, including the particle tracks and vertices, which represent the points where interactions or decays occur. The non-hermitic acceptance of the LHCb detector means that some particles may not be fully reconstructed, as many escape detection. These tracks are classified based on where the particle hits are recorded, as shown in Figure 2.3.1.

Each track type is useful for identifying different aspects of a particle and the event to which is belongs. If a particle is reconstructed only in the VELO, it is classified as a VELO track. VELO tracks are used to identify the primary vertex (PV) of an event. If particle hits are only recorded in the VELO and TT, the track is classified as an upstream track; typically, these particles leave the LHCb acceptance after the TT. Downstream tracks pass through the TT and the T1-T3 stations, while T tracks only pass through T1-T3. Finally, long tracks pass through the VELO and all tracking stations, therefore they measure momenta the most precisely.

Long tracks are particularly useful for physics analyses, including that which is presented in this thesis. The B^0 and $\overline{B}^0$ mesons are short-lived, so their decay vertices are close to the interaction point, requiring hits in the VELO. The final state particles, especially the muons, traverse the entire detector. Long tracks provide comprehensive information from all the tracking stations, which is essential for the accurate reconstruction of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode.

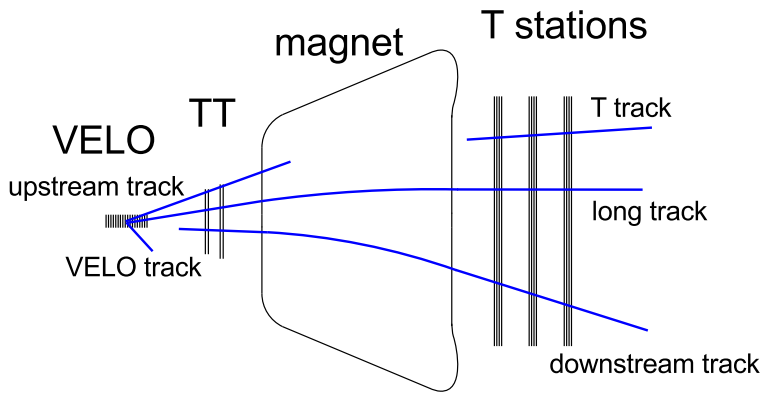


Figure 2.3.1: Track types in the LHCb detector [120], the details of each track type are found in the text.

2.3.1 Vertex Locator

The VELO [121] is a sub-detector which encloses the interaction point at the LHCb detector. During Run I and Run II the sub-detector consisted of two halves, each containing 21 modules of silicon strip sensors placed in an r, ϕ polar geometry. Measurements of track coordinates are performed allowing the identification of primary interaction vertices, and displaced secondary vertices (SV).

The VELO sensors are placed only 7 mm from the LHC beamline [122]. Performing measurements so close to the interaction point is pivotal for performing physics analysis of several observables key to the LHCb physics mandate. For instance, time-dependent measurements such as the rapidly oscillating $B_s^0 - \overline{B}_s^0$ meson system which requires a vertex resolution of ~ 0.1 mm [123, 124]. The VELO is able to provide spatial resolution of $\sigma_{xy} \sim 15\,\mu$m, $\sigma_z \sim 80\,\mu$m, where σ_{xy} refers to the resolution in the transverse plane, perpendicular to the beam axis, and σ_z to the resolution in the longitudinal plane, along the beam axis. The vertex resolution is optimised by placing the sensors of the VELO as close to the beam as possible. An important parameter measured by the VELO is the Impact Parameter (IP), defined as the closest distance between the PV and the extrapolated path of a track. The IP helps in determining whether a particle is prompt, meaning it was produced at the PV, or if it originates from a displaced secondary vertex, for example from a B meson decay.

The VELO is constructed from a series of silicon micro-strip modules arranged along the beam direction (z-axis). Since the sensors must be positioned close to the LHC beam for optimal tracking but would be at risk of damage during injection due to beam instabilities, the detector is built in two halves that come together to close when the LHC achieves stable beam circulation, as shown in Figure 2.3.2, centring on the interaction

region with a small overlap of 1.5mm. The overlap provides full angular coverage, and the proximity to the beam ensures excellent IP resolution.

Each module comprises two semi-circular sensors: the R-sensors measure the radial distances of particle hits from the beam position (the r coordinate), while the ϕ-sensors measure the azimuthal position (the ϕ coordinate). The z position is determined from the position of the module which registers the hits. A schematic view of the modules is shown in Figure 2.3.3. The VELO must provide angular coverage for all tracks in the forward region ($15 - 100$ mrad), ensuring that these tracks cross a minimum of three VELO stations. Two pile-up modules, located in the upstream regions of the VELO, also contain R sensors.

The VELO modules are contained inside of a vacuum separated from the LHC vacuum by an RF box, with inside surfaces (facing the beam) produced from 0.3 mm thick aluminium foil. This RF foil shields the VELO modules from RF pickup originating from the beam and prevents interference with the LHC vacuum.

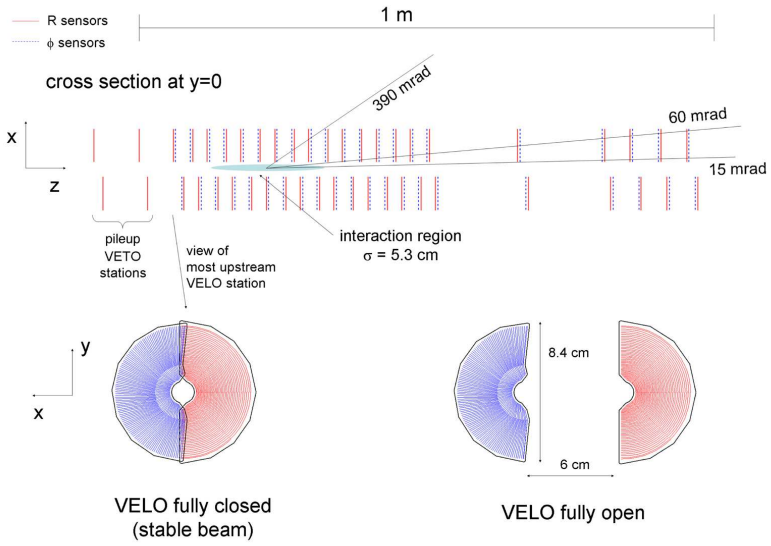


Figure 2.3.2: Cross section in the (x,z) plane of the VELO silicon sensors, at y = 0, with the detector in the fully closed position. The front face of the first modules is also illustrated in both the closed and open positions. The two pile-up veto stations are located upstream of the VELO sensors. Taken from Ref. [100].

2.3.2 Tracker Turicensis

Downstream of the VELO is the first of the tracking stations, the Tracker Turicensis (TT). The TT is a silicon strip detector, which measures the positions of particles before they pass through the dipole magnet, which aids in track reconstruction. This is particularly important for tracking particles that do not reach the downstream trackers.

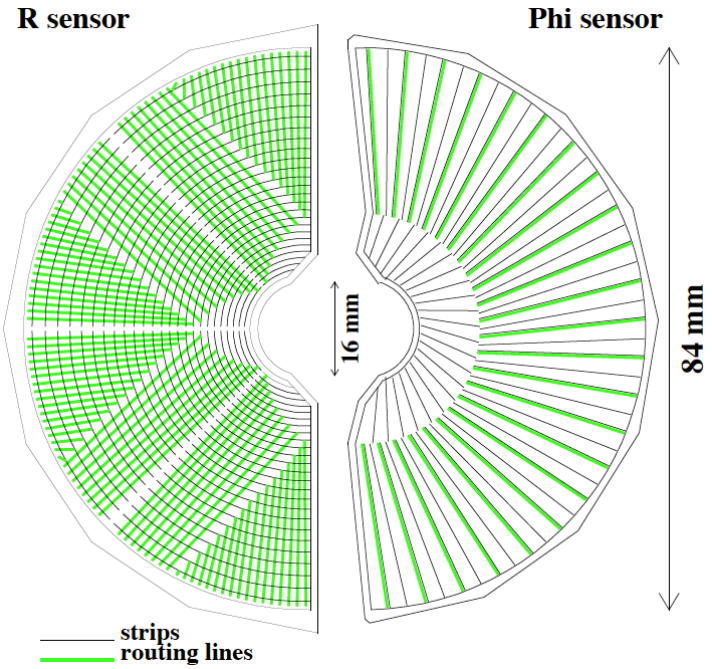


Figure 2.3.3: Schematic representation of an R and a Φ sensor in the VELO. The R sensor strips are arranged into four approximately 45° segments and have routing lines perpendicular to the strips. The Φ sensor has two zones with inner and outer strips. The routing lines of the inner strips are orientated parallel to the outer strips [122].

The TT is made up of four tightly packed layers that are thermally and electrically insulated, maintained at a temperature below 5°C to suppress radiation damage [125]. The detector is flushed with dry nitrogen gas (N_2) to avoid condensation on the cold surfaces. Each of the layers is assembled from two half-modules which join to cover the full height of the LHCb acceptance [119]. The four layers are arranged in two pairs (x,u) and (v,x), as shown in Figure 2.3.4, located 27 cm apart along the z-axis, with the first centred around z = 232 cm. The x-layers strips are vertical, while the inner two layers are oriented ±5°, respectively, with respect to the vertical axis.

Following the end of LHC Run II and 10 years of data taking, a report determined that the amount of radiation damage to the TT did not significantly impact its performance [126].

2.3.3 Magnet

Located between the TT and T1-T3 tracking stations is the LHCb dipole magnet. The purpose of the magnet is to allow the measurement of the momenta of charged particles by bending their trajectories in a known magnetic field, with the curvature allowing the determination of their momenta. The design of the LHCb detector requires a dipole

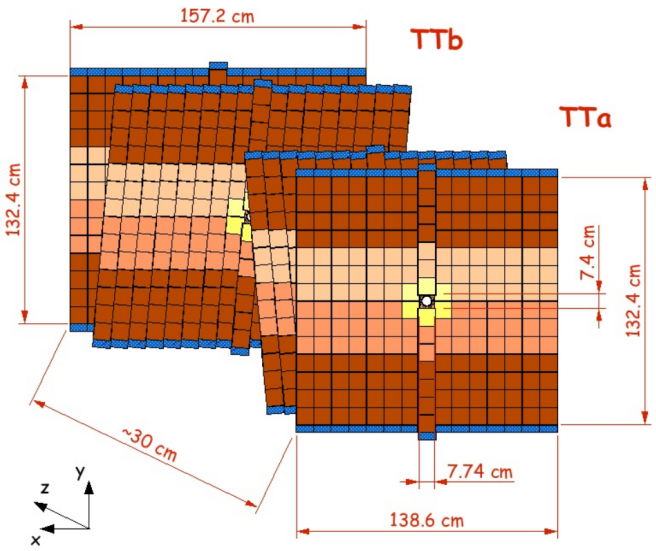


Figure 2.3.4: Layout of the LHCb TT silicon tracker. The LHC beam pipe passes through an opening in the centre of the detection layers. The four detection layers are arranged in two pairs: (x,u) and (v,x) [126].

magnet able to cover the whole of the LHCb acceptance. To meet these aims a warm magnet, shown in Figure 2.3.5, is constructed from two saddle-shaped coils of layers of aluminium, in a window-frame iron yoke, symmetric around the axis of the beam [127].

The magnet produces an integrated magnetic field of 4 Tm. Alongside the tracking detectors located either side of the the field, momentum measurements are performed for charged particles. A resolution of approximately $\frac{\delta p}{p} = 5 \cdot 10^{-3}$ is achieved for particles below 20 GeV/c. This resolution increases to about $8 \cdot 10^{-3}$ for particles around 100 GeV/c, with a systematic uncertainty of 0.03% on the momentum scale [101].

The magnetic field must be known precisely to accurately determine the momentum of a particle. The magnetic field is measured using Hall probes and mapped to determine the field strength at different locations within the detector. As shown in Figure 2.3.5, the field strength drops quickly when moving away from the magnet. This rapid decrease minimizes the field strength inside the VELO, ensuring that tracks remain straight and close to the interaction point, thereby preserving the VELO's resolution. The two Ring Imaging Cherenkov detector (RICH) subsystems, used for particle identification (PID), are located either side of the magnet. Although the field around the RICH detectors is not negligible, it is controlled to be below 2 mT, as the strong field, of approximately 60 mT [128], could affect the photon detectors by altering the trajectories of the photoelectrons. To achieve this, and maintain a high field strength between the VELO and the tracking stations, the RICH photodetectors are encased in iron shielding.

The magnet deflects oppositely charged particles in opposite direction in the x-z plane.

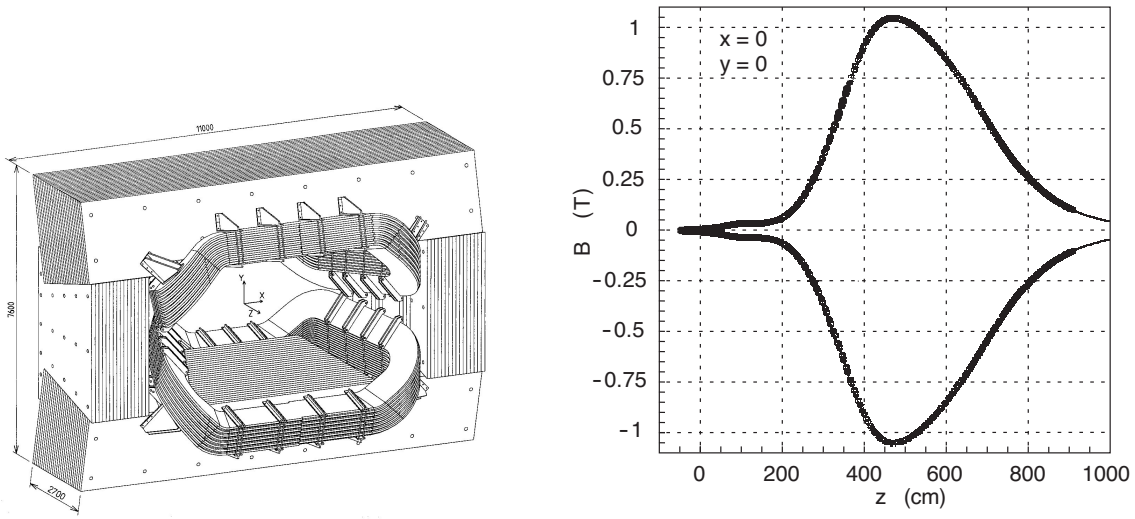


Figure 2.3.5: The figure on the left shows a perspective view of the LHCb dipole magnet (units in mm). The interaction point lies behind the magnet in this figure. Taken from the technical design report [127]. The figure on the right shows the magnetic field strength mapping of the LHCb dipole magnet as a function of the z-axis, measured by Hall probes at $x = 0$ and $y = 0$. The two curves show the field strength for both polarities of the magnet ("Up" and "Down"). Taken from Ref. [100].

To account for any potential asymmetries in the detector's response or in the uniformity of the magnetic field, the magnet's polarity is regularly switched between "Up" and "Down". This process helps produce datasets where the detector response to oppositely charged particles is symmetrical, effectively cancelling out detection asymmetries. Having a warm dipole allows for rapid ramping-up of the magnetic field synchronous with the LHC magnets, as well as facilitating regular field inversions.

2.3.4 Downstream Tracking Stations

The tracking stations, T1-T3, located downstream of the dipole magnet, each give independent measurements in order to achieve high momentum resolution and track reconstruction efficiency. They are split into two regions due to the high levels of radiation in the region close to the beam-pipe; the Inner Tracker, IT, and the Outer Tracker, OT. A diagram of the two tracking stations are shown alongside the TT in Figure 2.3.6.

Inner Tracker

The IT detector modules are made out of silicon microstrips, similarly to the TT. Due to its proximity to the beamline, the IT is exposed to a high density of charged particles, accounting for approximately 20% of the total of charged particles.

Each IT station consists of four detector "boxes" arranged around the beam pipe.

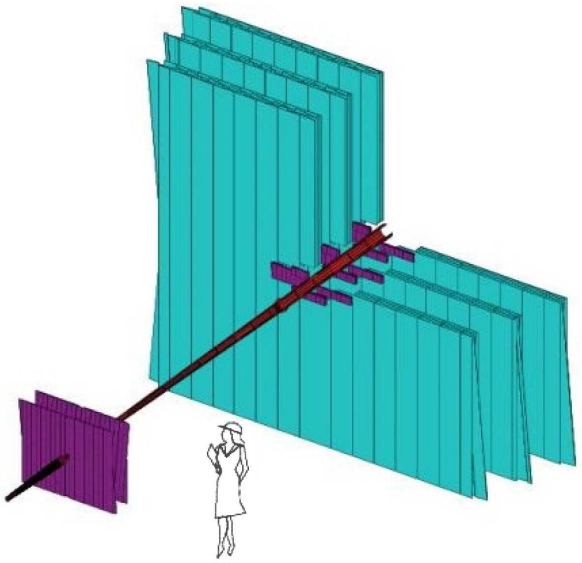


Figure 2.3.6: Sectional view of the LHCb tracking stations. The TT is shown in purple, upstream of the IT and OT. The IT is in purple, close to the beamline, surrounded by the OT (cyan) [119].

Similarly to the TT the IT is composed of four layers of silicon strips in two pairs (x,u) and (v,x), with the x strips vertical and the u, v strips at $\pm 5°$ from vertical. The layout is shown in Fig. 2.3.7.

The boxes located above and below the beam-pipe consist of one layer of 7 silicon sensors, while the boxes either side (A and C) contain two layers. The readout electronics are contained within the detector acceptance.

The IT is consistently cooled to temperatures below 5°C and flushed with C_6F_{14} (Perfluorohexane) to prevent condensation.

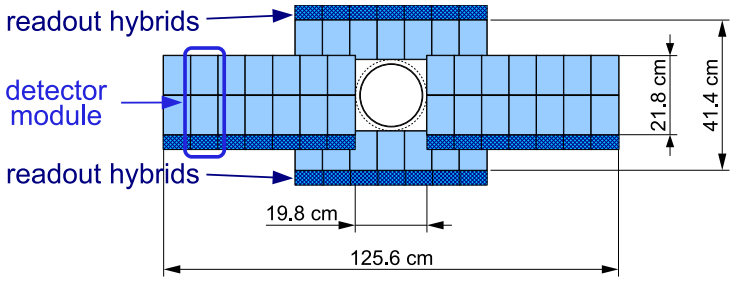


Figure 2.3.7: Layout of the first layer of silicon sensors in the IT. Figure taken from Ref. [100].

Outer Tracker

The OT modules are made from gas-filled drift tubes as opposed to the silicon microstrips of the other tracking modules. The stations are also arranged in a (x,u) and (v,x) geometry. The width of the OT is larger than its height to account for the path bending of the

charged particles which have passed through the dipole magnet.

The OT modules of the IT array are constructed of two layers of tubes of diameter 4.9 mm. The gas within the tubes is a mixture of 70 % argon and 30 % carbon dioxide, which is ionised by charged particles travelling through the straws. The ionised electrons drift in the electric field in the tube and induce an electrical current which is read by a carbon cathode. The OT tubes measure the drift-time of the electrons from the aluminium coating on the outside of the tubes to the gold anode wire at the centre. The drift time is combined with the drift velocity to measure the track position of the charged particles, giving a 200 µm hit-position resolution in the x direction [129]. This was improved in Run II to 170 µm [130], predominantly because of reconstruction software improvements (see offline processing discussed in Section 2.5).

2.4 Particle Identification

2.4.1 RICH

The Ring Imaging Cherenkov detectors (RICH) [131] subsystem comprises two detectors designed for particle identification (PID) of hadrons associated with reconstructed tracks. The first subdetector located upstream of the VELO is the RICH1, while the RICH2 subdetector is located downstream the dipole magnet and tracking stations. These detectors identify charged hadrons over wide momentum range of $1 - 100\,\mathrm{GeV}/c$ and an angular range of $25 - 300\,\mathrm{mrad}$.

The two RICH detectors have different dielectric radiator materials: RICH1 uses C_4F_{10} gas, while RICH2 uses CF_4, each have slightly different refractive indices. This difference allows the detectors to cover the full momentum range of LHCb. RICH1 [132] covers the lower momentum range over a large polar angular range, while RICH2 [133] covers a higher momentum range over a lower angular range. Specifically, RICH1 detects particles with a momentum range of approximately 2 to $60\,\mathrm{GeV}/c$, while RICH2 covers a momentum range of approximately 15 to $100\,\mathrm{GeV}/c$.

The RICH detector is designed to identify charged particles using Cherenkov emission. The photons are emitted in a cone, which is focused into a ring image via a combination of spherical and flat mirrors onto and array of hybrid photodetectors (HPDs). The detectors accept photons in a wavelength range of 200 nm $< \lambda <$ 600 nm. To protect the HPDs from strong magnetic fields that could affect the trajectories of the photoelectrons, they are enclosed in iron shielding and positioned outside the acceptance of the LHCb detector. The dielectric material emits photons when the velocity of the particles exceed the phase velocity of light in the medium, resulting in light being emitted in a cone with an angle of θ_c given by:

$$\cos\theta_c = \frac{1}{n\beta}, \tag{2.3}$$

where β is the ratio of the particle's velocity v to the speed of light in a vacuum, c, and n is the refractive index of the dielectric material. The observed images, focused by the mirrors, are compared to expected images for different hypotheses using a pattern recognition algorithm [134]. This algorithm reconstructs both the direction and velocity of the particles which produced the image. This information is then combined with the momentum measured by the tracking detectors to determine the mass of the charged hadron via the relation:

$$\cos\theta_c \simeq \frac{1}{n}\sqrt{1+\left(\frac{mc}{p}\right)^2}. \tag{2.4}$$

The Run I efficiency and reliability are considered high and effective in a performance review of the LHCb detectors in 2013 [131]. The performance of the RICH subdetector system is evaluated based on its efficiency in correctly identifying charged particles, such as pions, kaons, and protons. The log-likelihood ratio (ΔLL) compares the log-likelihood of two PID hypotheses for a given track. Three examples, examining performance during Run II, are given in Figure 2.4.1, with the top left figure, for instance, showing the separation of kaons and pions using $\Delta LL(K,\pi)$. The figures demonstrate the effectiveness of the RICH detectors in distinguishing between particles over a range of track momenta. Two selection are shown: a loose selection that yields high signal efficiency and a tight selection that provides good background rejection.

2.4.2 Calorimetry

The LHCb calorimeters [136, 137] are located downstream of RICH2 and the magnet. Their primary purpose is to select and identify hadrons, electrons, and photons, as well as to measure their transverse energy (E_{T}), a measure of the particle's energy derived from the transverse momentum (p_{T}) and the invariant mass of the particle, given by:

$$E_{\mathrm{T}} = \sqrt{p_{\mathrm{T}}^2 c^2 + m^2 c^4}. \tag{2.5}$$

However, in practice, E_{T} is often determined from the total energy E measured by the calorimeter and the polar angle θ via $E_{\mathrm{T}} = E\sin\theta$.

The detectors are comprised of alternating layers of scintillators and absorbers. As particles pass through, they interact with the absorber material, initiating electromagnetic showers primarily via bremsstrahlung, pair production, and Compton scattering at higher energies. These interactions produce a cascade of secondary particles, some of which then

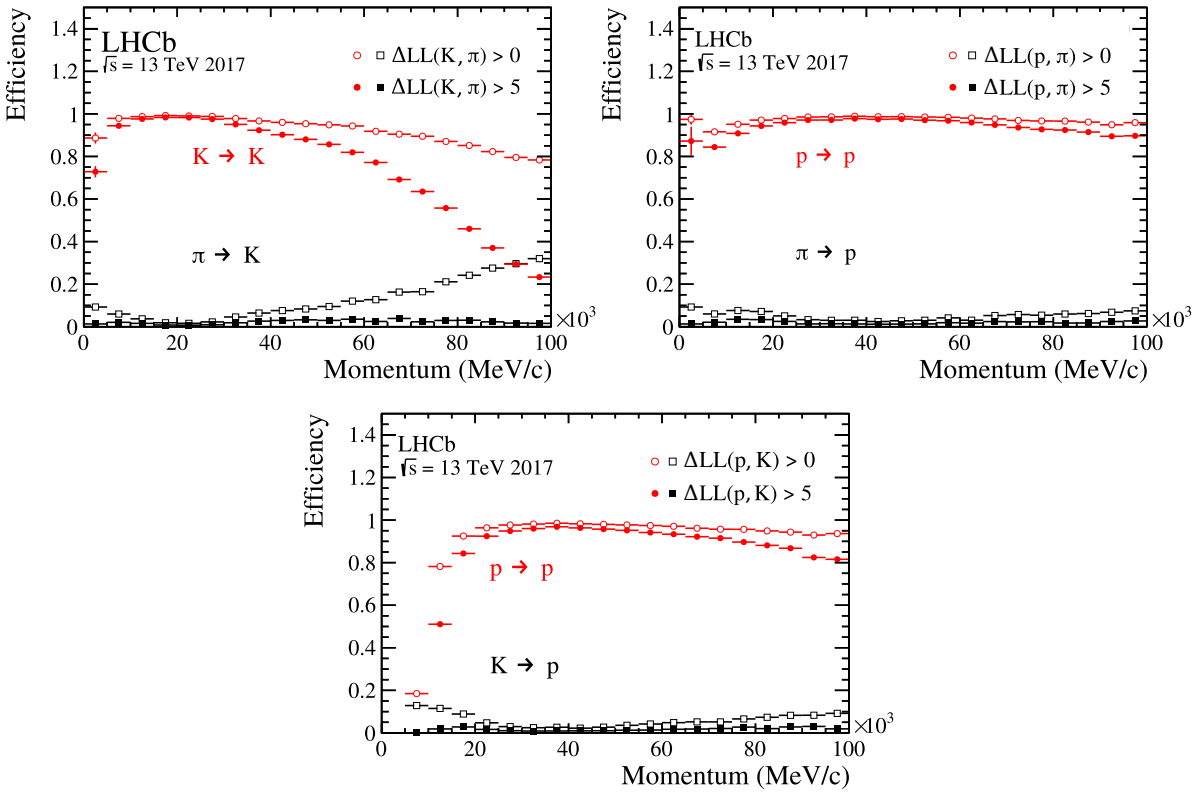


Figure 2.4.1: The efficiency of kaon selection (top left) and proton selection (top right and bottom), along with the associated misidentification rates of pions (top row) and kaons (bottom row), as a function of track momentum. Two selection criteria based on the log-likelihood are applied: a loose selection (hollow shapes) and a tight selection (solid shapes). Adapted from Ref. [135]

pass into the scintillating layers, where their energy is converted into light signals. The E_{T} is calculated by summing the energy of the particles in the shower.

As particles pass through the scintillating material, atoms or molecules become excited, and visible light is emitted during de-excitation. This scintillation light allows for the measurement of particle energy [138]. The light is then collected by a photomultiplier tube and converted into an electrical signal.

The detector array consists of the Scintillator Pad Detector (SPD) and Pre-Shower (PS) detectors, the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL).

The photons pass through the SPD, leaving no signals and interacting only with the lead plates. This interaction yields EM showers which produce signals in the PS. Conversely to photons, electrons and pions do leave signals in the SPD. The thickness of the lead plates is chosen to produce electron showers in the PS, while avoiding showers from the pions,

discriminating between the two particles. The most demanding task the calorimeters face is in identifying electrons, whose signal is affected by a large background from photons, many of which originate from neutral pion decays. There is a large background dominated by photons which are converted upstream in the spectrometers. The constraints needed in the calorimetry PID helped to define the characteristics of the calorimetry system outlined in Refs. [136, 139].

A 2.5-radiation-length lead foil is placed between the SPD and the PS, which initiates electromagnetic showers from the particles traversing the foil. Signals in the SPD are produced by charged particles, differentiating them from neutral particles. The PS is used to identify electrons via the detection of electromagnetic showers they produce. As particles pass through the detector, they interact with the lead and metal plates which absorb the energy of the particles, producing showers of secondary particles. These particles are detected by the aforementioned scintillating pads downstream, providing information which is used to distinguish electrons from charged pions. The showers produced by the charged pions typically deposit their energy earlier in the shower than the electrons.

The electromagnetic calorimeter (ECAL) is designed to contain and measure the complete showers of photons and electrons. Its design emphasises the need for a fast response, good radiation resistance, reliability, and acceptable energy resolution. The ECAL has a design energy resolution of $\sigma_E/E = 10\%/\sqrt{E[\,\mathrm{GeV}]}$, which allows for precise energy measurements. It is located 12.5 m from the interaction point, with the inner dimensions at $\rho_{xy} > 25\,\mathrm{mrad}$ and outer dimensions at $\rho_x < 300\,\mathrm{mrad}$ and $\rho_y < 250\,\mathrm{mrad}$, where ρ is the radial distance from the beamline perpendicular to the beam axis. The hit density drops steeply with distance from the beam axis, so the ECAL is divided into three sections. Each detector module is constructed from alternating layers of 2 mm thick lead sheets, followed by 120 µm of paper and 4 mm of scintillating polystyrene tiles. Each stack contains 66 layers, resulting in a combined thickness of 42 cm. This thickness corresponds to 25 radiation lengths for photons and electrons, ensuring that the showers are completely contained within the ECAL.

The hadronic shower penetrates through the ECAL into the hadronic calorimeter (HCAL). The HCAL is composed of alternating layers of 1 cm thick absorbing iron sheets and active scintillating material. It is oriented perpendicular to the beam-line axis, with a longitudinal depth equal to one hadron interaction length in steel. Each module contains 8 sub-modules, resulting in a total of 416 sub-modules across 52 modules in the detector. The maximum depth of the HCAL is constrained by the available space within the LHCb detector, limited to 1.6 m, which corresponds to an interaction length of 5.6 in steel. The HCAL has much larger readout cells than the ECAL, and along with its limited depth, this results in a design energy resolution of $\sigma_E/E = 69\%/\sqrt{E[\,\mathrm{GeV}]}$, which affects its

ability to measure energy accurately. As a result, showers from high-energy particles are not completely contained within the HCAL. However, the HCAL is still capable of efficiently identifying high E_T particles despite this limitation.

2.4.3 Muon System

The LHCb muon detectors [140, 141] are designed with the purpose of identifying muons that are able to pass through the calorimeters without being fully absorbed. While some muons are stopped in the calorimeters, many high-energy muons are able to penetrate further and reach the muon detectors. The five detectors are split into two section found either side of the calorimeter sub-detectors, with the M1 detector upstream and M2-M5 located downstream.

Muons are the only charged particle that traverse the entirety of the calorimeters. Identifying muons is an important aspect of flavour tagging and for various analyses, including the one presented in this thesis. The information produced by the muon modules provides information for the high-p_{T} hardware trigger (Section 2.5.1) and in muon PID in the software trigger (Section 2.5.2).

To ensure enough hits are recorded to precisely reconstruct the muon trajectories, iron absorbing sheets, each 80 cm thick, are placed between the M2 and M5 detectors. These sheets are designed to only allow muons with $p > 6\,\mathrm{GeV}$ to traverse the muon system, stopping any lower-momentum particles. As shown in Figure 2.4.2 (a), the inner and outer acceptances of the detectors progressively increase from 16 mrad to 258 mrad in the non-bending, vertical plane and 20 mrad to 306 mrad in the bending, horizontal plane. This projective geometry means that the dimensions of the detectors scale with distance from the p-p interaction point. The M1 detector is placed prior to the SPD/PS detectors to improve the p_{T} measurements for use in the trigger, before the muon passes through the dense calorimeter detectors. Consequently, M1, along with M2 and M3, is important for accurate determination of the muon track direction. The remaining two stations, M4 and M5, are primarily used to identify the most penetrating particles, distinguishing them from less energetic ones.

The inner M1 uses Gas Electron Multipliers (GEM), these $GEMs$ consist of three layers of foils between cathode and anode plates. The ionisation electrons produced in the drift-gap gas are attracted by electric fields through the three foils and into the anode in the induction gap, producing an induced current on the signal pads. The $GEMs$ are filled with a mixture of argon, carbon dioxide, and carbon tetrafluoride, at a ratio of $(45 : 15 : 40)$, resulting in a time resolution of 3 ns.

For each of the remaining stations Multi-Wire Proportional Chambers ($MWPC$) are used. The $MWPCs$ are constructed from two cathode plates with a separation of 5 mm,

with a similar ratio of gases to the $GEMs$, at $40 : 55 : 5$. Smaller avalanches are required for the $MWPCs$ and therefore a larger amount of inert gas is used. The cathodes are connected to a gold-wire which reads out with a time resolution of 5 ns.

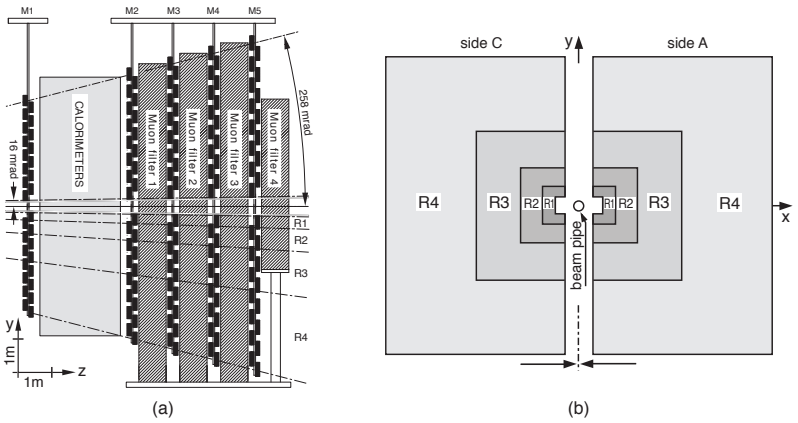


Figure 2.4.2: Side view of the LHCb Muon system. The diagram outlines (a) the side view of the LHCb Muon Detector, and (b) the station layout with the four regions R1–R4. [140]

2.5 Trigger and Data Processing

A trigger system is crucial in reducing the large amount of data produced during collisions. It filters out uninteresting background events and retains those that are relevant to physics analyses. The implementation of such a system optimises both storage and computing resources.

Events are selected once they meet a set of kinematic and topological selection requirements. Each set of selection requirements is unique to its specific use case and is chosen to optimise the conditions for selecting particular physics processes.

Due to large volumes of data collected during operation and limited storage space and CPU time, significant data reduction is required before more detailed software-based triggering can be applied.

For an event to pass the trigger selection, it must succeed in each of three stages [142, 143], which are discussed in the following sections. The three stages progressively refine event selection, starting with simple hardware-based criteria and progressing to more stringent software-based ones. This multi-stage implementation balances efficiency with computational loads. An event passes each stage by meeting one of the corresponding sets of trigger selection requirements. An illustration of the trigger system is given in figure 2.5.1.

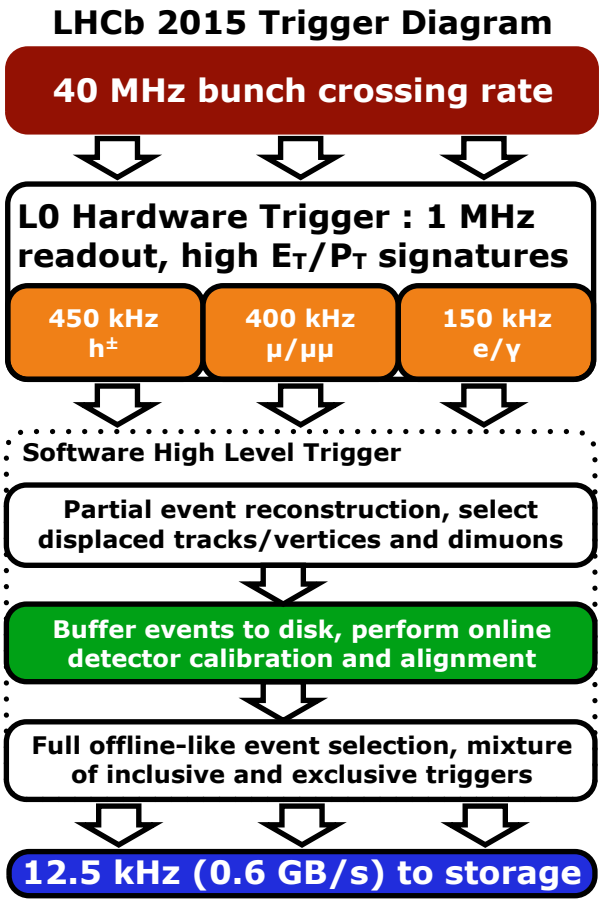


Figure 2.5.1: Diagram of the LHCb Run II trigger system. Taken from Ref. [144].

2.5.1 Hardware Trigger

The Level-0 (or L0) trigger is the initial trigger system, featuring electronics boards that process information read out by the calorimeters and muon system to quickly select events within 4 μs. The processed information is then passed to the L0 decision unit (DU). The L0 trigger reduces 40 MHz of data to 1 MHz.

In Run I and Run II, the L0 trigger comprises two independent systems: the Calorimeter and Muon triggers. The Calorimeter trigger identifies hadronic candidates with high E_T. The E_T information is combined into clusters of 2×2 cells and selects the cluster with the highest E_T for the DU. Photon candidates are identified based on hits in the PS, while subsequent hits in the SPD indicate the candidate is an electron.

The Muon trigger identifies candidate tracks with the two highest transverse momenta (p_{T}) recorded in the muon stations. Particles with p_{T} exceeding a specified threshold are selected by the hardware trigger. The successful candidates from both triggers are sent to the DU, which consolidates the data and performs basic logical operations to combine the signatures.

2.5.2 High Level Trigger 1

Those events which pass the hardware trigger are given to the High Level Trigger (HLT), which is split into two section, HLT1 and HLT2. The HLT is run on the Event Filter Farm (EFF), consisting of 1700 nodes running C++ software. However, due to CPU limitations and the output rate of L0 being 1 MHz, the HLT trims down the data while processing partial events and rejecting a large portion of the uninteresting events. A benefit of the HLT trigger being software based is that it is able to evolve over time and develop to follow the evolution of the LHCb physics programme.

The HLT1 trigger filters the information which passes the L0 trigger using fast algorithms, confirming the L0 decisions on particle tracking information such as candidate tracks, calorimeter clusters, and muon information, as well as neutral hadrons and photons if no track is present. The HLT1 performs basic reconstruction using this information, but does not perform a full fitting of the tracks.

Loose selection cuts are applied on kinematic variables of the tracks, such as the p_{T}, IP, and the distance of closest approach of the detected particles in relation to the beam axis. The tracks are used to form composite particles such as $K^{\star} \to K^{+}\pi^{-}$ or $\phi \to K^{+}K^{-}$.

The HLT1 takes the 1 MHz of data output from the L0 trigger and reduces this to 30 kHz and 110 kHz of data during Run I and Run II respectively. The remaining data is passed to the HLT2 trigger.

2.5.3 High Level Trigger 2

The HLT2 trigger specialises in specific decay selection using all of the information from the LHCb detectors for the remaining candidates. The HLT2 uses, as in HLT1, C++ programming language to apply a range of algorithms which select specific physics processes. The algorithm explicitly selects tracks likely to be the decay products (daughters) of a decaying particle (mother). The inclusive selection involves a partial reconstruction of a mother particle.

Once the data is selected for offline storage and processing, it is reduced to 5 (12.5) kHz during Run I (Run II). Events are filtered based on selection triggers that identify candidates capable of triggering the event. The HLT2 trigger is crucial for its adaptability to new selection criteria, particularly for handling complex event classification into specific decay modes.

2.5.4 Offline Processing

The raw data that passes the HLT2 trigger requirements is moved to offline storage for further analysis and processed through software packages [145], starting with Brunel.

During Run I, Brunel reconstructed data from the LHCb in a similar manner to HLT2. However, since Brunel operates offline, it is subject to less stringent time constraints, allowing for more rigorous processing. For Run II, this process was modified [146] so that the HLT2 triggers and the offline reconstruction operate with the same software. The HLT1 ran continuously over periods of up to two weeks, processing data in real time before passing it to HLT2. This allows time for calibration alignment and efficient utilisation of HLT resources during periods without collisions. This change resulted in a fourfold increase in data volume, even though the integrated luminosity only doubled.

Brunel uses the information from the tracking and PID detectors to produce track reconstruction and PID for each of the recorded events. The data output from Brunel is passed to the DaVinci package which performs candidate reconstruction and selection on the event data and 'strips' the events into the identified decay modes via 'stripping selection', implemented by the user to construct datasets of the desired decay mode. The DaVinci output datasets can then be used by the user to analyse the desired physical phenomena.

2.6 LHCb Simulation

Integral to physics analyses performed by the LHCb collaboration are simulated events designed to closely replicate real data. In order to simulate accurately the known physical processes involved in pp-collisions the Monte Carlo, MC, data is processed in the same way, through the LHCb reconstruction and trigger software.

The first step in the workflow involves Gauss [147], within which pp collisions are generated using the MC event generator Pythia [148], which models the underlying hard-scattering process, parton showering, and hadronisation. The LHCb configuration of Pythia [149] is tuned to match experimental data as closely as possible. Unstable particles produced in these collisions are passed to EvtGen [150], which propagates their decays. EvtGen describes hadronic and semileptonic decays using experimental branching fractions and decay dynamics. Within EvtGen, Photos [151] is used to model final-state radiation, ensuring a realistic treatment of QED effects. Once the event kinematics and decays are simulated, the particles are passed to Geant4 [152, 153], which simulates the interactions of particles with the detector materials. This step replicates the experimental conditions under which real data is collected.

The simulated events are then processed through several LHCb software packages [145] which emulate the reconstruction through the LHCb detector. First they are passed to Boole, which mimics the response of the front-end electronics in the actual detector, producing the same digital output as the data acquisition system. These signals are treated

in the same way as the real LHCb data and are processed through the same software. The L0 hardware trigger response is emulated using MOORE, which also provides the HLT software trigger emulation. MOORE is followed by offline processing with BRUNEL, which reconstructs complex objects, and DAVINCI, which is used for applying selection requirements and reconstruction for specific physics analyses.

2.7 The LHCb Upgrade and Future Operations

The LHC has now completed Long Shutdown 2 (LS2), during which major upgrades were made to the LHCb detector [154]. These upgrades have prepared the detector for the aim of taking $50\,\text{fb}^{-1}$ of data in Run III, a factor five increase in instantaneous luminosity over the combined Run I and Run II. The trigger system has been overhauled with the removal of the L0 hardware trigger moving the LHCb to a fully software HLT. Trigger selections have been rewritten and some sub-detectors have been upgraded. Additionally, the front-end electronics and readout electronics of all of the sub-detectors have been replaced, meaning the LHCb detector can read out information at a bunch crossing rate of 40 MHz. The LHCb has proved an extremely capable experiment for the analysis of rare decays and *CP* violation processes, with these improvements, the LHCb experiment will continue to play a crucial role in these studies in Run III and beyond.

Chapter 3

Data preparation aspects for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ analysis

The analysis presented in this thesis uses pp collision data collected by the LHCb experiment during the first run of data taking years, 2011–2012, and the second run, 2015–2018. The 2011–2012 dataset, referred to hereafter as Run I data, corresponds to an integrated luminosity of 3 fb^{-1}, while the 2015–2018, Run II, dataset contains an additional 5.7 fb^{-1}. The data was collected at beam energies of $\sqrt{s} = 7$, 8 and 13 TeV for 2011, 2012 and 2015-2018 respectively. The LHCb datasets used in this analysis are summarised in Table 3.1.

Year	Run	$\sqrt{s}$ [TeV]	$\int \mathcal{L}$ [fb^{-1}]
2011	1	7	1.0
2012	1	8	2.0
2015	2	13	0.4
2016	2	13	1.6
2017	2	13	1.7
2018	2	13	2.0

Table 3.1: Summary of data taken by the LHCb experiment during Run I and Run II of data taking. A total of 8.7 fb^{-1} of data is used in this analysis.

This chapter details the preparation of both the LHCb data and the simulation samples, which are used to measure the physical observables of interest in this thesis. It begins by describing the selection criteria applied to the data, followed by corrections to the simulated samples, an analysis of the backgrounds present in the dataset, and the selection efficiency. Several geometry and particle identification variables are introduced in the following chapters, with a summary of their definitions provided in Appendix B.

3.1 Simulation samples

Modern analyses in high-energy physics (HEP) involve the use of simulated datasets generated using Monte Carlo (MC) techniques, which account for both the underlying physics of the pp scattering process and subsequent interaction of particles in the detector, to produce realistic samples. For each year of data taking, simulated samples are generated under conditions that match the configuration of the LHCb detector. As discussed in Section 2.3.3, the LHCb data samples are split roughly evenly between the two magnetic polarities, "Up" and "Down". The simulation samples are separated in the same way.

The simulated MC samples are used in this analysis to aid in the optimisation of selecting signal decay candidates in the dataset, determine the shapes of kinematic distributions, such as the four-body invariant mass of the parent b meson, and to determine the model for the detector efficiency. These samples, produced using EvtGen [150], are used to model kinematic distributions, such as the four-body invariant mass of the b-hadron, of each of the decay modes.

Simulated physics samples are used for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $c\bar{c}$ resonant $B^0 \to J/\psi\pi^+\pi^-$ decay modes for each year of data-taking, as well as for the corresponding B_s^0 modes, $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$. Additionally, samples are produced for the $B^0 \to J/\psi K^{*0}$ control mode and the non-resonant $B^0 \to K^+\pi^-\mu^+\mu^-$ decay. These samples are reconstructed as both the $K^+\pi^-\mu^+\mu^-$ and under the mass hypothesis of $\pi^+\pi^-\mu^+\mu^-$, representing a misidentified background in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ datasets.

Other physics samples are produced, reconstructed as $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays, for backgrounds present in the dataset, which are identified in Section 3.6. These background samples are used to estimate the relative size of the backgrounds compared to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal decay mode.

3.2 Online selection

The candidates are identified either as "Trigger On Signal" (TOS), where the reconstructed object triggers the event to be stored, or "Trigger Independent of Signal" (TIS), where a track is stored because particles from elsewhere in the event are responsible for the triggering decision. The online selection requirements are designed to search for broad conditions at the L0 level, such as the presence of a dimuon system. At the HLT level the events are partially reconstructed and selected if the topology of the system resembles that of b-hadron decays. An event is able to pass multiple requirements.

At the L0 stage, the online selection is based on information from the calorimeters and muon stations. In this analysis, the selection focuses on muons due to their clean

signatures. The first of two criteria requires that a single muon track must have a high transverse momentum (p_{T}). The second is that the product of the p_{T} of a muon pair is large, indicating a dimuon system is present in the decay chain. These two requirements are applied independently. Additionally, there is a requirement that the number of hits recorded in the SPD subdetector is sufficiently small, to control the amount of data being processed, easing the reconstruction of the event.

The HLT1 software makes selection decisions based on the quality of particle tracks. At this stage, the aim is to identify the b-meson decays where at least one track of a final-state particle has a high p_{T} and is displaced from the PV. In Run I, a simpler cut-based approach was used, primarily focusing on muon tracking. In Run II, a Multi-Variate Analysis (MVA) technique was introduced, enabling more refined selection processes which identified tracks with poor reconstruction quality and rejected them. The MVA method considers a range of variables, such as the p_{T} of the particle, the track reconstruction quality, and the number of hits in subdetectors. This multi-variable approach allows for the consideration of correlations between the variables, which the cut-based approach of Run I could not account for.

The HLT2 software aims to select events based on their topology, specifically those of b-hadron decays [155]. A Boosted Decision Tree (BDT) was employed during Run I [156] to determine those events with a decay vertex, that of the J/ψ for instance, with a significant displacement from the primary vertex of the b-meson. The b-meson is required to have decayed into two, three, or four particles. In Run II, this BDT was replaced by the MatrixNet algorithm [157, 158], a BDT method that improves classification performance by dynamically optimising the boosting process. The selection of displaced b-hadron decay vertices is more computationally efficient.

3.3 Stripping the raw data

The loose requirements of the online selection allow many candidates from decay modes which are not the desired signal to pass through. These unwanted candidates need to be removed from the dataset. The initial step in refining the dataset is to apply a series of requirements on kinematic and topological variables related to particle detection and identification. These requirements are generalised to satisfy basic criteria, for a wide range of decay modes.

The requirements at this stage are standardised for analyses studying decays with similar characteristics, specifically to loosely identify $B \to X\mu^+\mu^-$ signal candidates. The selection used in this analysis is widely applied at LHCb for $b \to q\mu^+\mu^-$ analyses. The decay must include a neutral dimuon system in the final state and a hadron, which can be

either a single hadron or a resonance. The numerical values of these requirements, which are based on established cuts used in previous analyses, are summarised in Table 3.2. Some differences exist in the requirements between data-taking years, due to changing detector conditions, the tightest selection is applied uniformly to all samples for consistency, and only these are shown.

For a candidates to be retained in the dataset, it must pass loose requirements on the invariant mass of the four-body final state, ensuring consistency with decays originating from a b-meson. Additionally, the B-meson lifetime of 1.5 ps is exploited to distinguish the meson from other particles which decay closer to the PV. This lifetime corresponds to a typical flight distance (FD) of the B-meson is typically a few millimetres from the interaction point. The IP of reconstructed particles is used to determine whether the track is inconsistent with originating from a PV. The B-meson is required to be consistent with coming from the PV, the IP must be small and the momentum vector should be consistent with the direction vector between the PV and the decay vertex of the B-meson.

Criteria are applied to muons for particle identification, such as a requirement on the $\Delta L(\mu|\pi)$ variable, which represents the difference in the log-likelihood between two hypotheses: that the track corresponds to a muon versus a pion. Further requirements are applied more generally to the events. For instance, the probability that a track is a "ghost" is required to be small. A "ghost" track is a track reconstructed from a combination of different tracks originating from multiple particles; these tracks do not correspond to a single real particle but to background noise. During reconstruction, each track is assigned a probability of being a ghost track, determined using a trained machine learning tool. A threshold is applied to this probability to reduce this source of background track.

3.4 q^2 and $m(\pi^+\pi^-)$ binning scheme

The di-lepton mass squared (q^2) bins used in this analysis are those regularly used in LHCb analyses (see Table 3.3) and align with the branching fraction measurement of this decay mode [89]. The binning for both the q^2 and $m(\pi^+\pi^-)$ distributions are illustrated in Figure 3.4.1.

The low q^2 bin starts at 1.1 GeV^2/c^4, which is above the mass of the $\phi(1020)$. The bins comprising the two $c\bar{c}$ resonances, J/ψ and $\psi(2S)$, are chosen to be broad enough to fully contain the peaks and tails of their respective shapes. The data from these two bins are used for fit validation. The upper limit of the high q^2 bin is $19\,\text{GeV}^2/c^4$, which is close to the kinematic limit. Above $19\,\text{GeV}^2/c^4$, some of the particles are close to being at rest, therefore the angles become ill-defined, making it difficult to parametrise the efficiency in the q^2 dimension.

Particle	Variable	Requirement
B^0	m($\pi^+\pi^-\mu^+\mu^-$)	$\in$ [4.9,7.0] GeV/c^2
	Vertex χ^2/ndf	$<$ 8.0
	$\chi^2_{\rm IP}$	$<$ 9.0
	$\cos\theta_D$	$>$ 0.9999
	$\chi^2_{\rm FD}$	$>$ 121.0
Hadron	m(X)	$<$ 6200 MeV/c^2
	DOCA $\chi^2_{\rm IP}$	$<$ 20.0
	Vertex χ^2/ndf	$<$ 12.0
	$\cos\theta_D$	$>$ -0.9
	Flight Distance χ^2	$>$ 16.0
Dimuon	m($\mu^+\mu^-$)	$<$ 7100 MeV/c^2
	Vertex χ^2/ndf	$<$ 12.0
	$\cos\theta_D$	$>$ -0.9
	Flight Distance χ^2	$>$ 9.0
μ	isMuon	True
	$\Delta L(\mu\|\pi)$	$>$ -3.0
Tracks	Ghost Probability	$<$ 0.35 (Run I)
	Ghost Probability	$<$ 0.5 (Run II)
	min. $\chi^2_{\rm IP}$	$>$ 9.0
Global	SPD Hits	$<$ 600

Table 3.2: Requirements for stripping $B \to X\mu^+\mu^-$ signal candidates from LHCb data. The requirements are separated by particle, variable, and cut. A summary of the definitions of the variables is provided in Appendix B.

The analysis is performed using a single bin of $m(\pi^+\pi^-) \in [620, 920]\,\text{MeV}/c^2$. The $m(\pi^+\pi^-)$ distribution has been shown to contain many resonances in the $B^0 \to J/\psi\pi^+\pi^-$ dataset [159], as introduced in Section 1.2.6, with the ρ^0 (770) the predominant resonance, with a mass of $775.26 \pm 0.23\,\text{MeV}/c^2$. The range aims to remove a large part of the $f_2(1270)$ and potential contributions from the K^0_s at $498\,\text{MeV}/c^2$.

Bin	q^2 [GeV$^2/c^4$]		
Low	1.1	-	6.0
J/ψ	8.0	-	11.0
$\psi(2S)$	12.5	-	15.0
High	15.0	-	19.0

Table 3.3: Binning scheme for q^2. The bins used for the measurements of the rare mode are shaded in blue.

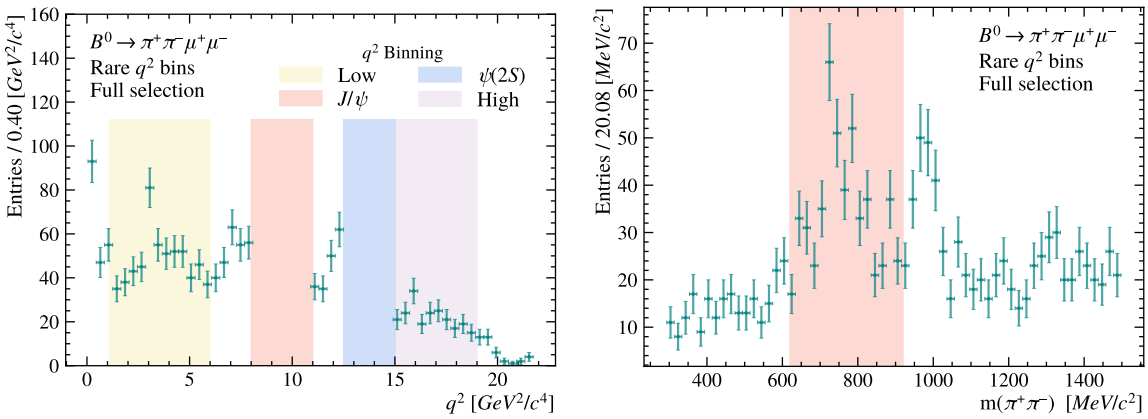


Figure 3.4.1: Binning scheme in the dimensions of q^2 and $m(\pi^+\pi^-)$. Distributions of real data, with a loose selection applied, including the veto of the resonant q^2 bins, are shown alongside illustrations of the different bins used in this analysis. The q^2 bins are summarised in Table 3.3, while a single bin is used for $m(\pi^+\pi^-)$ in the range $\in [0.6, 1.1]\,\text{GeV}/c^2$.

3.5 Combinatorial background suppression

Following stripping selection requirements, many backgrounds can be present in the dataset which pass through the selection due to, for example, similarities in the kinematics or mis-identification of particles. One such background is known as combinatorial background, cases where random combinations of tracks appear to the selection as the signal decays.

Several processes are established for reducing backgrounds, including the previously discussed triggers and offline pre-selection requirements. There are, however, many candidates which will pass these filters and which require a more dedicated approach. A common procedure for reducing the combinatorial background is to perform a multivariate analysis (MVA) which involves training an algorithm to distinguish between a signal and background candidate.

For the algorithm to determine which candidates are signal and which are background two samples are provided as proxies. The signal uses $B^0 \to \rho^0\mu^+\mu^-$ MC, a simulated dataset containing only signal candidates with realistic kinematic distributions. The background sample is taken from the LHCb dataset from a region of the B^0 invariant mass distribution which contains purely combinatorial events, $5800 < m(\pi^+\pi^-\mu^+\mu^-) < 6500\,\text{MeV}/c^2$, which is above the fitting region used in the final analysis. Both samples are processed through the selection stages previously discussed.

The MVA algorithm used in this analysis is a Boosted Decision Tree (BDT), implemented using the XGBoost package [160]. A BDT is a predictive model that combines successive decision trees iteratively to improve accuracy. The process of *boosting* involves training each successive tree to correct the errors made by the previous tree. Each tree

is not accurate individually, but when combined through boosting the process becomes increasingly robust. This iterative process enables BDTs to achieve a high level of accuracy when performing classification tasks, such as determining whether a candidate is signal or background.

A test of the discriminating power of the BDT is the Receiver Operating Characteristic (ROC) curve. The curve plots the true positive rate, the proportion of signal candidates correctly identified, against the false positive rate. The area under the curve, AUC, is a metric used to determine the performance of the BDT. An AUC score of 0.5 (and therefore a diagonal ROC) shows that the classifier cannot distinguish between the samples. A score of 1.0 is optimal in such a case, as the BDT perfectly separates the two samples [161].

Overtraining occurs when the model does not generalise well to independent samples due to performing exceptionally well on the training sample. This happens when the model not only learns the underlying patterns in the data but also irrelevant details, such as noise or small fluctuations. In such an occurrence, the model may learn about spurious correlations or statistical features unique to the training sample, but which are not artefacts of the independent samples. Care must be taken to reduce the risk of overtraining by applying techniques such as K-Fold cross validation [162], ensuring the model is well-trained and capable of accurately categorising the events.

3.5.1 K-Fold Cross Validation

K-Fold Cross Validation is used to prevent overfitting and improve the generalisation of the model. A model is generalised well when it performs well on unseen data, not just the data is has been trained on. This technique involves the splitting of the dataset into multiple subsets, or k "folds" [163]. The model is trained on *k-1* folds and validated on the remaining sample. Each fold is subsequently used as both training and testing samples, until all the samples have been been used as the validation sample. The overall performance of the model is averaged across each of the individually-trained models. This reduces the bias in choosing which of the candidates are used to train and test the algorithm, as each fold is used to both train and test. The testing on multiple folds also provides a measure of the ability of the model to generalise across different, unseen, samples.

In this analysis, the signal and background proxy samples are combined and split into the training and test samples. The training sample is used to teach the BDT algorithm to discriminate between signal and background, while the test sample is reserved for evaluating the model for signs of over-training.

The optimal number of folds is determined by comparing the AUC of the trained model using a range of values for k. The events are shuffled and split into each fold according to

the modulus of the index in the dataset. No significant difference is found when using more than 8 folds, although the computational complexity increases significantly for a larger numbers of folds. In this study, therefore, $k = 8$ folds are determined to be optimal.

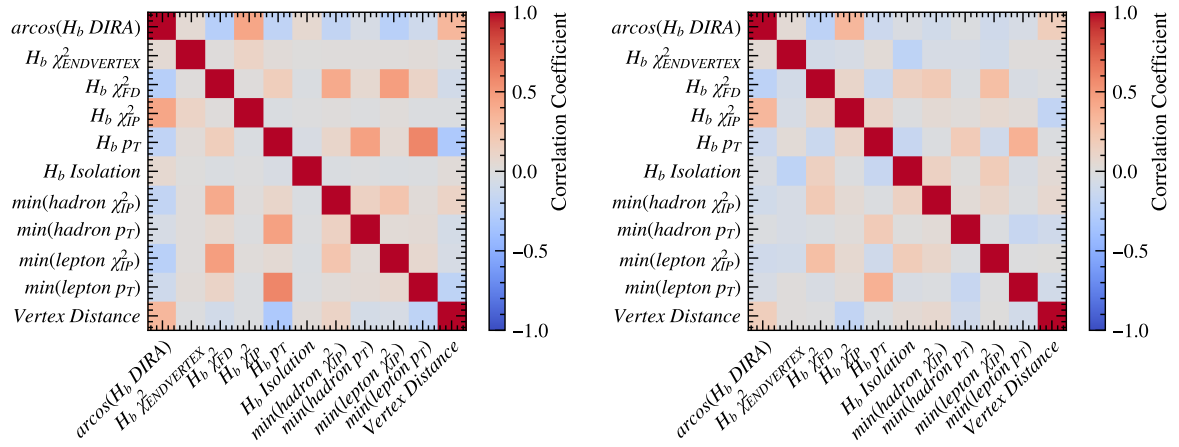
3.5.2 Training features

There is no single ideal set of variables, or features, for training the MVA algorithm. Well motivated variables include those related to kinematics and vertexing information. To avoid overtraining the classifier, a sensible selection of variables is necessary. Starting with a long-list of candidate features, the number can be reduced using various methods.

To narrow down the list of features to an optimal set the Recursive Feature Elimination technique is utilised [164]. The technique works by recursively training the model, removing the least performing feature in each iteration. The process continues until an optimal set is obtained, where removing further features reduces the AUC score of the model by more than 0.005. This threshold was chosen heuristically to balance feature reduction with model performance, ensuring that only features contributing meaningfully to classification are retained. The performance of a feature is determined by the F-score, which reflects both the number of times the feature is used to split the data across all of the trees and the improvement in model performance associated with those splits. A split occurs when the BDT selects a threshold value for a feature that maximises the separation of the target categories, thereby increasing the homogeneity of the resulting subsets.

When using machine learning algorithms, such as a BDT, features are most effective when they contain information which is distinct from one another, including those with correlations which differ between the two samples. While a BDT is generally robust to correlated features, high correlations can lead to inefficiencies and potentially to overtraining. Therefore, minimising the correlation between features is likely to improve the performance of the trained model. The correlations between features in the signal and background proxy samples are given in Figure 3.5.1, for the combined Run I and Run II datasets.

The features used in the training are given in Table 3.4. The features include a range of geometric and kinematic features, as well as the "isolation" parameter. Unlike traditional isolation definitions, which typically count tracks or sum transverse momentum in a cone around the candidate, the isolation parameter used here is determined by performing a vertex fit on tracks within the cone. This approach provides a measure of how well the nearby tracks are consistent with forming a common vertex. The probability derived from the minimum vertex χ^2 of this fit is used as the isolation parameter, with a lower value indicating that the track is more likely to be part of the true signal candidate rather than an unrelated background track. The definitions of these features are given in Appendix



(a) Signal

(b) Background

Figure 3.5.1: Correlations of features used to train the BDT to categorise signal and background data. These correlations use combined Run I and Run II samples although the training is performed separately.

B. An example of the ranking of the feature importance is given in Figure 3.5.2. The trend for the importance is similar for the model trained on both runs and across all folds, the Isolation feature performs the strongest followed be roughly equal importance from the rest of the features. The importance is determined by the F-score, which is a measure of the number of times the feature is used in a split in the tree. In Figure 3.5.3 the distributions of the signal and background proxies are shown for each of the features. These distributions are given for the combined Run I and Run II sample, although the training is performed for each run separately.

3.5.3 Hyperparameter optimisation

Hyperparameters are preset in the XGBoost algorithm, but they can be tuned to optimise the performance of the trained BDT model. While the default hyperparameters are determined to give good performance, tuning can further improve results and potentially reduce the effects of overtraining.

Several hyperparameters are tuned to optimise the performance of the BDT. To determine the optimal values of the hyperparameters, a method is followed which involves randomly searching across ranges of parameter values and evaluating the performance of the model using the AUC score. The set of values which provide the highest AUC score is chosen as the optimal set.

Performing the tests randomly on continuous distributions of values, as opposed to sequentially through a grid of defined values, has several benefits [165]. The random search is more efficient than testing all possible combinations of parameter values in a

Particle	Variable
B^0	χ^2_{vtx}
	χ^2_{IP}
	χ^2_{FD}
	p_{T}
	θ_D
	Isolation
Tracks	$\min(p_{\text{T}})_{\text{hadron}}$
	$\min(p_{\text{T}})_{\text{lepton}}$
	$\min(\chi^2_{\text{IP}})_{\text{hadron}}$
	$\min(\chi^2_{\text{IP}})_{\text{lepton}}$
	Vertex distance ($\pi^+\pi^- \rightarrow \mu^+\mu^-$)

Table 3.4: List of features used to train the BDT algorithm to categorise signal and background candidates. There are a total of six features related to the B^0 and five to the tracks of the final state particles. A description of each feature is given in Appendix B. This is the final list of features obtained by reducing a long-list using the Recursive Feature Elimination [164].

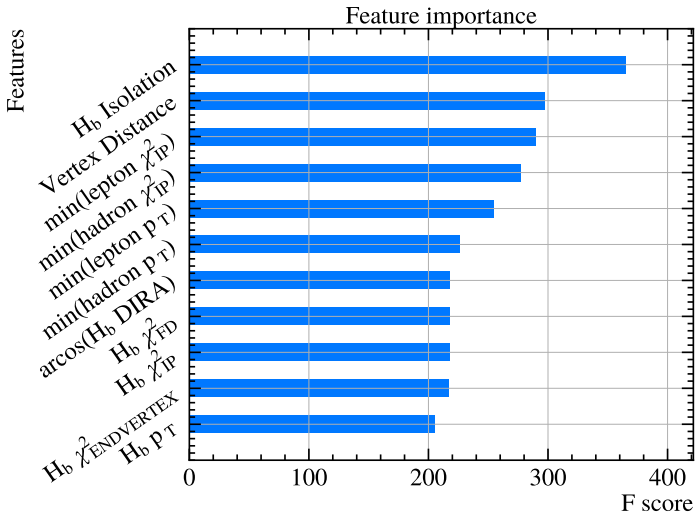
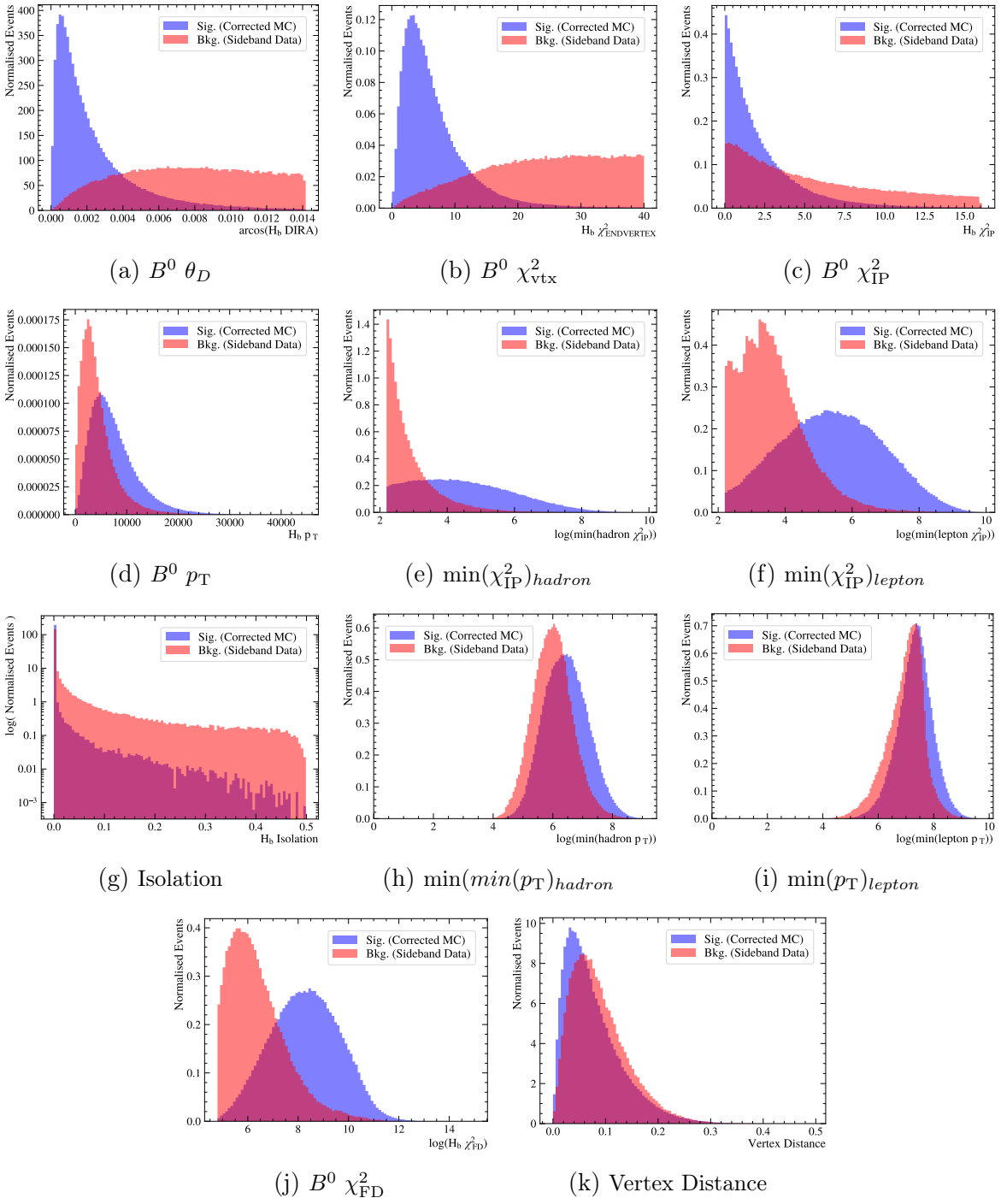


Figure 3.5.2: Example of the ranking of the importance of features for the BDT model. While the training is performed over several folds, the trend is consistent, with the Isolation feature performing the strongest. Here, H_b is a shorthand for the B-meson.

grid and, thus reducing the use of computational resources while achieving similar results. This approach is beneficial when dealing with the hyperparameters, as it allows scalability and avoids the exponential growth in the number of tests involved in the grid search. The tests are performed with the same number of k-folds as in the nominal training of the BDT, ensuring the set of parameters is robust.

The hyperparameters chosen for tuning are commonly optimised when training a gradient boosted decision tree, they are defined as follows:



(a) B^0 θ_D (b) B^0 χ^2_{vtx} (c) B^0 χ^2_{IP}

(d) B^0 p_{T} (e) $\min(\chi^2_{\text{IP}})_{hadron}$ (f) $\min(\chi^2_{\text{IP}})_{lepton}$

(g) Isolation (h) $\min(min(p_{\text{T}})_{hadron}$ (i) $\min(p_{\text{T}})_{lepton}$

(j) B^0 χ^2_{FD} (k) Vertex Distance

Figure 3.5.3: Features used in the XGBoost BDT to categorise events as either signal or background. The distributions are shown for the signal and background proxy samples for the combined Run I and Run II datasets.

- Number of Estimators: Total number of trees in the model. More trees improves the performance of the model but increase the risk of overtraining.
- Learning Rate (η): Controls the contribution of each tree to the final prediction. A lower rate improves the generalisation of the model but increases computation time.
- Maximum Depth: Limits the number of splits in a tree. A deeper tree may model noise, while a shallower tree may overlook patterns in the features.
- Minimum Split Loss: Adds a penalty to the decision of whether to split a node. A larger penalty reduces the likelihood of unhelpful splits, thus preventing overfitting.
- Regularisation: Each node is assigned a weight corresponding to the probability that, if an event ends up in that node, it will be classified into that category. The regularisation parameter applies a penalty to nodes with large weights, promoting simpler models which may generalise better to unseen datasets.
- Feature Subsampling: The fraction of input features used in each trees. This parameter introduces randomness which can help prevent overfitting by reducing the over reliance on a small subset of features.

The optimised parameters for each of the BDTs, which are trained separately on each run, are detailed in Table 3.5. The optimised hyperparameter values differ between Run I and Run II due to variations in data-taking conditions, detector performance, and event selection criteria across the two runs. These differences affect the optimal balance between model complexity and generalisation, making it necessary to tune the hyperparameters separately for each dataset. While it is possible that different feature selections could be optimal for each dataset, no significant evidence suggested that a different set of features would improve classification for either Run I or Run II. Therefore, the same set of features was used for consistency.

A quantification of the optimisation with the default hyperparameter values is given in Table 3.6. While the improvement in AUC is relatively small for both samples, this is expected given the already high baseline performance of the classifier. A more relevant performance indicator in the context of background suppression is the background efficiency at a fixed signal efficiency of 90%. The optimised classifier reduces the background efficiency in both cases by more than 15%. This demonstrates a tangible improvement in separating signal from background. The reduction in background contamination, while maintaining high signal efficiency, highlights the benefit of hyperparameter tuning.

Hyperparameter	Optimised value		Default value
	Run I	Run II	
max depth	8	6	6
n estimators	359	429	100
learning rate (η)	0.05	0.08	0.3
feature subsampling	0.88	0.83	1.0
minimum loss reduction	0.54	0.63	0.0
regularisation	0.52	0.39	1.0

Table 3.5: Optimised values of hyperparameters for the training of a BDT to categories candidates as signal or background. The final column shows the default values for the XGBoost package [160].

Dataset	AUC Score			Background Efficiency		
	Default	Optimised	Gain (%)	Default	Optimised	Gain (%)
Run I	0.996	0.997	4.9	0.0045	0.0038	15.2
Run II	0.997	0.998	3.2	0.0036	0.0030	16.9

Table 3.6: Comparison of AUC scores and background efficiency before and after hyperparameter tuning for the BDT training on Run I and Run II datasets. The gain in AUC represents the percentage improvement of the optimised model relative to the default model, while the improvement in background efficiency reflects the reduction in background contamination at 90% signal efficiency. A lower background efficiency indicates improved background suppression.

3.5.4 Results

The output of a BDT is a score, $f(x)$, which is the sum of the decision tree outputs, which can be transformed into a probability y that the event belongs to category k given the input features x. Specifically, we look at the probability that an event is classified as a signal event. To obtain the probability, the score is transformed using a Logistic (Sigmoid) Function,

$$p(y = k \mid x) = \frac{1}{1 + \exp(-f(x))}, \tag{3.1}$$

which returns a value between 0 and 1. The trained model can be tested for overtraining by comparing the probability distributions of the training and independent testing samples for both the signal and background events. The results for each of the k-folds are given in Figure 3.5.4. The similar distributions for the training and testing samples across all folds suggest that the model generalises well to unseen data, with minimal overtraining.

The distribution of probabilities obtained from the trained BDT model is used to determine a threshold value to cut upon, reducing the combinatorial background component in the datasets. The determination of this threshold is discussed in Section 3.5.5.

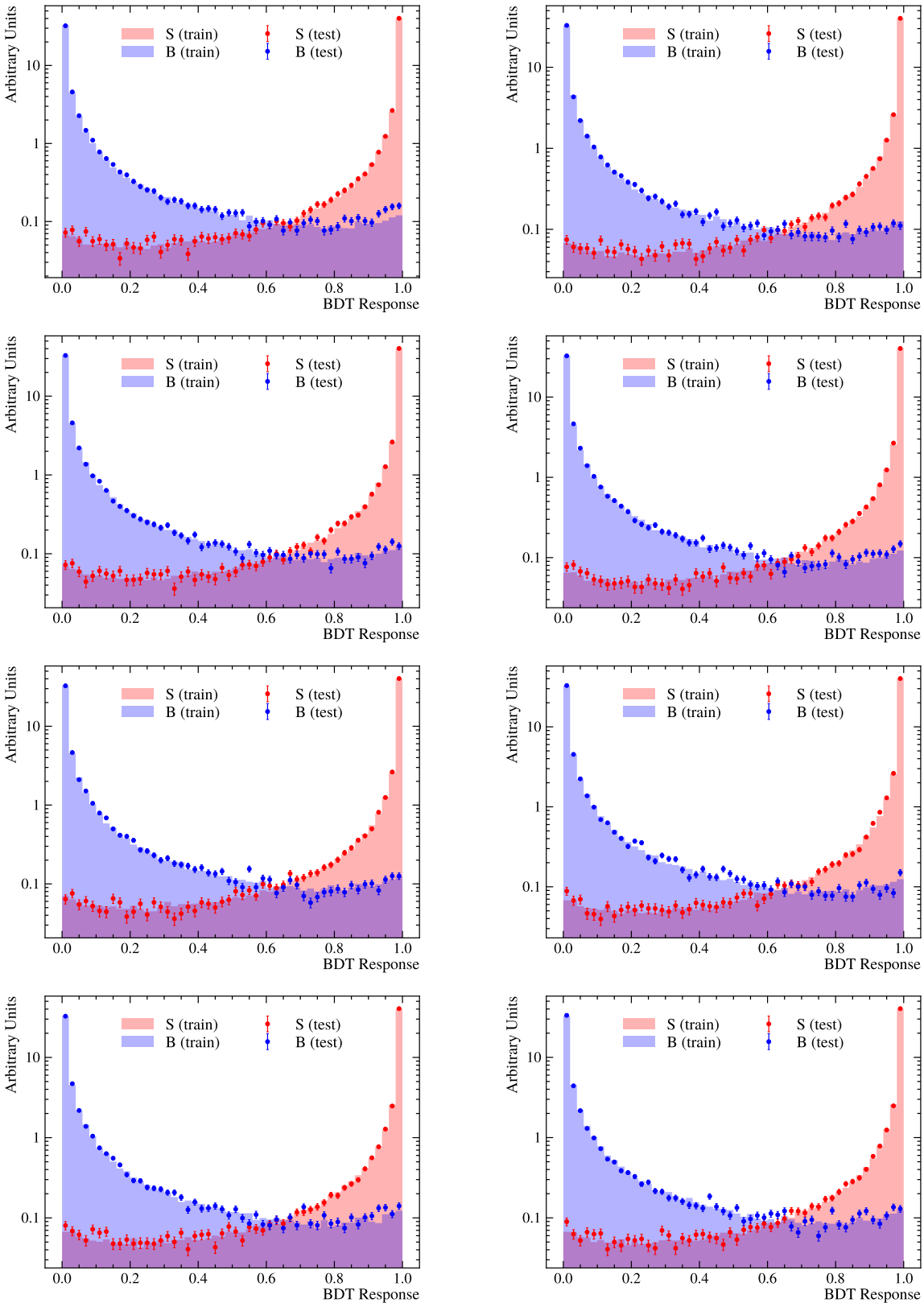


Figure 3.5.4: Output from the BDT trained to classify signal and background candidates using Run II datasets. Each figure shows one of the k-folds used in the cross validation, the training and testing results are shown for the signal and background classes.

3.5.5 Optimisation of BDT and hadron PID requirements

In Section 3.5, a BDT model was trained to classify candidates as either signal or combinatorial background. This section focuses on optimising selection thresholds for the output of the BDT model (O_{BDT}) and variables relating to PID information. This selection aims to minimise both the combinatorial background and those backgrounds resulting from particle misidentification, while maximising sensitivity to the angular observables.

The PID variables represent the probabilities that a track originates from a specific particle species. These probabilities are determined using a trained neural network [166]. Each final state hadron has a separate probability estimate for it to be a proton, pion, or kaon. Since protons are rarely misidentified as other particles, they are omitted from this selection process. A variable is defined with the probability that the hadron is of the desired species, combining the conditional probability that a final-state hadron is a pion, $P(x \mid \pi)$, with the probability that the same hadron is not a kaon, $(1 - P(x \mid K))$:

$$P(h \mid h) = P(x \mid \pi) \times (1 - P(x \mid K)), \tag{3.2}$$

This combined probability enhances the rejection of backgrounds where kaons are misidentified as pions. A loose requirement is also placed on the final-state muons, $P(x \mid \mu) > 0.1$, in the pre-selection stage, which is sufficient, due to the excellent muon PID performance of the LHCb detector.

The distributions of O_{BDT} are found predominantly close to either 0, indicating a background classification, or 1, indicating signal. To better manage this variable, O_{BDT} is transformed as $\log(1 - O_{BDT})$.

Example distributions of $\log(1 - O_{BDT})$ and $P(h \mid h)$ are shown in Figure 3.5.5. The simulation samples of the prominent decay modes are normalised to unit area to compare how the two variables affect each of the components. The hadron PID ($P(h \mid h)$) has a similar shape for the $B^0 \to \rho^0\mu^+\mu^-$, $B_s^0 \to f_0(980)\mu^+\mu^-$, and combinatorial background components, indicating the majority of the candidates are positively identified as $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays. A large portion of the $B^0 \to K^{*0}\mu^+\mu^-$ candidates decline as the variable increases, indicating a loose selection cut would remove a sizeable portion of this misidentified decay mode. The transformed BDT output ($\log(1 - O_{BDT})$) shows that the combinatorial background will be greatly reduced with a loose cut to this variable, and that the other decay modes were identified as signal by the trained BDT.

A two-dimensional scan of these variables is performed to identify the optimal selection thresholds. This scan involves the application of different selection requirements on the $B^0 \to J/\psi\pi^+\pi^-$ dataset and determining the number of candidates related to the signal

and background decay modes. The number of candidates in the rare decay mode are obtained using the ratio of branching fractions and selection efficiencies,

$$N_{rare} = N_{control} \times \frac{\mathcal{B}_{rare}}{\mathcal{B}_{control}} \times \frac{\epsilon_{rare}}{\epsilon_{control}}, \tag{3.3}$$

where N is the number of candidates, $\mathcal{B}$ is the branching fraction, and ϵ the selection efficiency determined using simulation samples. The subscripts *control* and *rare* refer to the $B^0 \to J/\psi\pi^+\pi^-$ and $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay modes, respectively. This normalisation approach using the control mode applies only to the rare decay signal. In contrast, since no simulation samples exist for the combinatorial background, it is determined directly from the upper mass sideband of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset ($5800 < m(\pi^+\pi^-\mu^+\mu^-) < 6500\,\text{MeV}/c^2$), which was also used in training the BDT.

The starting point for the scan uses the selection of $P(h \mid h) > 0.2$ and $\log(1-O_{BDT}) < 0.5$, applied to the $B^0 \to J/\psi\pi^+\pi^-$ dataset. At each point in the scan, the selection efficiency is used to scale the number of events calculated at the original working point. For the combinatorial background the scale is obtained from the efficiency of placing the requirement on the upper mass sideband dataset.

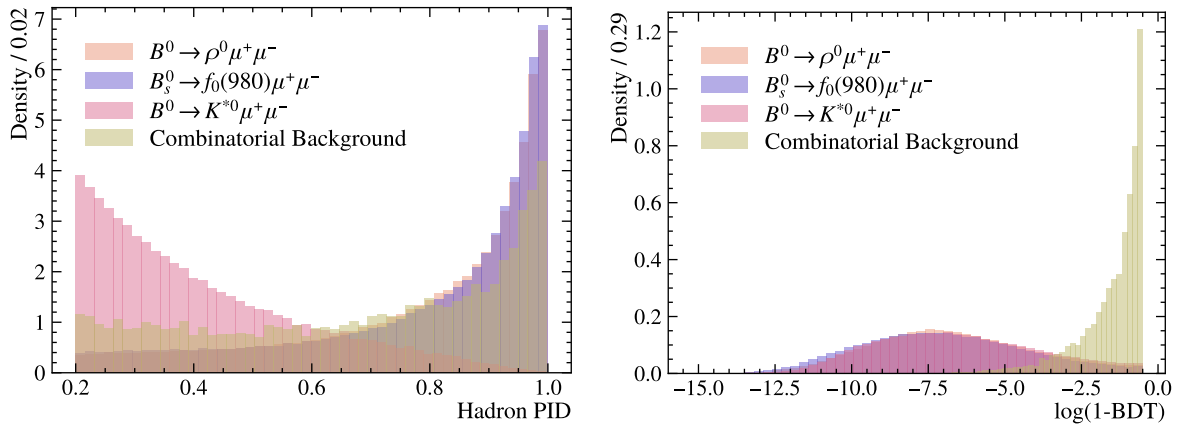
A total of 1000 pseudoexperiments are generated at each point in the scan. The backgrounds included in the study are the peaking backgrounds discussed in Section 3.6, as well as the combinatorial background. For each of the pseudoexperiments, the nominal fitting procedure is then followed, which is discussed later in Section 4. Two metrics are used to judge the results of the study: the first is the significance of the signal, estimated using Wilks' theorem [167], and the second assesses the uncertainties of the angular observables F_L, S_3, A_6, and A_9..

Wilks' theorem is used to compare two hypotheses by employing a likelihood ratio test. The test compares the likelihood for the signal-plus-background model ($\mathcal{L}_{model}$) to the background-only model (($\mathcal{L}_{null}$)), where the signal shape is omitted from the mass model under the null hypothesis. The significance is a measure of the strength of the signal hypothesis, it is calculated using:

$$\sqrt{2 \cdot (\ln(\mathcal{L}_{model}) - \ln(\mathcal{L}_{null}))}. \tag{3.4}$$

The maximum value of the significance across the two-dimensional scan indicates the optimal selection cut.

The results of the scans are shown for the Wilk's test in Figure 3.5.6, and the angular observables in Figure 3.5.7. The heat-maps show agreement between each of the distributions of the uncertainty on the angular observables and the Wilk's test, with the uncertainties exhibiting similar flat regions. The optimal requirements are determined to be $P(h \mid h) > 0.44$ and $\log(1 - O_{BDT}) < -4.7$.



(a) $P(h \mid h)$ (b) $\log(1\text{-}O_{BDT})$

Figure 3.5.5: Distributions of hadron PID ($P(h \mid h)$) and transformed BDT output ($\log(1\text{-}O_{BDT})$), normalised to unit area, using simulation samples of $B^0 \to \rho^0\mu^+\mu^-$ and $B_s^0 \to f_0(980)\mu^+\mu^-$ (dominant P-wave and S-wave $\pi^+\pi^-$ resonances), and $B^0 \to K^{*0}\mu^+\mu^-$ (dominant $K^+\pi^-$ background). Combinatorial background is from the upper mass region of the real $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset. Selection requirements are optimised via a 2D scan of the distributions.

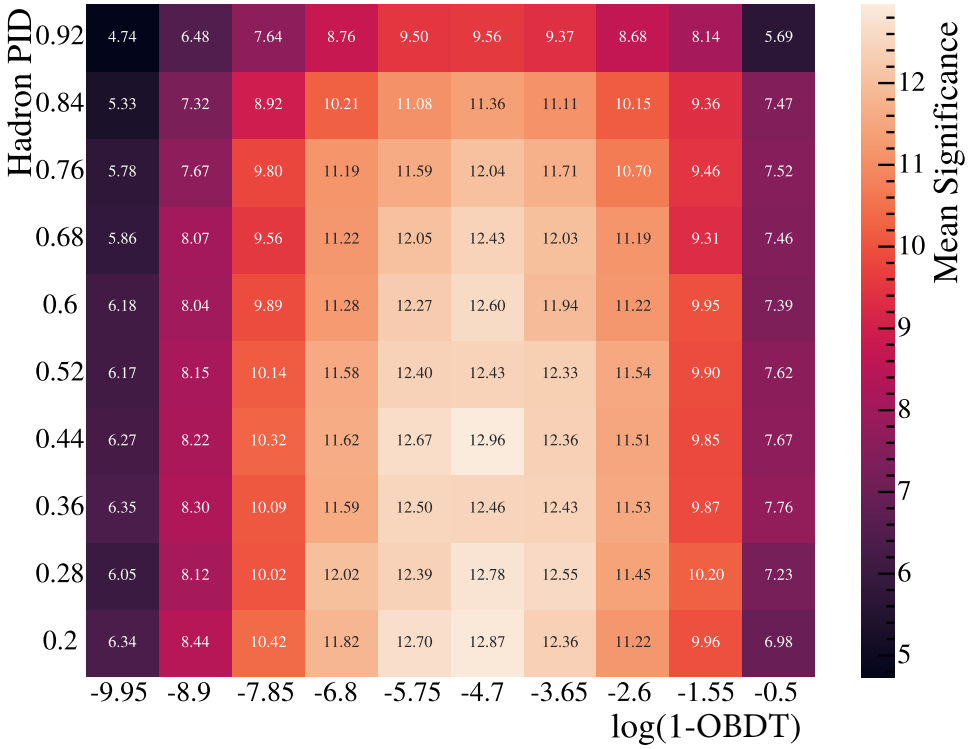
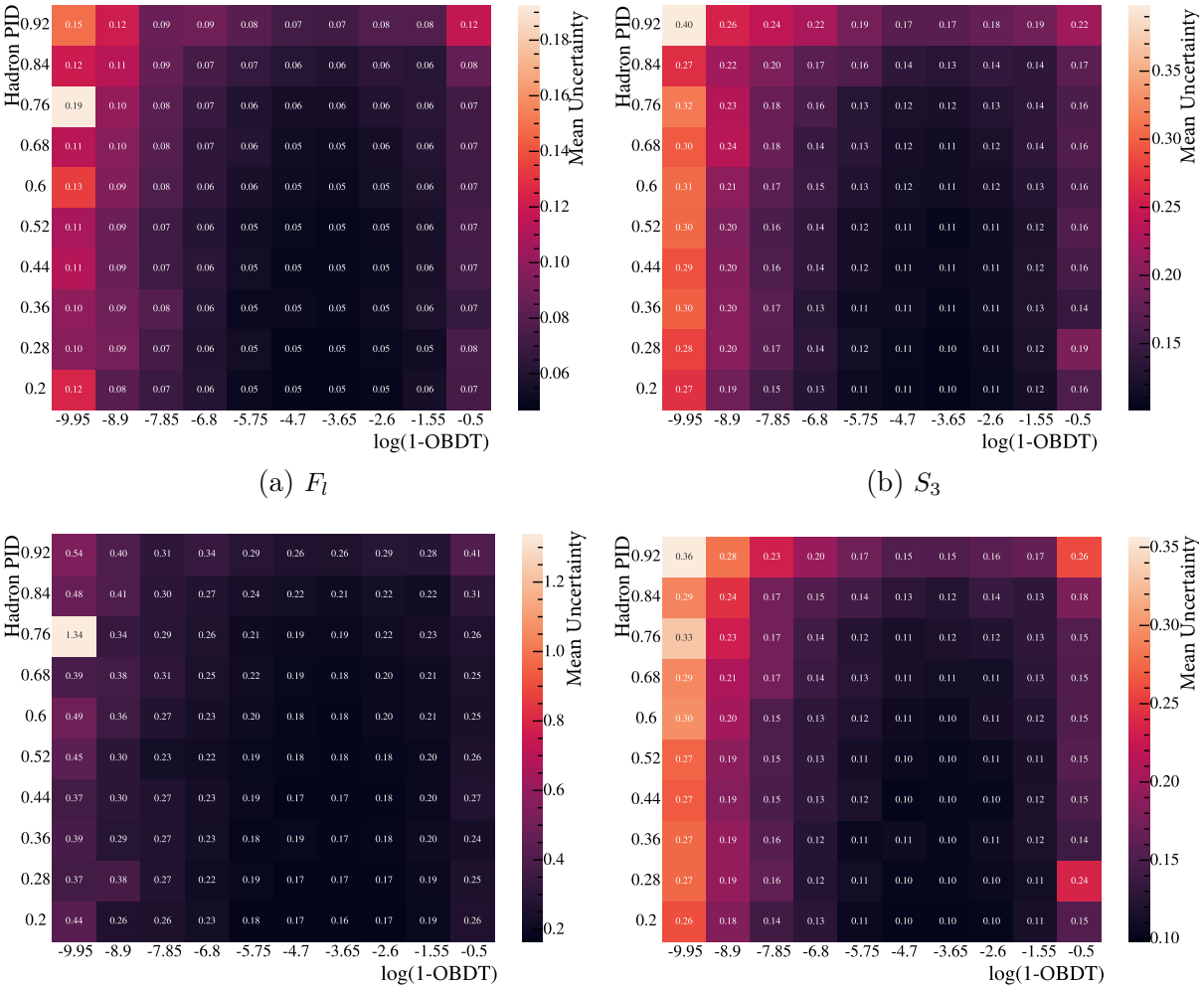


Figure 3.5.6: Heatmap of the 2D scan of the transformed BDT output ($\log(1 - O_{BDT})$) and probability that a hadron is a pion and not a kaon ($P(h \mid h)$). The z axis shows the mean significance from performing Wilks' test on 1000 pseudodatasets at each working point. The highest significance is observed at $P(h \mid h) > 0.44$ and $\log(1 - O_{BDT}) < -4.7$, suggesting an optimal working point for separating signal and background.



(a) F_l

(b) S_3

(c) A_6

(d) A_9

Figure 3.5.7: Heatmap of the 2D scan of the transformed BDT output ($\log(1 - O_{BDT})$) and probability that a hadron is a pion and not a kaon ($P(h \mid h)$). The z axis shows the mean uncertainty of the measured observable on 1000 pseudodatasets at each working point. The highest significance is observed at $P(h \mid h) > 0.44$ and $\log(1 - O_{BDT}) < -4.7$, suggesting an optimal working point for separating signal and background.

3.6 Peaking backgrounds

We turn now to the identification of specific backgrounds associated with the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to J/\psi\pi^+\pi^-$ decay modes that may be present in the data. This investigation looks at the backgrounds within the resonant $B^0 \to J/\psi\pi^+\pi^-$ dataset, with the assumption that similar backgrounds exist in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ sample. The backgrounds tend to derive from the misidentification of b-hadron decays into four final state particles.

The backgrounds which do slip through the selection are suppressed relative to the

selection of the signal decay mode, however due to the small branching fraction of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ compared to other similar decays many backgrounds have a significant presence in the data. This section studies several sources of background and aims to quantify the rate at which they appear.

3.6.1 Misidentified final-state particles

The LHCb detector has demonstrated excellent PID performance during Run I and Run II [131]. However, there are cases where particles are misidentified, leading to backgrounds which are identified as the signal decay but contain a misidentified particle. These cases can become a considerable source of background. As discussed in Section 3.5.5, the variable $P(h \mid h)$, representing the probability that the particle species is correctly identified, has been optimised to select signal candidates. This variable should remove a large contribution from the misidentified backgrounds. In this section, these backgrounds are identified and their impact on the dataset is assessed.

The contribution of each potential background is investigated using the full Run I and Run II datasets in the J/ψ q^2 bin, collected by the LHCb detector. The online selection and stripping requirements are applied to all samples, collectively referred to as "pre-selection". The requirements on $\log(1 - O_{BDT})$ and $P(h \mid h)$ are applied to show the effect of the optimised variables on the backgrounds. When the requirements on these variables are applied, this is referred to as "full selection".

The background contributions in data are investigated by substituting the mass of the reconstructed final state particle, in this case either the relevant π or μ, with the mass hypothesis of the relevant particle for the concerned background, while preserving the measured momentum vector. The momentum is correctly measured by the LHCb tracking system irrespective of any misidentification of particle species. Since the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal decay is charge-symmetric, potential charge misidentifications do not affect the overall charge balance.

3.6.1.1 Backgrounds with hadronic misidentification

The misidentification of either, or both, of the final state hadrons gives rise to many potential backgrounds. Since both the hadrons in the signal decay are pions the study can be limited to the $K \to \pi$ and $p \to \pi$ backgrounds. The RICH detectors are used to identify hadrons in the LHCb detector. At low and high momentum. the reliability of identification is reduced [131], especially for the identification of protons, as seen in Figure 2.4.1.

The potential sources of background are summarised in Table 3.7. The single misidentification backgrounds are given in the top half of the table, although the rate of misidentifi-

cation is low the branching fraction of the backgrounds (some $\mathcal{O}(10^{-5})$) are in many cases larger than the that of the rare signal ($\mathcal{O}(10^{-8})$), some backgrounds are also relatively small due to Cabibbo suppression. An estimate of the probability of misidentifying the pions (ϵ(misID)) is given for each case, from which one can determine a "misidentification" branching fraction,

$$\mathcal{B}_{\text{misID}} = \mathcal{B} \times \epsilon(\text{misID}). \tag{3.5}$$

The probability, ϵ(misID), is determined using an LHCb PID lookup table, produced using the software package PIDCalib [168]. The table takes input values of the PID threshold applied in this analysis, momentum, pseudorapidity, and track multiplicity, and returns a corresponding PID efficiency determined from LHCb data. The values chosen to obtain these estimates are $p = 19\,\text{GeV}$, $\eta = 3.25$, and multiplicity = 50. These values correspond to a region of the detector where misidentified background tracks are likely to appear, with moderate-to-high momentum, forward pseudorapidity, and typical event multiplicity. This makes them representative for estimating the impact of this type of background.

In the following sections these backgrounds are discussed with several figures provided to illustrate the different backgrounds in data. Unless stated otherwise these all show the $B^0 \to J/\psi\pi^+\pi^-$ dataset due to the abundance of the data compared to the rare and $\psi(2S)$ samples.

Misidentification	**Decay Mode**	$\mathcal{O}$**(BF)**		ϵ_{misID}
		J/ψ decay	rare decay	
–	$B^0 \to \pi^+\pi^-\mu^+\mu^-$	10^{-6}	10^{-8}	–
	$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	10^{-5}	10^{-8}	–
$K \to \pi$	$B^0 \to K^+\pi^-\mu^+\mu^-$	10^{-5}	10^{-7}	2.2×10^{-2}
	$B_s^0 \to K^+\pi^-\mu^+\mu^-$	10^{-5}	10^{-8}	2.2×10^{-2}
$p \to \pi$	$\Lambda_b^0 \to p\pi^-\mu^+\mu^-$	10^{-6}	10^{-8}	2.4×10^{-1}
$K^+ \to\pi^+$ & $K^- \to\pi^-$	$B^0 \to K^+K^-\mu^+\mu^-$	10^{-7}	10^{-9}	4.6×10^{-4}
	$B_s^0 \to K^+K^-\mu^+\mu^-$	10^{-5}	10^{-7}	4.6×10^{-4}
$p \to\pi^+$ & $\overline{p} \to\pi^-$	$B^0 \to p\overline{p}\mu^+\mu^-$	10^{-8}	Unknown	5.9×10^{-2}
	$B_s^0 \to p\overline{p}\mu^+\mu^-$	10^{-7}	Unknown	5.9×10^{-2}
$p \to\pi^+$ & $K^- \to\pi^-$	$\Lambda_b^0 \to pK^-\mu^+\mu^-$	10^{-7}	10^{-5}	5.2×10^{-3}

Table 3.7: Hadronic backgrounds from hadron misidentification potentially present in the dataset. The order of magnitude of the branching fractions (BF) of each decay mode [22] is given, both for modes including a $J/\psi \to\mu^+\mu^-$ resonance and for rare decays. BFs give a relative sense of likelihood compared to the signal. Charge conjugation is implied. The quoted BFs for J/ψ modes include $J/\psi \to\mu^+\mu^-$. Entries marked "Unknown" are seen but have no recorded BF. The final column gives an estimated misidentification probability.

$K \to \pi$ misidentification

The $K \to \pi$ background is seen by reassigning the mass hypothesis of the pion to its true mass, that of the kaon. The resulting invariant mass distribution of the four body $K^+\pi^-\mu^+\mu^-$ system, within the resonant $J/\psi \to \mu^+\mu^-$ dataset, is compared to the true mass distribution of the signal decay in Figure 3.6.1. The 2D planes show a comparison of the two distributions, with a one-dimensional (1D) histogram of each sample shown on the relevant axis. The bins with no events have no colour and the less populated bins, in purple, are generally filled with combinatorial background events, which are reduced considerably with the application of the BDT selection.

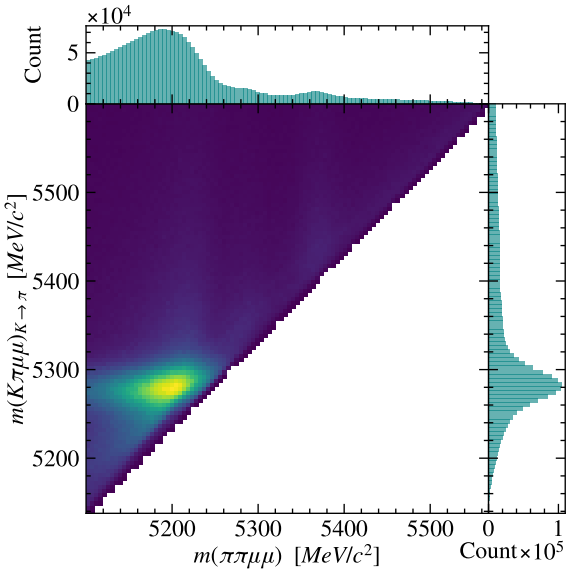
This background is shown, in Figure 3.6.1c, to originate predominantly from the $K^{*0}(892)$ resonance, with a sizeable contribution from the $K^{*0}(1430)$. This background is the largest of those coming from misidentification and therefore handling it is crucial. The reduction of this background is considerable when applying a requirement to the hadron PID by choosing the hadron to identify as a hadron and not as a kaon. These PID conditions are combined as a product and a requirement is placed upon the PID, which is optimised in Section 3.5.5, to be $\text{ProbNN}_\pi \times (1 - \text{ProbNN}_K) > 0.44$.

The correlation can be seen between the two samples in the 2D plane in Figure 3.6.1 prior to applying the BDT and the hadron PID requirements the $B^0 \to J/\psi K^{*0}$ background dominates the sample, a small contribution can be seen from the B_s^0 decay but the signal is swamped. Once the full selection is applied to the dataset one can clearly identify the peaks of the signal and B_s^0 decays and while this background is reduced the $K^+\pi^-\mu^+\mu^-$ peak is still identifiable. This background cannot be removed by a dedicated veto due to the proximity to the signal peak in the four-body mass and the complicated spectrum in the dihadron system. Thus the background requires care when modelling the invariant mass in order to differentiate the events from the signal events.

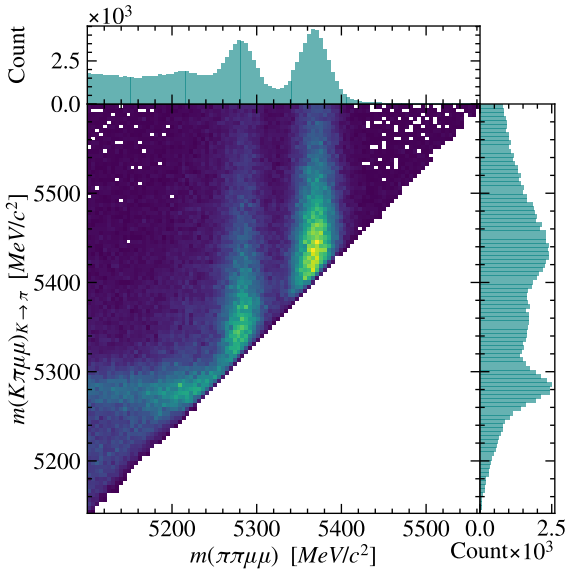
$p \to \pi$ misidentification

Making the substitution of the mass of a pion for that of a proton allows one to examine the distribution of the $p\pi^-\mu^+\mu^-$ four-body final state, corresponding to the $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ decay mode (or its conjugate). In Figure 3.6.2, a diagonal band is visible starting from the upper shoulder of the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ signal peak. This diagonal smearing arises because misidentifying a proton as a pion causes a shift in the reconstructed mass, which depends on the momentum of the misidentified particle. Higher-momentum protons experience a larger mass shift than lower-momentum ones, leading to the observed slanted structure in the distribution. In the 1D projection of the invariant mass distribution, there is no visibly peaking resonance at $5620\,\text{MeV}/c^2$ corresponding to the Λ_b^0 is before the full selection is applied.

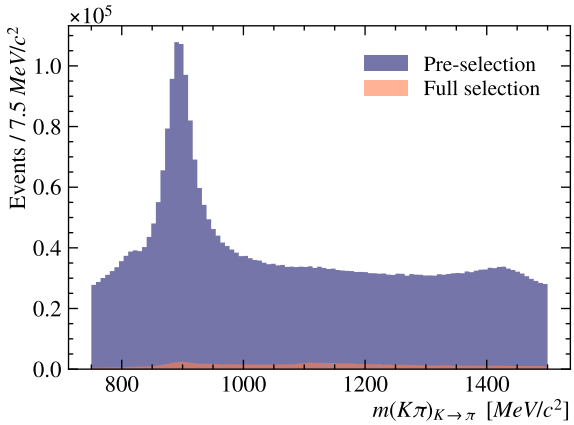
A rough estimate of the relative contribution of this background, based on the branching



(a) Selection without BDT and PID.

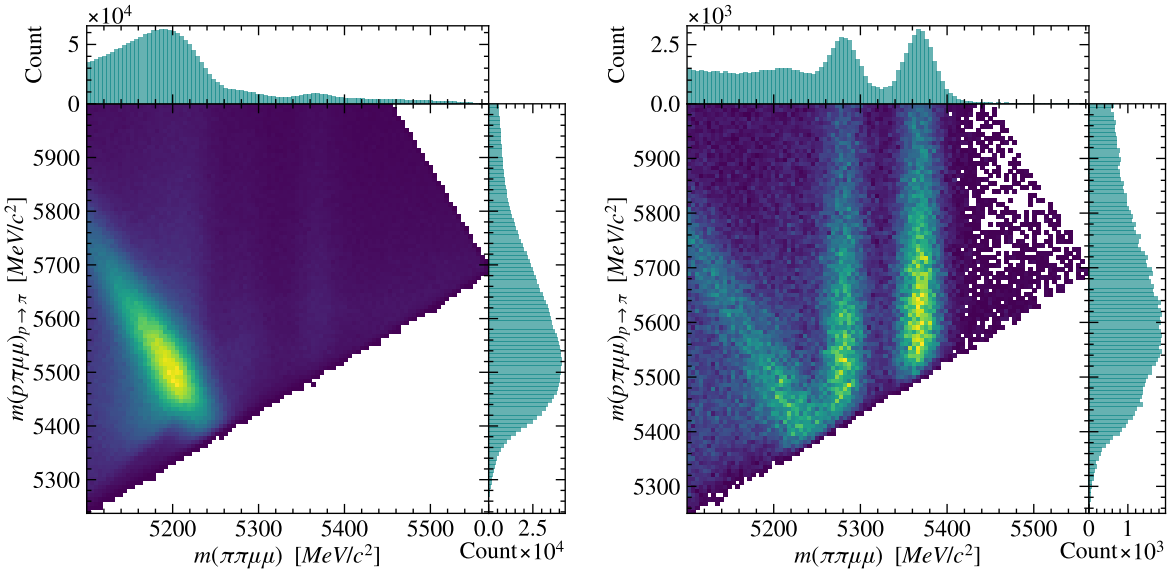


(b) Full selection.



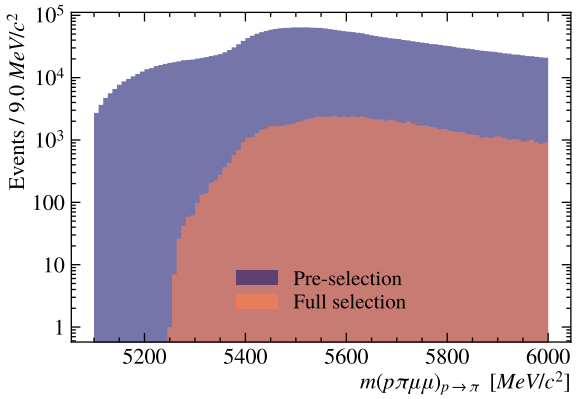
(c) 1D histogram of $m(K^+\pi^-)$ distribution.

Figure 3.6.1: Study of the background of the misidentification of a $K \to \pi$. Figures (a) and (b) compare the invariant mass of the four body $\pi^+\pi^-\mu^+\mu^-$ final state particles in the actual dataset and with the mass hypothesis of a kaon in place of a pion. Figure (a) shows the combined $9\,\text{fb}^{-1}$ LHCb data with triggers and stripping selection applied, (b) has the full selection with the addition of BDT and PID cuts. Figure (c) shows the 1D distribution of the $m(K^+\pi^-)$ invariant mass, with the pre-selection and full selection applied. The range of the mass window is widened from the fit window to illustrate the wider contribution of the component.



(a) Selection without BDT and PID.

(b) Full selection.



(c) $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$

Figure 3.6.2: Study of the background of the misidentification of a $p \to \pi^+$. Figures (a) and (b) compare the invariant mass of the four body $\pi^+\pi^-\mu^+\mu^-$ final state particles in the actual dataset and with the mass hypothesis of a proton in place of a pion. The first figure shows the combined $9\,\text{fb}^{-1}$ LHCb data with triggers and stripping selection applied, the second has the full selection with the addition of BDT and PID cuts. The bottom figure shows the invariant mass of the $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ at 5620 MeV alongside a $\Lambda_b^0 \to pK^-\mu^+\mu^-$ from misidentified $K^- \to \pi^-$. The two Λ_b^0 peaks are removed with the application of the full selection.

fraction $\mathcal{B}(\Lambda_b^0 \to p\pi^-\mu^+\mu^-)$, the misidentification probability from Table 3.7, and the branching fraction of the signal mode, suggests a contribution of around 25%. However, this estimate does not account for the additional selection criteria applied, which significantly reduce this background in practice. Given also that the Λ_b^0 peak lies well above the signal region, the contamination in the signal fit window is minimal. Therefore, the background from $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ is considered negligible in the selected dataset.

$K \to \pi$ double misidentification

Other backgrounds arise from the misidentification of both hadrons in the final state. A prominent example is the double misidentification of K^+K^- as $\pi^+\pi^-$, which can lead to background events from the $B_s^0 \to J/\psi K^+K^-$ decay passing the selection. These kaons predominantly originate from the $\phi(1020) \to K^+K^-$ resonance, which has a branching fraction of $(1.04 \pm 0.04) \times 10^{-3}$ [22], and through non-resonant contributions at $(7.9 \pm 0.7) \times 10^{-4}$.

The double misidentification occurs less frequently than the single-hadron case. As shown in Figure 3.6.3, this background is visible in the pre-selected $B^0 \to J/\psi\pi^+\pi^-$ data. However, the application of the full selection, particularly the PID requirements, suppresses both the K^+ and K^- misidentifications, reducing the background significantly. The contribution of this background is estimated, using Table 3.7, to be less than 1% relative to the signal. No additional vetoes are applied to reduce this background.

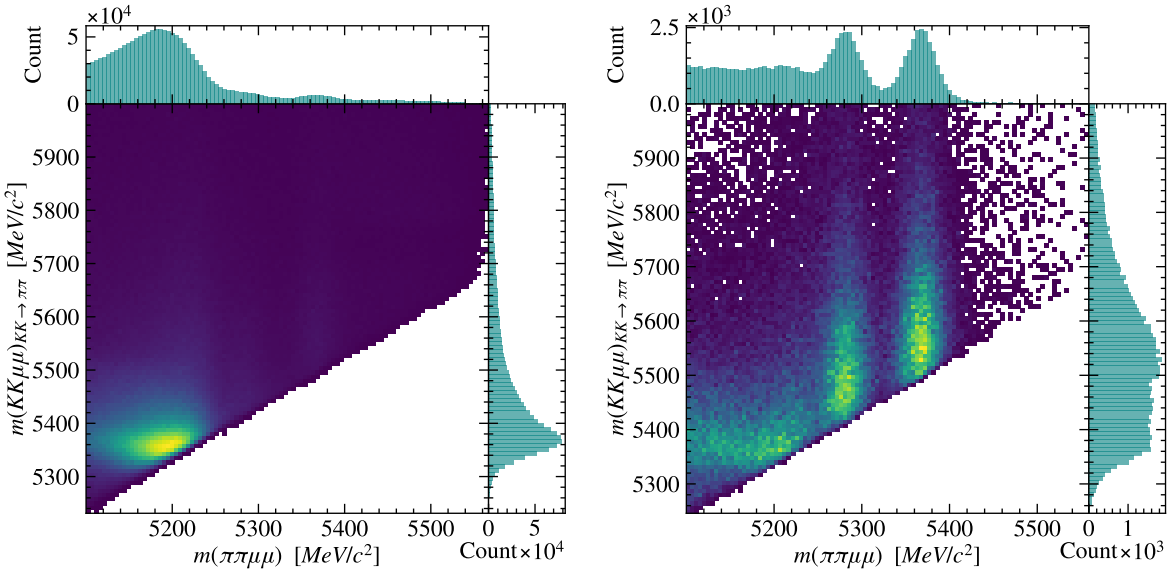
$p \to \pi$ double misidentification

The two decay modes that include a J/ψ and a diproton system are $B^0 \to J/\psi p\bar{p}$ and $B_s^0 \to J/\psi p\bar{p}$. The branching fractions of these decays are relatively low, measured at $(4.5 \pm 0.6) \times 10^{-7}$ and $(3.6 \pm 0.4) \times 10^{-6}$, respectively [169]. No measurements of the corresponding rare decay modes exist at the time of writing. Due to the combined suppression from both the small branching fractions and the low probability of double misidentification from protons to pions, the number of background candidates passing pre-selection is expected to be minimal. An estimate based on Table 3.7 suggests that the contribution is less than 1% relative to the signal.

As shown in Figure 3.6.4, the $\pi^+\pi^-$ dihadron mass spectrum within the fitting range $620\ \text{MeV}/c^2 < m(\pi^+\pi^-) < 920\ \text{MeV}/c^2$ exhibits no significant structure associated with this background. No additional selection is applied to suppress this contribution.

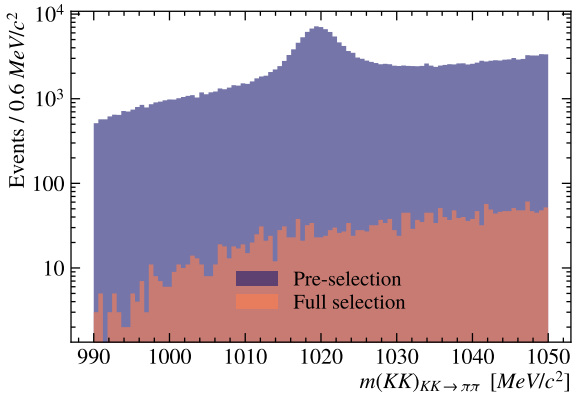
$p \to \pi$ and $K \to \pi$ double misidentification

The $\Lambda_b^0 \to \pi^+K^-\mu^+\mu^-$ decay mode can contribute to the dataset due to the misidentification of $p \to \pi$ and $K \to \pi$. Its branching fraction is significantly larger than that of the $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ decay [88, 170], which is relevant to the $p \to \pi^+$ misidentification background. However, the double misidentification reduces the contribution of this background relative to the signal. Based on estimates from Table 3.7, this background is expected to contribute less than 1% of the signal. The full selection, including the BDT and PID requirements described in Section 3.5.5, significantly suppresses this background, as illustrated in Figure 3.6.5.



(a) Selection without BDT and PID.

(b) Full selection.

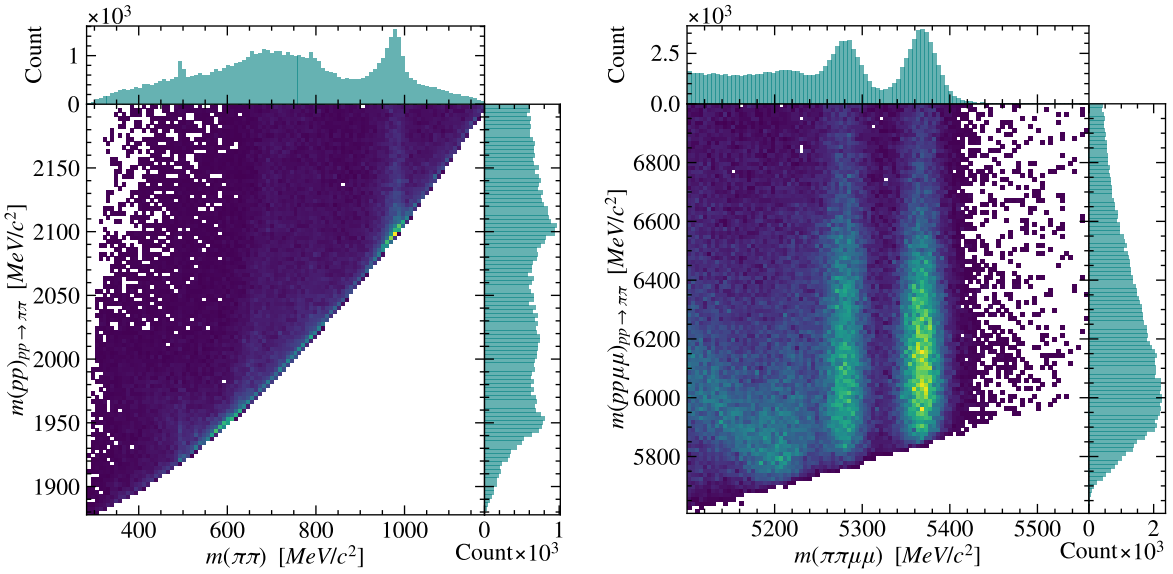


(c) 1D

Figure 3.6.3: Study of the background of the double misidentification of $KK \to \pi\pi$. Figures (a) and (b) compare the invariant mass of the four body $\pi^+\pi^-\mu^+\mu^-$ final state particles in the actual dataset and with the mass hypothesis of two kaons in place of the pions. The first figure shows the combined 9 fb^{-1} LHCb data with triggers and stripping selection applied, the second has the full selection with the addition of BDT and PID cuts. The bottom figure shows the 1D distribution of the $m(K^+K^-)_{K^+K^-\to\pi^+\pi^-}$ invariant mass, with the pre selection and with the full selection applied.

3.6.1.2 Backgrounds with $\mu \to \pi$ misidentification

The most problematic background resulting from leptonic misidentification arises from the $B^0 \to J/\psi\pi^+\pi^-$ decay, where both $\mu \to \pi$ and $\pi \to \mu$ misidentifications occur. This can be visualised by exchanging the mass hypotheses of the muon and pion tracks. In Figure 3.6.6, a small peak is visible near the J/ψ mass, originating from this effect. The peak is substantially reduced after the application of the full selection, particularly due to



(a) $m(\pi^+\pi^-)$ vs $m(p\overline{p})_{p\overline{p}\to\pi^+\pi^-}$

(b) $m(\pi^+\pi^-\mu^+\mu^-)$ vs $m(p\overline{p}\mu^+\mu^-)_{p\overline{p}\to\pi^+\pi^-}$

Figure 3.6.4: Study of the background of the double misidentification of $p\overline{p} \to \pi^+\pi^-$. The invariant mass of the four body $\pi^+\pi^-\mu^+\mu^-$ final state particles in the $B^0 \to J/\psi\pi^+\pi^-$ dataset and with the mass hypothesis of a $p\overline{p}$ pair in place of the dipion system. The figure shows the combined $9\,\text{fb}^{-1}$ LHCb data with the full selection of the trigger and stripping requirements applied as well as PID and BDT requirements.

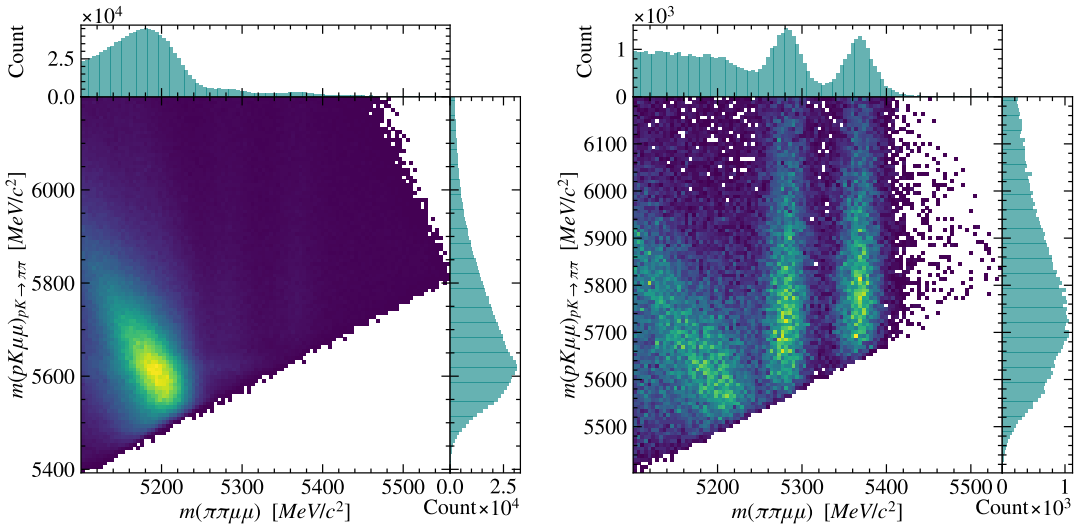
the optimised hadron PID requirements.

Using the same approach employed to derive the misidentification probabilities in Table 3.7, the combined probability for this double misidentification is determined to be $p(\mu \to \pi \& \pi \to \mu) \approx 9.4 \times 10^{-4}$. This yields an estimated contribution of less than 0.3% relative to the signal. While small, this background is not negligible, and a dedicated veto of $\mid m(\mu\pi) - 3097 \mid > 35\,\text{MeV}/c^2$ is applied to suppress it.

3.6.2 Partially Reconstructed Backgrounds

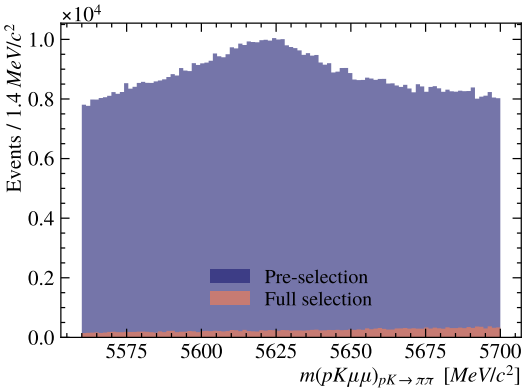
Some decay modes are picked up by the selection which have final states containing the same particles as the signal mode $\pi^+\pi^-\mu^+\mu^-$, but with a missing particle, such as a γ or a π^0. Such examples include the decay modes $B^0_s \to J/\psi(\phi \to \pi^+\pi^-\pi^0)$ and $B^0_s \to J/\psi(\eta' \to \rho^0\gamma)$, which include $\rho^0 \to \pi^+\pi^-$ decays. The same processes exist for the signal, with no $c\bar{c}$ resonance. These decays produce a shape in the $m(\pi^+\pi^-\mu^+\mu^-)$ spectrum smeared to lower values than $m(B^0)$ due to the missing energy of the π^0 or photon. The fit window of the four body invariant mass, discussed later in Section 4.1, is $m(\pi^+\pi^-\mu^+\mu^-) \in [5150, 5650]\,\text{MeV}/c^2$. The lower bound of this window suppresses the contribution of the partially reconstructed backgrounds.

By comparing the branching fractions and selection efficiencies of rare and J/ψ decay



(a) Selection without BDT and PID.

(b) Full selection.



(c) $m(pK^-\mu^+\mu^-)$

Figure 3.6.5: Study of the background of the misidentification of a $p \to \pi$ and $K \to \pi$. Figures (a) and (b) compare the invariant mass of the four body $\pi^+\pi^-\mu^+\mu^-$ final state particles in the $B^0 \to J/\psi\pi^+\pi^-$ dataset and with the mass hypothesis of the pK^- in place of the dipion system. The first figure shows the combined $9\,\text{fb}^{-1}$ LHCb data with triggers and stripping selection applied, the second has the full selection with the addition of BDT and PID cuts. The bottom plot shows the invariant mass at the Λ_b^0 peak with a clear reduction when the full selection is applied.

modes, the backgrounds contribute which contribute significantly in the sample can be determined. The expected yield of each background (N_x) can be determined relative to the control mode yield ($N_{B^0\to J/\psi\pi^+\pi^-}$) using

$$N_x = N_{B^0\to J/\psi\pi^+\pi^-} \times \frac{\epsilon_x}{\epsilon_{B^0\to J/\psi\pi^+\pi^-}} \times \frac{\mathcal{B}(x)}{\mathcal{B}(B^0 \to J/\psi\pi^+\pi^-)} \times \frac{f_q}{f_d}, \tag{3.6}$$

where ϵ is the efficiency of the selection, $\mathcal{B}$ is the branching fraction of the decay mode, and f_q/f_d is the ratio of fragmentation fractions, where $q = d, u, s, c$. The branching

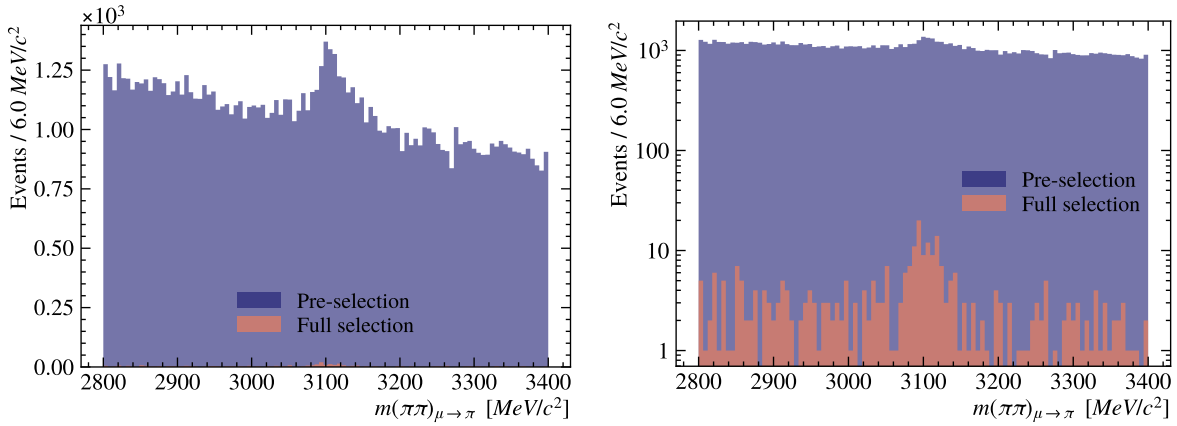


Figure 3.6.6: Study of the background of the double misidentification of $\mu^+\mu^- \to \pi^+\pi^-$. The invariant mass of the two body $\pi^+\pi^-$ system in the rare dataset is shown with pre selection applied and the full selection, in linear and logarithmic scale. A peak is seen above the combinatorial background at the J/ψ peak prior to the application of the optimised hadronic PID requirements.

fractions of several backgrounds and their estimated contributions as a percentage of $N_{B^0\to J/\psi\pi^+\pi^-}$ are given in Table 3.8. There are a considerable number of possible decay modes which can be considered here, however only a selection of those with branching fractions comparable to the signal mode are listed in the table, for brevity.

The calculations are performed using samples generated using the RapidSim package [171], generated within the LHCb geometry and reconstructed as the signal decay mode. The efficiency, ϵ_x, is the product of applying fiducial cuts, $m(B^0)$ and $m(\pi^+\pi^-)$ cuts, q^2 cuts, and PID selection. It is assumed that the efficiencies for each background with the same final-state particles is similar to the signal mode. As the selection primarily depends on kinematics and PID requirements, the efficiencies are expected to be similar. This assumption averts the need to produce dedicated MC-simulated samples for each potential background. In general, this leads to a potential overestimate of the yield, as the majority of the backgrounds will have a lower selection efficiency than the signal.

The ratio of fragmentation fractions f_s/f_d has been measured at 0.2539 ± 0.0079 (at 13 TeV centre-of-mass energy) [172], while the ratio f_c/f_d is predicted to be much smaller, ≈ 0.014 [173]. The ratio f_u/f_d is taken to be 1 as the relation $f_u = f_d$ is assumed to hold, due to isospin symmetry [172]. Backgrounds deriving from the B_c^+ meson are considered, such as the $B_c^+ \to (J/\psi \to \mu^+\mu^-)\pi^+\pi^+\pi^-$ decay. These modes are smeared down in mass from the B_c^+ at $6274.47 \pm 0.27 \pm 0.17$ MeV [174]. Given the large mass gap from the B^0, of 995 MeV, and both the selection efficiencies and the predicted f_c/f_d ratio, the contribution from the partially-reconstructed B_c^+ backgrounds is determined to be negligible.

Similarly, potential partially reconstructed backgrounds from Λ_b^0 decays are considered. There are not many decay modes originating from an Λ_b^0 baryon which are potential

backgrounds of this type, the one with the largest branching fraction is the $\Lambda_b^0 \to (J/\psi \to \mu^+\mu^-)p\pi^+\pi^-K^-$ decay. Here, the decay would be reconstructed if two particles are missed by the selection algorithms. Therefore, the contribution of this mode is expected to be small, and is estimated to be negligible.

Using the results shown in Table 3.8, it is expected that the $B_s^0 \to J/\psi(\phi \to \pi^+\pi^-\pi^0)$ and $B_s^0 \to J/\psi(\eta' \to \rho^0\gamma)$ decays are significant in the J/ψ dataset. The shapes of these two backgrounds are shown in Figure 3.6.7, illustrating their distributions in the $m(\pi^+\pi^-\mu^+\mu^-)$ fit window, smeared down in mass from the B^0 mass peak.

In the rare dataset, the same backgrounds can be considered as are predicted to be significant in the J/ψ dataset. Although naturally a smaller sample, the partially reconstructed backgrounds may still be present. The branching fraction of the $B_s^0 \to \phi\mu^+\mu^-$ decay mode is $(8.4 \pm 0.4) \times 10^{-7}$, it is an order larger than the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ branching fraction of $(2.1 \pm 0.5) \times 10^{-8}$ [22]. However, the ratio of fragmentation fractions is $f_s/f_d = 0.2539 \pm 0.0079$ and the efficiency of applying the selection to the simulated $B_s^0 \to \phi\mu^+\mu^-$ candidates is approximately 0.2%, as estimated using the large pseudodataset. Consequently, the expected contribution of $B_s^0 \to \phi\mu^+\mu^-$ in the fit region is about 2% of the signal yield. Given the available data sample size, this component is too small to be considered significant. The $B^0 \to \eta'\mu^+\mu^-$ branching fraction has not been measured but can be estimated using the $B_s^0 \to \phi\mu^+\mu^-$ branching fraction as

$$\mathcal{B}(B^0 \to \eta'\mu^+\mu^-) \simeq \mathcal{B}(B_s^0 \to J/\psi\eta') \times \frac{B_s^0 \to \phi\mu^+\mu^-}{B_s^0 \to J/\psi\phi}. \tag{3.7}$$

The ratio of this estimated branching fraction to that of $B^0 \to \pi^+\pi^-\mu^+\mu^-$ is 0.014 ± 0.001, and approximately 13% of simulated $B^0 \to \eta'\mu^+\mu^-$ candidates fall within the fit region. Consequently, the expected contribution of this background is only about 0.2% of the signal yield, making it negligible given the available data sample size.

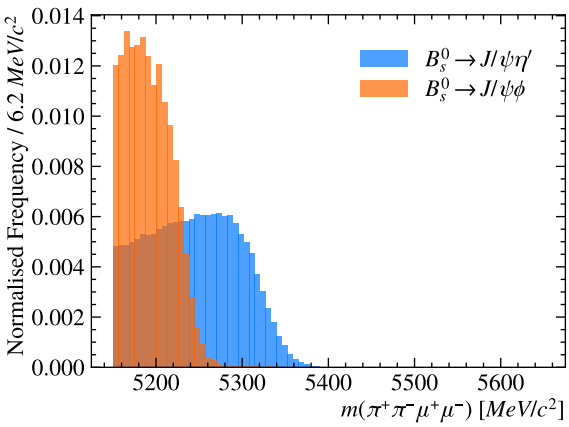
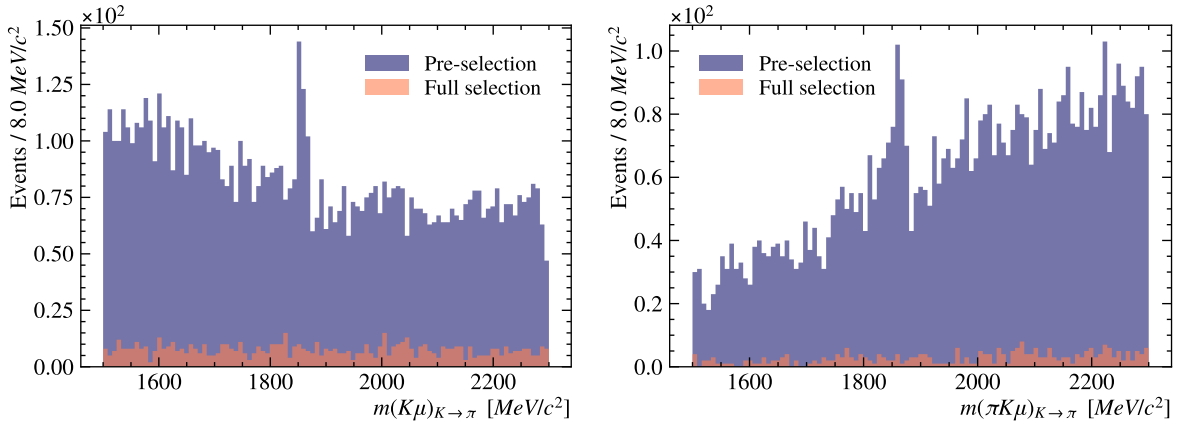


Figure 3.6.7: The shapes of partially reconstructed decay modes, $B_s^0 \to J/\psi(\phi \to \pi^+\pi^-\pi^0)$ and $B_s^0 \to J/\psi(\eta' \to \rho^0\gamma)$, in the $B^0 \to J/\psi\pi^+\pi^-$ dataset are shown in the B^0 mass window. The decays are missing either a photon or π^0, which shifts the shape down in mass from the B^0 due to the loss in momentum. The samples are from MC simulation and are generated within the LHCb geometry without selection cuts, as they are used to estimate the fraction of events within the signal region rather than for direct comparison with the signal. The distributions are normalised to unit area.

Decay Mode	***J/ψ* decay $\mathcal{O}$(BF)**	**Fraction of *J/ψ***
$B^0 \to J/\psi(\omega \to \pi^+\pi^-\pi^0)$	10^{-8}	0.00
$B^0 \to J/\psi(\eta \to \pi^+\pi^-\pi^0)$	10^{-7}	0.00
$B^0 \to J/\psi(\eta' \to \pi^+\pi^-\pi^0)$	10^{-9}	0.00
$B_s^0 \to J/\psi(\phi \to \pi^+\pi^-\pi^0)$	10^{-7}	3.11
$B_s^0 \to J/\psi(\eta' \to (\omega \to \pi^+\pi^-)\gamma)$	10^{-9}	0.01
$B_s^0 \to J/\psi(\eta \to \pi^+\pi^-\gamma)$	10^{-5}	0.01
$B_s^0 \to J/\psi(\eta \to \pi^+\pi^-\pi^0)$	10^{-6}	0.00
$B_s^0 \to J/\psi(\eta' \to \rho^0\gamma)$	10^{-6}	29.32
$B_s^0 \to J/\psi(\eta' \to \pi^+\pi^-\pi^0)$	10^{-8}	0.01
$B^+ \to K^+\pi^+\pi^- J/\psi$	10^{-6}	0.00
$B^+ \to K^+K^-\pi^+ J/\psi$	10^{-6}	0.00
$B_c^+ \to (J/\psi \to \mu^+\mu^-)\pi^+\pi^+\pi^-$	†	0.00
$\Lambda_b^0 \to (J/\psi \to \mu^+\mu^-)p\pi^+\pi^- K^-$	10^{-6}	0.00

Table 3.8: Partially reconstructed background decays in the dataset. The branching fraction (BF) of each mode [22] is shown for decays with a J/ψ resonance decays. Charge conjugation is implied throughout. For J/ψ modes, the BF includes $J/\psi \to \mu^+\mu^-$. Entries labelled "†" have been seen but lack experimentally measured values, with contributions estimated from selection efficiency. The final column gives the background yield relative to J/ψ events, where zero indicates a predicted yield of $\ll 1\%$.



(a) $m(K^+\mu^-)_{K^+\to\pi^+}$ (b) $m(K^+\pi^-\mu^-)_{K^+\to\pi^+}$

Figure 3.6.8: Study of the background of the $b \to c \to s$ cascading decays. Figures (a) and (b) show the invariant mass spectra of the two body $m(K^+\mu^-)$ and three body $m(K^+\pi^-\mu^-)$ distributions, with the misidentification of a K^+ to a μ^+. A peak at the $D^\pm$ mass results are the result of reconstructing candidates with missing energy from unreconstructed neutrinos and K^+ to π^+ misidentification. Both figures show the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset in the rare q^2 regions with the pre selection and with the full selection applied.

3.6.2.1 Cascading decays

Another source of background comes from partially reconstructed cascade decays. These are decays where an event is reconstructed with a missing intermediate decay mode. The predominant decays of this type are semileptonic decays, whereby neutrinos are not reconstructed and the b-meson mass is shifted due to the missing energy. There are two main contributions to this background for this analysis. The first are from decays which proceed through the $b \to c \to d$ transitions via the decay mode $B^0 \to D^-\mu^+\nu$ and subsequently the $D^- \to \rho^0\mu^-\overline{\nu}$ decay mode, where the ρ^0 decays to a dipion system. A second series of decays proceed through the $b \to c \to s$ transitions, with the D^- instead proceeding through the $D^- \to K^{*0}\mu^-\overline{\nu}$ decay mode. Finally there is a misidentification of the K^+ to π^+ to give the final state of $\pi^+\pi^-\mu^+\mu^-$ with missing ν and $\overline{\nu}$.

In data, no peak is observed at the $D^\pm$ mass in resonant $c\bar{c}$ regions of q^2. In Figure 3.6.8 the masses of the $m(\pi^+\pi^-\mu^-)_{K^+\to\pi^+}$ and $m(\pi^+\mu^-)_{K^+\to\pi^+}$ are studied in the rare dataset. A small contribution is seen in the two body mass, this is completely reduced with the application of the full selection. A veto could be applied to $\mid m(\mu\pi) - 1865 \mid > 20\,\mathrm{MeV}/c^2$. Since no significant resonance remains after the full selection, and the residual background contribution is negligible, an additional veto is not required.

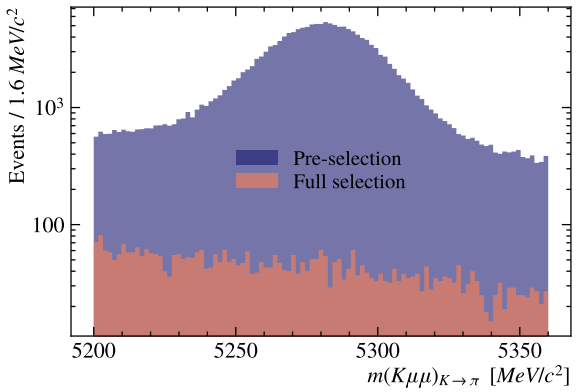


Figure 3.6.9: The invariant mass of the $K^+\mu^+\mu^-$ from the over reconstructed $B^+ \rightarrow J/\psi K^+$ decay mode. The 3-body final state system is associated with a random π^- and the K^+ is misidentified as a π^+. The decay is thus reconstructed as the signal $B^0 \rightarrow \pi^+\pi^-\mu^+\mu^-$ decay mode. The spectrum is shown with pre selection and with the full selection applied, a prominent peak is reduced entirely by the optimised PID requirement.

3.6.3 Over-reconstructed backgrounds

An event can pass selection due to over-reconstruction, where an unrelated, random particle is mistakenly reconstructed alongside the true final state particles. The predominant example in this analysis' dataset is the $B^+ \rightarrow J/\psi\pi^+$ decay, which could be reconstructed alongside a random π^-, allowing it to pass the selection requirements as an $B^0 \rightarrow \pi^+\pi^-\mu^+\mu^-$ decay.

Such three-body decays that include a random particle tend to have higher invariant mass in the four-body distribution, appearing above the mass of the B^0. This background can present a challenge in a mass fit, as in this region it is assumed that the data is purely combinatorial, i.e. consisting of only random combinations of particles.

There is also the potential for a background originating from the combination of an over-reconstructed background and a particle misidentification. the prime example from this dataset is the $B^+ \rightarrow J/\psi K^+$ decay. In the three-body distribution, the background has a distinct peak at the B^+ mass. When applying the full selection requirements this background is significantly reduced. An additional veto is applied around the B^+ mass region to ensure that this background is suppressed further, of $\mid m(K^+\mu^+\mu^-) - 5279 \mid > 50\,\mathrm{MeV}/c^2$. Other hadronic misidentifications were investigated, and no significant peak was observed in the $m(\pi^+\mu^+\mu^-)$ distribution prior to selection. As a result, no additional veto was deemed necessary.

3.6.4 Summary of background search

Following the investigation into the backgrounds we can conclude which backgrounds remain in the signal window following the application of the full selection. The list of decay modes which require consideration when modelling are summarised in Table 3.9. The combinatorial background also remains following the application of the trained BDT.

Those backgrounds resulting from the misidentification of hadrons are heavily suppressed by the application of hadronic identification requirements. However, the $B^0 \to J/\psi K^{*0}$ decay mode continues to have a large contribution near the B^0 peak. The leptonic misidentification backgrounds are less of a concern, as they are almost negligible following the full selection process. The other backgrounds which remain in the signal window are partially reconstructed. In the J/ψ and $\psi(2S)$ resonant datasets their contribution requires treatment when modelling the invariant mass of the B^0, in the rare sample their contribution is predicted to be small enough to be negligible with the current sample size.

Decay Mode	Reason for consideration
J/ψ and $\psi(2S)$	
$B^0 \to J/\psi K^+\pi^-$	Misidentification
$B^0_s \to J/\psi\pi^+\pi^-$	Indistinguishable from signal
$B^0_s \to J/\psi(\phi \to \pi^+\pi^-\pi^0)$	Partially reconstructed
$B^0_s \to J/\psi(\eta' \to \rho^0\gamma)$	Partially reconstructed
Rare	
$B^0 \to K^+\pi^-\mu^+\mu^-$	Misidentification
$B^0_s \to \pi^+\pi^-\mu^+\mu^-$	Indistinguishable from signal

Table 3.9: Backgrounds which remain following the investigation into the peaking backgrounds present in the LHCb dataset following BDT and PID requirements.

3.7 Aligning simulated samples to LHCb data

There are disagreements between the LHCb data and the MC simulation samples, this discrepancy is corrected using weights to bring the MC into alignment. Correcting the simulation samples is important for optimising the multivariate selection discussed in Section 3.5 and for the modelling of the efficiency effects on the angular distributions in Section 4.3.

The applied corrections, discussed in the following sections, include adjustments to the particle identification of the final state particles, kinematics in multiple dimensions,

and corrections to the decay model. The different weights are calculated independently, with the total weight being the product of individual corrections.

3.7.1 Particle Identification (Hadrons)

The hadron PID variables exhibit some differences between the simulation and LHCb data samples, which can lead to inaccuracies in the evaluation of selection efficiency and background studies of decays involving misidentified particles. These differences arise due to imperfections in the detector response modelling in simulation, such as variations in detector conditions. For example, the response of the RICH detectors, which provide particle identification information, depends on factors such as the refractive index of the gas radiators and the calibration of photon detection, both of which may not be perfectly reproduced in the simulation.

The tool used for these corrections is PIDCorr [175], developed at LHCb. Four distributions: the PID response, $log(p_{\mathrm{T}})$, η, and $log(multiplicity)$, where multiplicity is the number of candidates in an event, are described using kernel density estimation (KDE). These variables are used to construct a function which transforms the PID variables in the simulation data. The KDE is modelled using datasets of $D^* \to D^0\pi$ decays, which can be selected cleanly, without PID information. Correlations are preserved between PID variables for the same track, such as those concerning the identification of a π or a K on the same track.

3.7.2 Particle Identification (Muons)

A correction is applied to the PID of the reconstructed muons due to differences between the PID selection efficiencies in data and simulation in the stripping process. These differences stem from an imperfect modelling of the muon system response and the efficiency of muon identification algorithms in simulation. Effects such as limitations in the simulation of multiple scattering and energy loss in the muon stations can contribute to these discrepancies. For instance, the efficiency of correctly identifying a muon in data may be slightly lower than in simulation due to additional detector noise or variations in the response of individual muon chambers.

The weights are calculated, to adjust the muon PID variables, using the PIDCalib package [175]. The weights are determined in bins of p_{T} and η, as the ratio of the efficiencies of the data and simulation samples. Samples of $B^+ \to J/\psi K^+$ decay candidates are used, with the subsequent $J/\psi \to \mu^+\mu^-$ decay. The decays are selected using a tag-and-probe method, whereby PID requirements are applied to one muon and the other is selected in an unbiased manner. In MC the events are reconstructed using the same stripping

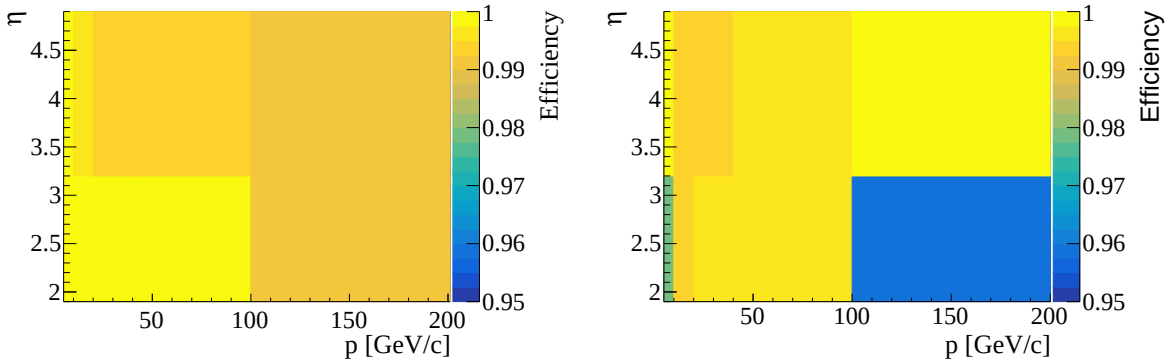


Figure 3.7.1: Track reconstruction efficiency as a function of the momentum p and pseudorapidity η of a particle. The efficiency is shown for 2011 (left) and 2018 (right) samples.

requirements as the data but without the μ PID. The corrections are produced and applied separately for each year of data taking and each polarity.

3.7.3 Track Efficiency

The tracking efficiency in simulation differs from data due to imperfections in the detector modelling, including misalignment effects, variations in detector response over time, and inaccuracies in the material description. For example, multiple scattering and energy loss in the detector material may not be perfectly simulated, leading to small differences in track reconstruction efficiency between data and simulation.

Corrections are derived using candidates from the abundant $J/\psi \to \mu^+\mu^-$ decay mode, employing a tag-and-probe approach. The tag track is identified as a muon and fully reconstructed, while the probe track is partially reconstructed, omitting data from at least one subdetector. The efficiency is determined by searching for a fully reconstructed track that corresponds to the probed track, providing the efficiency of the unused subdetectors.

The tracking efficiency is parametrised as a function of momentum, pseudorapidity, and track multiplicity, as these variables strongly influence the probability of reconstructing a given track. The use of $J/\psi \to \mu^+\mu^-$ decays allows for an unbiased measurement of tracking efficiency across these kinematic variables, as the J/ψ provides a clean and well-understood source of muons with a wide kinematic coverage. The efficiency as a function of momentum and pseudorapidity is shown in Figure 3.7.1, for the 2011 and 2018 samples. The corrections are computed using the LHCb TrackCalib package [176] and applied separately for each year of data taking and both magnetic polarities.

3.7.4 L0 hardware trigger efficiency

The response of the L0 hardware trigger is corrected using a data-driven method [177]. Samples of $B^+ \to J/\psi K^+$ decays are reconstructed, and the stripping selection is applied. An additional requirement is made to ensure that candidates identified as kaons satisfy $P(K \mid K) > 0.1$.

The L0 trigger efficiency is calculated in bins of the transverse momentum (p_{T}) of the μ^+ and μ^-, as the trigger response is a function of the muons' momentum. While the L0 trigger is applied to events based on the presence of high-momentum particles, the efficiency correction derived from $B^+ \to J/\psi K^+$ decays is assumed to be valid for the $\pi^+\pi^-\mu^+\mu^-$ decay mode. This is because the L0 trigger response primarily depends on the presence of the two muons in the event, which are common to both the signal and $B^+ \to J/\psi K^+$ decay modes. As a result, the same efficiency correction procedure can be applied to both signal and control samples.

For both the data and simulation datasets, the number of candidates in two samples are determined. The first sample contains those events which are triggered by non-signal particles in the event (TIS), independently of the specific signal particles. The number of events in this sample is denoted by N_{TIS}. The second sample contains the subset of events from the first sample where the signal also triggered the event, with the sample size denoted by $N_{\mathrm{TISandTOS}}$. The efficiency, determined separately for data and simulation samples, is given by

$$\epsilon = \frac{N_{\mathrm{TIS\ and\ TOS}}}{N_{\mathrm{TIS}}}. \tag{3.8}$$

The efficiency of the real data is determined by fitting the invariant mass of candidates in both samples separately, to obtain the respective yields. The correction weights are then calculated as the ratio of the data to MC trigger efficiencies, determined separately for each year of data taking.

3.7.5 Kinematics of the B^+ meson

The kinematics of the B-hadrons are known to be described incorrectly in simulation data. To correct for this, three variables, the p_{T} and η of the parent meson, and event multiplicity are used to produce a multi-dimensional weight. The event multiplicity is included because it correlates with the interaction dynamics and can impact the B^0 kinematics, particularly its p_{T} and η distributions. It is typically measured as either the number of hits in the scintillating pad detector or the number of tracks in the event.

Samples of LHCb and MC simulated data for the $B^+ \to J/\psi K^+$ decay mode are used to derive the kinematic weights due to the clean signal peak found in the LHCb dataset.

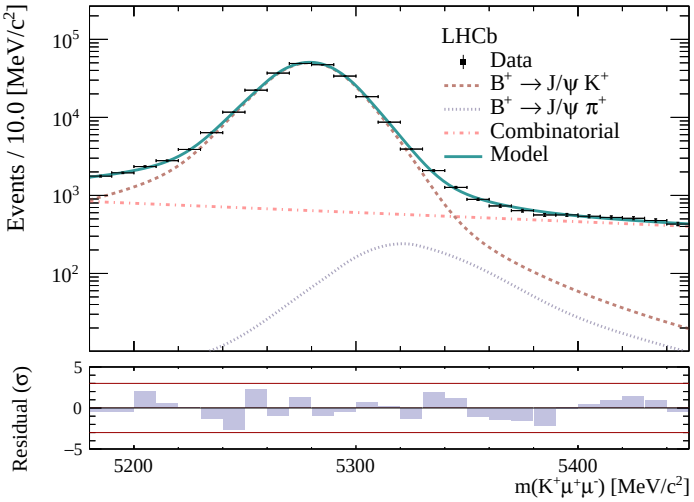


Figure 3.7.2: Fitting of the invariant mass fits of the $B^+ \to J/\psi K^+$ decay mode. The fit is performed on the 2016, magnet polarity down LHCb dataset.

Due to isospin symmetry there should be no difference in B^+/B^0 collisions at the LHC.

The selection requirements placed on the $B^+ \to J/\psi K^+$ candidates, for the pre-selection and trigger, are the same as those applied $B^0 \to \pi^+\pi^-\mu^+\mu^-$ events. The selection of $P(h \mid h) > 0.44$ is applied to the kaon, with the opposite definition to Equation 3.2, selecting those events with a probability of being a kaon and not being a pion. The MC corrections which have thus far been discussed are also applied to the data.

The invariant mass distribution of the B^+, with the range $[5180, 5450]$, is fitted in data in order to separate the signal from the background candidates. The decay modes present in the sample are the $B^+ \to J/\psi K^+$ signal, the $B^+ \to J/\psi \pi^+$ (where a π^+ is misidentified as a K^+), and combinatorial background. The shapes for the two decay modes are determined using MC samples which are both modelled using the sum of an asymmetric double-sided *Crystal Ball* [178], a Gaussian function with power-law tails, and a *Gaussian*. The mean of the signal shape and a scale on the widths of both the signal and $B^+ \to J/\psi \pi^+$ are floated in the fit to data, with the rest of the parameters fixed to the values obtained from the simulation modelling. The combinatorial background is modelled in data using an exponential. The mass shapes are detailed in Section 4.1.1. The *sPlot* method [179] is used to statistically separate the $B^+ \to J/\psi K^+$ from the background components [180]. An example fit is shown in Figure 3.7.2 for the 2016, magnet polarity down data sample.

Gradient Boost Reweighter

The corrections weights for the variables are determined in three dimensions. To account

for the correlations between these dimensions, a BDT algorithm called Gradient Boost Reweighter (GBR) [181, 182] is employed. GBR calculates weights that align an input with a target distribution.

The signal data and simulation samples are divided into training and testing datasets. The training is performed using k-fold cross validation, similarly to the use in Section 3.5.1. The weights for each event are determined using the trained model from the fold in which that event was not part of the training set. The use of multiple folds ensures the model is robust to unseen samples and reduces the risk of overtraining.

Agreement between data and MC samples of $B^+ \to J/\psi K^+$ candidates

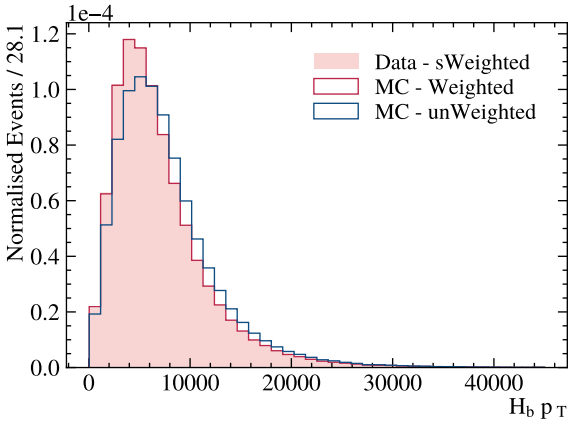
The kinematic corrections are applied to the simulated $B^+ \to J/\psi K^+$ data with the intention of bringing the MC and data samples into agreement in multiple dimensions. Figure 3.7.3 shows, for each of the three variables, the MC samples prior to and post training alongside the corresponding LHCb data.

3.7.6 Decay model

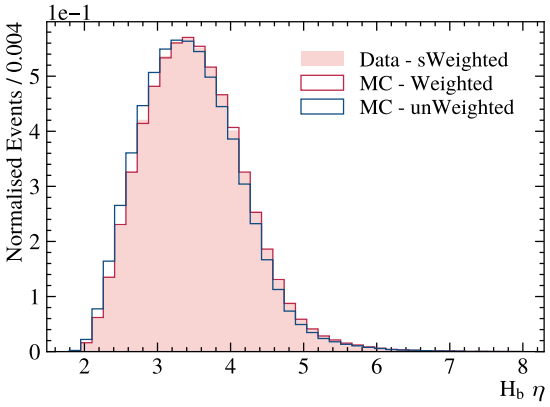
The $B^0 \to J/\psi \pi^+\pi^-$ control mode is used to validate the agreement between simulation and data. However, in the simulation, this decay is generated as a pure $B^0 \to J/\psi \rho^0$ decay, meaning only the dominant ρ^0 resonance is present. In reality, the $\pi^+\pi^-$ invariant mass distribution includes additional resonant structures beyond the dominant ρ^0 contribution, such as other $\pi^+\pi^-$ resonances. If uncorrected, this mismatch in the $\pi^+\pi^-$ spectrum can lead to misleading conclusions when comparing simulation to data.

To address this, decay model corrections are applied based on measured amplitude distributions. The structure of the $\pi^+\pi^-$ mass spectrum in $B^0 \to J/\psi \pi^+\pi^-$ decays has been well studied [53], allowing the derivation of weights that transform the simulated $B^0 \to J/\psi \rho^0$ spectrum to match the experimental mass distribution. The impact of these corrections is illustrated in Figure 3.7.4, where the weighted simulation aligns with the expected resonance contributions. The corrected model corresponds to the equivalent component in the complete $m(\pi^+\pi^-)$ spectrum.

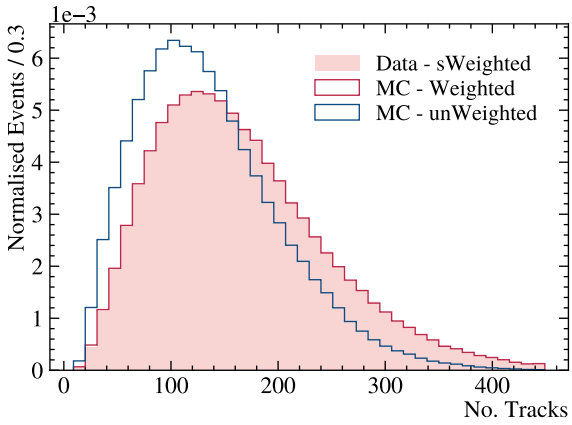
These weights are not applied to the rare decay $B^0 \to \pi^+\pi^-\mu^+\mu^-$ because no equivalent amplitude model exists for this decay. While the $B^0 \to J/\psi \pi^+\pi^-$ mode has been studied in detail, the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay contains a mix of resonant and non-resonant contributions that are not well understood. However, this does not affect the validity of the approach, as these weights are used exclusively to check the agreement between simulation and data in the control mode and are not used for modelling the rare decay.



(a) B^+ p_{T}



(b) B^+ η



(c) Occupancy (Number of tracks)

Figure 3.7.3: Data/MC agreement for the $B^+ \to J/\psi K^+$ samples in (a) B^+ p_{T}, (b) B^+ η, and (c) track multiplicity. All figures are for samples configured from the 2018, magnetic polarity down, dataset.

3.7.7 Luminosity

The simulated MC samples used in this analysis have been produced separately for each year of data taking. The samples are generated according to the changing run conditions throughout this period. However, the sizes of the samples are not coordinated for consistency between the sizes of the datasets taken each year. A correction is applied to scale the relative sizes of the simulations samples to concur with the luminosity, $\mathcal{L}$, for each year. The yearly luminosities are summarised in Table 3.1. Weights per year and magnetic polarity, $w_{y,p}$, are calculated according to the ratio of the number of B^0 mesons between the yearly samples of LHCb data samples:

$$w_{y,p} = \frac{\mathcal{L}_y \times R_B(s)}{N_{y,p}}, \tag{3.9}$$

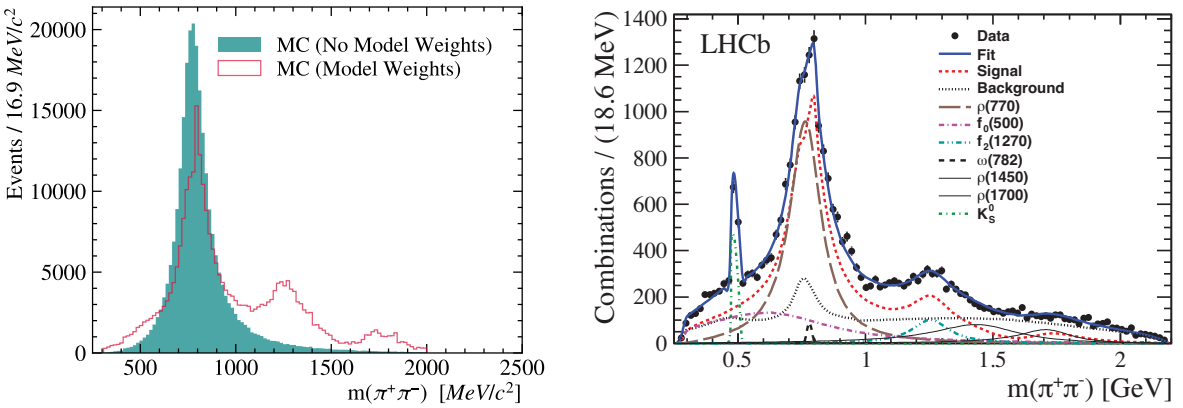


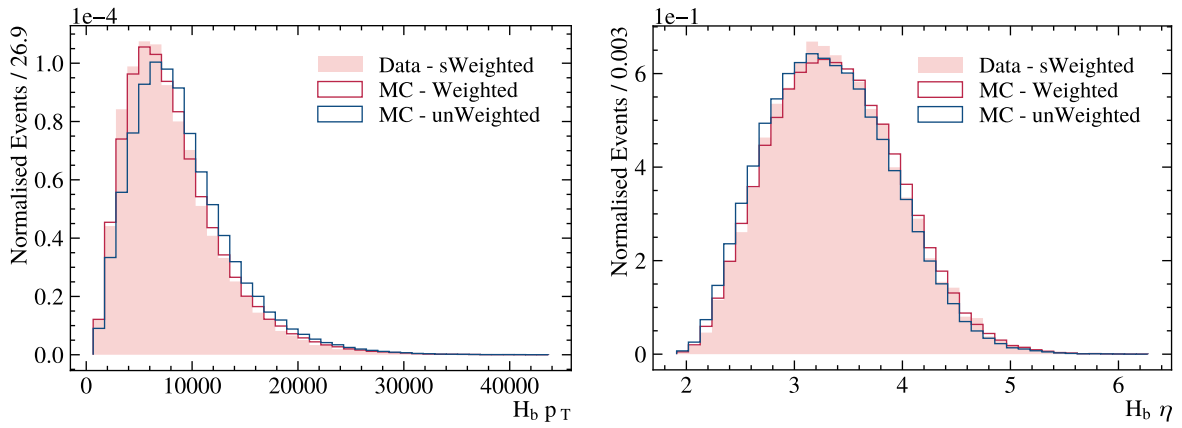
Figure 3.7.4: The figure on the left shows the true invariant mass distribution of the dipion system in MC simulated samples. The samples are generated for the $B^0 \to J/\psi\rho^0$ decay mode, the sample shown is for 2017, magnet polarity up. The spectrum is shown with and without weights correcting the decay mode. The second figure is the same distribution from the amplitude analysis from which the weights were derived [53], the ρ^0 resonance is shown with a dashed red.

where $N_{y,p}$ is the number of generated candidates for each year y and polarity p, taken directly from the generated samples. The production of B^0 mesons depends on the luminosity and the centre-of-mass energy, $\sqrt{s}$. The term $R_B(s)$ represents the ratio of the B^0 meson production probability at a given $\sqrt{s}$, relative to the lowest energy of 7 TeV.

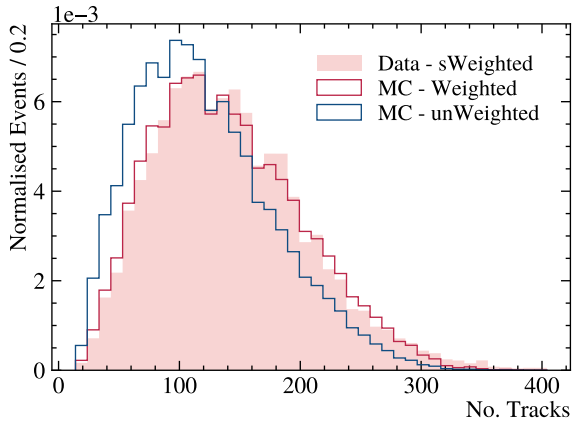
The corrections account for the differences in the number of events generated across each of the simulation samples. Therefore, the statistics of the combined sample reflect the conditions of data taking. This weighting approach ensures the relative proportions of data-taking periods and polarities in simulation match those in data, without relying on production cross-sections.

3.7.8 Comparison of MC samples and LHCb data

After calculating the weights which are produced to align the MC simulation samples with the LHCb datasets, the alignment is assessed by comparing one-dimensional distributions of variables relevant to the weighting procedure of the kinematic variables, discussed in Section 3.7.5. The MC samples, both pre- and post-weighting, are compared to signal candidates from the $B^0 \to J/\psi\pi^+\pi^-$ control mode dataset. The signal candidates are separated from the background by fitting the invariant mass distribution and employing the *sPlot* technique [179], as discussed in Section 4.1. The one-dimensional comparisons are shown in Figure 3.7.5.



(a) B^0 p_{T} (b) B^0 η



(c) Occupancy (Number of tracks)

Figure 3.7.5: Comparison of LHCb data (rose), with background removed, to MC simulation samples with (red) and without (blue) corrections. All samples correspond to the $B^0 \to J/\psi\pi^+\pi^-$ decay mode. The datasets are the combined Run I and Run II samples.

3.7.9 Effect of weighting on angular distributions

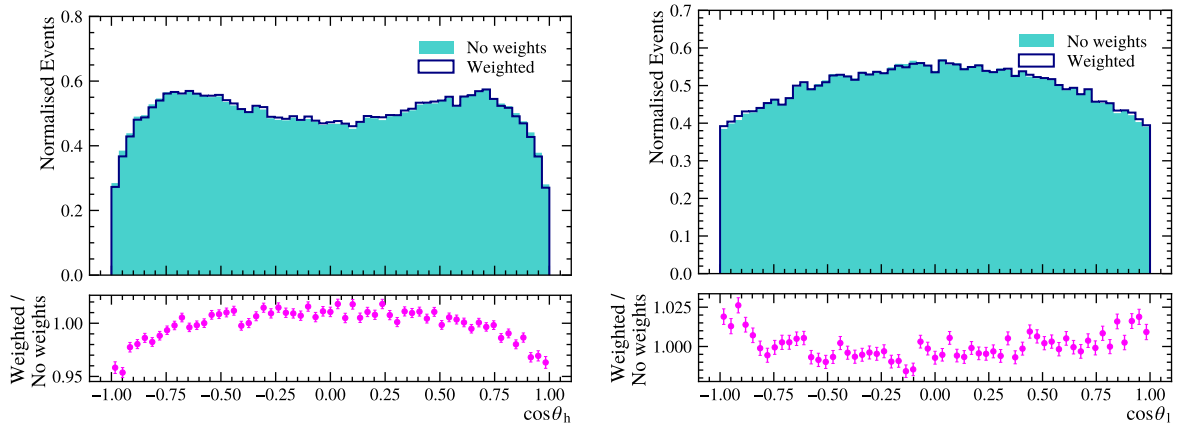
The weighting procedures are intended to align the MC simulated sampled to the LHCb data, as seen in the previous section this is achieved for several variables of interest. It is also important to review the impact of the corrections on the angular distributions, $\cos\theta_h$, $\cos\theta_l$, and ϕ, which will be used to parameterise efficiency and test fitting processes.

The rare $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC samples are weighted similarly to the J/ψ resonant sample, with the exception of the model weighting. Therefore, no adjustments are made to the weighting when parameterising the acceptance. The effects of the corrections on the angular distributions for the rare $B^0 \to \pi^+\pi^-\mu^+\mu^-$ simulation samples are given in Figure 3.7.6. The p-values obtained from the Kolmogorov–Smirnov test [183], shown in Table 3.10, are all much greater than 0.05, meaning we cannot reject the null hypothesis

that the distributions are the same. Consequently, the corrections have minimal impact on the overall shape of the angular distributions, and the corrected samples are suitable for determining the effects of the selection on the angular distributions.

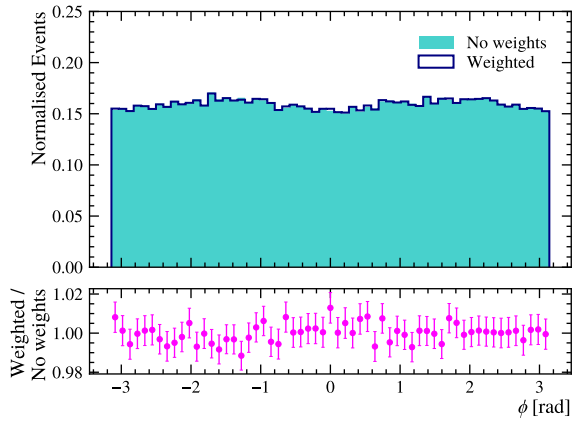
Distribution	p-value
$\cos\theta_h$	0.46
$\cos\theta_l$	0.97
ϕ	0.96
$m(\pi^+\pi^-)$	0.85

Table 3.10: Results from Kolmogorov-Smirnov tests [183] comparing angular distributions in the simulation data from the combined Run I and Run II datasets. A p-value greater than 0.05 indicates that we cannot reject the null hypothesis, suggesting that the distributions are statistically consistent. The K-S tests were performed using a seed value of "2021".

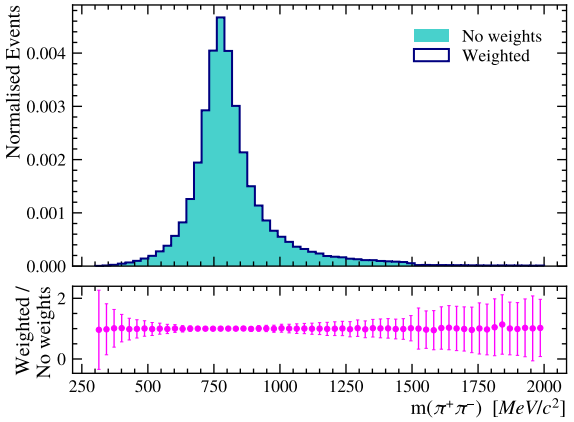


(a) $\cos\theta_h$

(b) $\cos\theta_l$



(c) ϕ



(d) $m(\pi^+\pi^-)$

Figure 3.7.6: Distributions of the three angles, $\cos\theta_h$, $\cos\theta_l$, and ϕ, and the dipion mass. The figures show the combined simulation samples for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, for all data taking years in Run I and Run II. With no correction weights applied and with the product of all of the corrections applied to the sample. The corrected and uncorrected samples align in the angular distributions.

Chapter 4

Angular analysis of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay

This section presents the methodology employed to model the data and extract the angular observables, introduced in Section 1.2.3. It will first address the fitting of the invariant mass distribution, followed by the approach to modelling the angular distributions, including an explanation of how the efficiency of the detector acceptance, coverage, and selection requirements on the data affect the shape of the angular distributions are accounted for. In Chapter 5, systematic uncertainties will be discussed and evaluated.

The invariant mass distribution of the b-hadron is modelled to distinguish between candidates from the signal and the background decay modes which are not vetoed by the selection requirements. An unbinned maximum likelihood fit is performed to statistically separate the signal and background decay modes. The shape of each of the components, with the exception of the combinatorial background, is determined using dedicated simulation samples, and subsequently used in the modelling of LHCb data samples. The $B^0 \to J/\psi\pi^+\pi^-$ and $B^0 \to \psi(2S)\pi^+\pi^-$ decay modes are used as control modes, where the $B^0 \to J/\psi\pi^+\pi^-$ decay was used in Section 3.7.8 to determine agreement between data and simulation samples. Both decay modes are used to validate the modelling strategy and the observables are extracted in Section 4.5.2. The mass fit is conducted separately for the $B^0 \to J/\psi\pi^+\pi^-$, $B^0 \to \psi(2S)\pi^+\pi^-$, and the rare $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay modes. The fits to each sample are detailed in Section 4.1.

After modelling the mass distribution, the sPlot technique [179] is employed to isolate the signal decay mode from the background in the dataset. The technique uses the invariant mass to project out the distribution of the three angular dimensions, $\vec{\Omega} = (\cos\theta_h, \cos\theta_l, \phi)$.

In Section 1.2.3, the differential decay rate was introduced in Equation 1.51, parametrised as a function of the observables to be measured. The $9\,\text{fb}^{-1}$ LHCb dataset includes a limited number of signal candidates, which precludes a direct three-dimensional fit to measure the angular observables. Therefore, the function is simplified by integrating

the decay rate to a two-dimensional function in $\cos\theta_h$ and $\cos\theta_l$, and a one-dimensional function in ϕ, improving the stability of the fit. This integration removes the ability to measure some of the angular observables, however the fit is stable and four observables, F_L, S_3, A_6, and A_9, can be measured. This simplification and its consequences are described in Section 4.2.1.

The impact of selection efficiency on the angular distributions is examined in Section 4.3, with a discussion on how this efficiency is incorporated into the measurement. The validation of the angular modelling using pseudodatasets, in Sections 4.4.3 to 4.4.4, is followed by a description of the modelling of simulation samples, in Section 4.4. The measurement of the observables on the J/ψ and $\psi(2S)$ data samples, and the blinded measurements of the data in the rare q^2 bins, as detailed in Sections 4.5.1 to 4.5.3. In Chapter 5, the systematic uncertainties will be evaluated.

4.1 Invariant Mass

Following the application of requirements designed to select $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays in the LHCb dataset there are other decay modes remaining as backgrounds. To separate the signal from the background an unbinned maximum likelihood fit is performed to the mass of the four body final state, $m(\pi^+\pi^-\mu^+\mu^-)$.

The fit is performed by maximising the extended likelihood function ($\mathcal{L}$), which represents the probability of observing both the number of events and their distribution in the dataset, given a set of parameters. The likelihood function is a combination of the Poisson probability of observing the total number of events and the product of probability density functions (PDFs) for each event in the dataset. Here, an unbinned extended fit is performed, with likelihood,

$$\mathcal{L} = e^{-\mu(\theta)} \prod_{i=1}^{N} f(m_i;\theta), \tag{4.1}$$

where N is the number of observed events $\mu(\theta)$ is the expected number of events based on the model parameters θ, and $f(m_i;\theta)$ is the PDF that describes the distribution of the variable m (here, the invariant mass) for each event i.

The fit maximises $\mathcal{L}$, thus determining the best values of the set of parameters in θ. In practice, the negative log-likelihood (NLL) is used for computational efficiency and numerical stability. For an extended likelihood fit, the NLL is given by:

$$-\ln\mathcal{L} = \mu(\theta) - \sum_{i=1}^{N} \ln f(m_i;\theta). \tag{4.2}$$

Here, $\mu(\theta)$ accounts for the expected total number of events, and the second term is the sum of the log-probabilities for each event.

The total PDF, in the fit to data, is the sum of the contributions from different components in the dataset, weighted by the number of candidates associated to each component. The generalised form of this total PDF is:

$$\mathcal{P}_{\text{total}}(m;\theta) = \frac{1}{N_{\text{total}}} \sum_k N_k \cdot F_k(m;\theta_k), \tag{4.3}$$

where N_k is the number of events for component k, $F_k(m;\theta_k)$ is the PDF, and θ_k the set of parameters, for the same component. The total number of events is denoted by $N_{\text{total}} = \sum_k N_k$.

The mass distributions of four data samples are modelled. The fit is performed separately in the J/ψ and $\psi(2S)$ q^2 bins, and in the two rare q^2 bins. To approximate the shape of the individual signal and background components in the samples, PDFs are fitted to dedicated simulation samples. These samples are subject to selection requirements, and are corrected to align with the LHCb data.

4.1.1 The $B^0 \to J/\psi\pi^+\pi^-$ mass distribution

The modelling of the $B^0 \to J/\psi\pi^+\pi^-$ decay mode, in the J/ψ-q^2 bin, is detailed in this section. The model is a sum of models for each of the components in the dataset. In addition to the signal component, there are a total of five background components in the total PDF, given by

$$\begin{aligned}\mathcal{P}_{\text{total}}(m) = \frac{1}{N_{\text{total}}} \Bigg[& N_{B^0\to J/\psi\pi^+\pi^-} \cdot \mathcal{P}_{B^0\to J/\psi\pi^+\pi^-}(m) + N_{B_s^0\to J/\psi\pi^+\pi^-} \cdot \mathcal{P}_{B_s^0\to J/\psi\pi^+\pi^-}(m) \\ & + N_{B^0\to J/\psi K^{*0}} \cdot \mathcal{P}_{B^0\to J/\psi K^{*0}}(m) + N_{\text{Part. Reco.}} \cdot \mathcal{P}_{\text{Part. Reco.}}(m) \\ & + N_{\text{Combinatorial}} \cdot \mathcal{P}_{\text{Combinatorial}}(m) \Bigg].\end{aligned} \tag{4.4}$$

The parameters N_i are the number of candidates for each component i, and N_{total} is the total number of modelled candidates. The individual components will be discussed in the following subsections.

Unless stated otherwise, the models for the individual shapes in the fit to LHCb data are determined using fits to dedicated simulation datasets, for each component.

The mass window

The range for the invariant mass of the final state particles, $m(\pi^+\pi^-\mu^+\mu^-)$, for the modelling of the J/ψ and $\psi(2S)$ resonant decay modes and the rare mode is $m(\pi^+\pi^-\mu^+\mu^-) \in [5150, 5650]\,\text{MeV}/c^2$. As mentioned in Section 3.6.2, the lower bound of this range suppresses the contribution of the partially reconstructed backgrounds. The range of the mass of the dihadron system, as discussed in Section 3.4, is $m(\pi^+\pi^-) \in [620, 920]\,\text{MeV}/c^2$. This range predominantly selects the ρ^0 (770)$\rightarrow\pi^+\pi^-$ resonance.

$B^0 \rightarrow J/\psi\pi^+\pi^-$* and *$B^0_s \rightarrow J/\psi\pi^+\pi^-$

The invariant mass distributions of the signal decay mode, $B^0 \rightarrow J/\psi\pi^+\pi^-$, and corresponding B^0_s mode, are fit with the same shape. The two decay modes share similar kinematics and are separated by the known mass difference, between the B^0 and B^0_s mesons, of $87.0 \pm 0.6\,\text{MeV}/c^2$ [22]. The model used to fit the shapes is sum of an asymmetric double-sided Crystal Ball function [178] (DCB) and a Gaussian function (G),

$$\mathcal{P}(m) = f_{\text{fit}} \cdot DCB(m; \mu, \sigma_L, \sigma_R, \alpha_L, \alpha_R, n_L, n_R) + (1 - f_{\text{fit}}) \cdot G(m; \mu, \sigma_G), \tag{4.5}$$

where f_{fit} is the fit fraction that controls the relative contribution of each component to the PDF. The DCB is described by:

$$DCB(m; \mu, \sigma_L, \sigma_R, \alpha_L, \alpha_R, n_L, n_R) = \frac{1}{\mathcal{N}} \begin{cases} A_L \left(B_L - \frac{m-\mu}{\sigma_L} \right)^{-n_L}, & \text{if } \frac{m-\mu}{\sigma_L} \leq -\alpha_L, \\ e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma_L}\right)^2}, & \text{if } -\alpha_L < \frac{m-\mu}{\sigma_L} \leq 0, \\ e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma_R}\right)^2}, & \text{if} 0 < \frac{m-\mu}{\sigma_R} \leq \alpha_R, \\ A_R \left(B_R + \frac{m-\mu}{\sigma_R} \right)^{-n_R}, & \text{if } \frac{m-\mu}{\sigma_R} > \alpha_R, \end{cases} \tag{4.6}$$

where μ is the mean of the distribution, σ_L and σ_R are the widths of the left and right sides of the distribution, respectively. The parameters α_L and α_R determine the points where the distribution transitions from the core to the power-law tails. The parameters n_L and n_R control the steepness of the left and right power-law tails, respectively, $\mathcal{N}$ is the total integral over m, it ensures the distribution is normalised to unity.

The normalisation constants A_L, and A_R, B_L, and B_R are given by:

$$A_i = \left(\frac{n_i}{|\,\alpha_i\,|} \right)^{n_i} \cdot \exp\left(-\frac{|\,\alpha_i^2\,|}{2} \right) \quad \text{and} \quad B_i = \frac{n_i}{|\,\alpha_i\,|} - |\,\alpha_i\,|\,. \tag{4.7}$$

where i represents either L or R for the left and right tails, respectively. The shape has a Gaussian core and power-law tails, which have independent parameters. The power-

law tails makes the model effective in capturing non-Gaussian characteristics resulting from detector effects or energy loss. The shape is commonly used to modelling mass distributions in b-hadron decays.

The additional Gaussian component is included in the model to account for detector resolution effects, which introduce a broadening of the peak. The Gaussian is described by a standard form:

$$G(m;\mu_G,\sigma_G) = \frac{1}{\sigma_G\sqrt{2\pi}} \exp\left(-\frac{(m-\mu_G)^2}{2\sigma_G^2}\right), \tag{4.8}$$

where μ is the mean, shared with the DCB shape, and σ_G is the width of the Gaussian component. The fits to the simulated samples are shown in Figure 4.1.1.

Misidentified backgrounds

The background associated with the $B^0 \to J/\psi K^{*0}$ decay mode arises from the misidentification of a kaon as a pion, resulting in a mass shape smeared towards lower invariant mass.

The shape is modelled using the sum of a Gaussian distribution and a Johnson's S_U (JSU) distribution [184]. The JSU distribution is obtained by transforming a normally distributed variable, m, by

$$z = \gamma + \delta \sinh^{-1}\left(\frac{m-\xi}{\lambda}\right), \tag{4.9}$$

where ξ is the peak position in m, λ is the scale parameter which stretches or compresses the shape, and γ and δ control the skewness and kurtosis of the distribution, influencing the tails and asymmetry. The corresponding PDF is:

$$JSU(m;\xi,\lambda,\gamma,\delta) = \frac{\delta}{\gamma\sqrt{2\pi}} \cdot \frac{1}{\sqrt{1+\left(\frac{m-\xi}{\lambda}\right)^2}} \exp\left(-\frac{1}{2}z^2\right). \tag{4.10}$$

The fit to the dedicated simulation sample is shown in Figure 4.1.1.

Partially reconstructed backgrounds

There are components in the dataset which result from partially reconstructed decays, as discussed in Section 3.6.2. In the mass fit window, two such components are considered: the $B_s^0 \to J/\psi(\phi \to \pi^+\pi^-\pi^0)$ and $B_s^0 \to J/\psi(\eta' \to \rho^0\gamma)$ decay modes, where the ρ^0 decays to $\pi^+\pi^-$. Each contribution is modelled with the sum of a Gaussian and a *Johnson's* S_U shape, similarly to the misidentified $B^0 \to J/\psi K^{*0}$ background. These shapes are determined separately using dedicated simulation samples.

The partially reconstructed backgrounds overlap significantly in the invariant mass

distribution, so in the fit to data the shapes are summed into a combined component. To help the fit converge, the fit fraction is fixed to the ratio of the product of their branching fractions and selection efficiencies (determined using the dedicated simulation samples):

$$\frac{\mathcal{B}(B_s^0 \to J/\psi\phi) \cdot \epsilon(B_s^0 \to J/\psi\phi)}{\mathcal{B}(B_s^0 \to J/\psi\eta') \cdot \epsilon(B_s^0 \to J/\psi\eta')} = 0.53. \tag{4.11}$$

This component is referred to as "Part. Reco.", in Equation 4.4. The fits to the dedicated simulation samples, for each partially reconstructed background, is shown in Figure 4.1.1.

Combinatorial background

The combinatorial background is produced by a random combination of tracks passing the selection. Section 3.5 discussed the reduction of this background using an MVA. Following this reduction, the component has an exponential shape, decreasing moving upwards in mass.

There are no simulation samples for the combinatorial background component, from which the fit parameters can be determined. By extending the fit window beyond the signal peak, a region dominated by combinatorial background events can be used to constrains the slope parameter in the fit to data. The combinatorial background is modelled using an exponential function given by:

$$E(m;\lambda) = C \cdot e^{-\lambda m}, \tag{4.12}$$

where λ is the slope parameter and C is a normalisation scale factor, here representing the number of candidates. This parameter is allowed to vary during the fit to data.

Fit to LHCb data

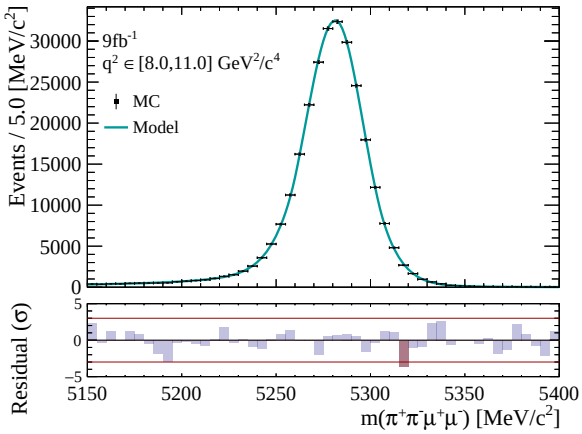
In the fit to the LHCb dataset, the total PDF contains 28 parameters, of which 20 are fixed to the values obtained from the fits to the simulation samples. The remaining parameters are allowed to vary, they include those which are not able to be predicted by these fits to the simulation datasets. They include the number of candidates for each component, an offset on the mean of each component shared across the contributions,

$$\mu_{\text{data}} = \mu_{\text{simulation}} + \mu_{\text{offset}}. \tag{4.13}$$

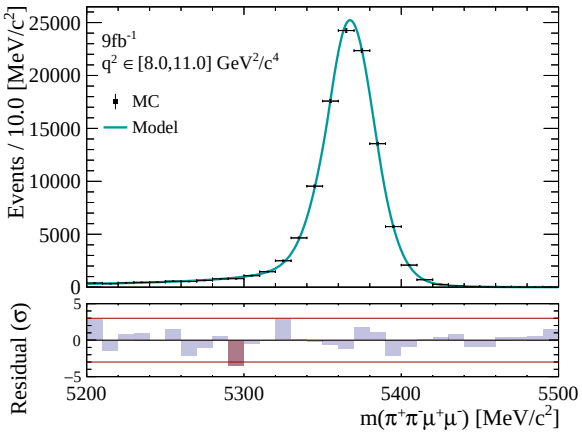
and a scaling factor on the widths,

$$\sigma_{\text{data}} = \sigma_{\text{simulation}} \times \sigma_{\text{scale}}. \tag{4.14}$$

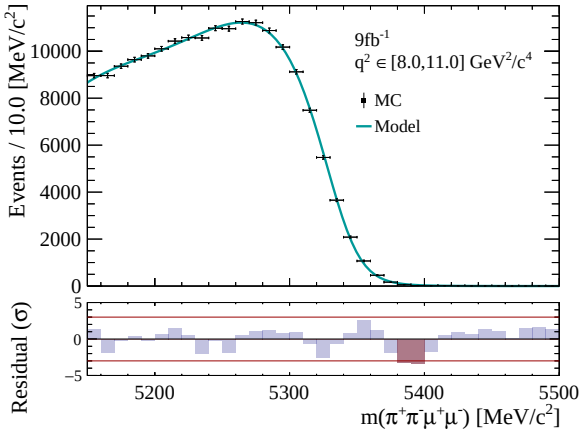
The width scale accounts for differences in the resolution between simulation and data samples. Table 4.1 summarises the values of the parameters allowed to vary, returned



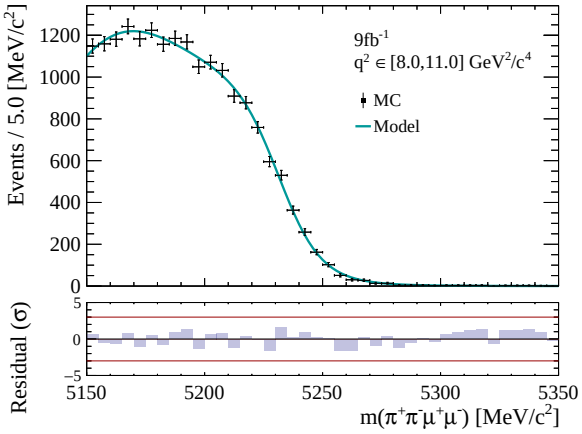
(a) $B^0 \to J/\psi\rho^0$



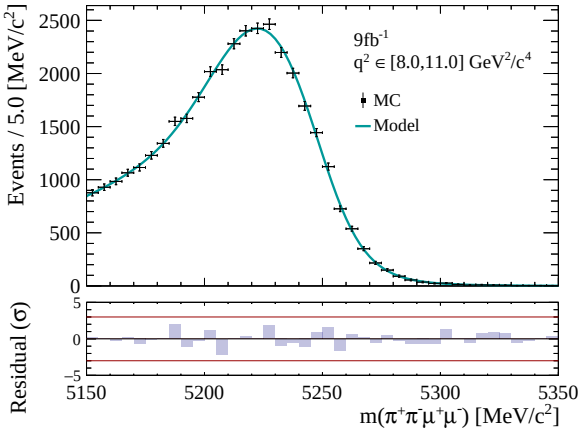
(b) $B_s^0 \to J/\psi\rho^0$



(c) $B_s^0 \to J/\psi\eta'$

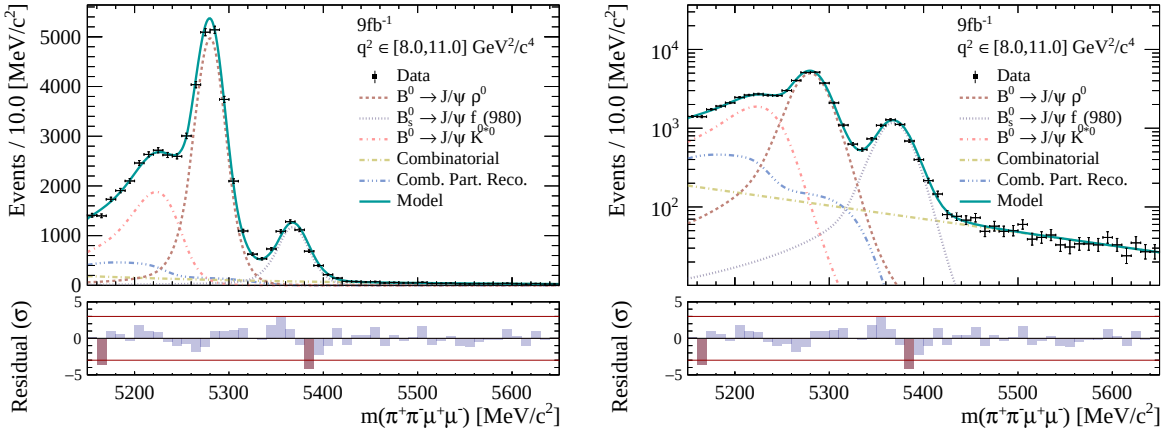


(d) $B_s^0 \to J/\psi\phi$



(e) $B^0 \to J/\psi K^{*0}$

Figure 4.1.1: Invariant mass fits to the simulated samples of different decay modes reconstructed as $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays, in the J/ψ-q^2 bin. The samples are combined Run I and Run II simulation datasets.



(a) Linear (b) Log

Figure 4.1.2: Invariant mass fits to the $B^0 \to J/\psi\pi^+\pi^-$ decay mode in the combined Run I and Run II LHCb datasets. The figure is shown in linear and logarithmic scales.

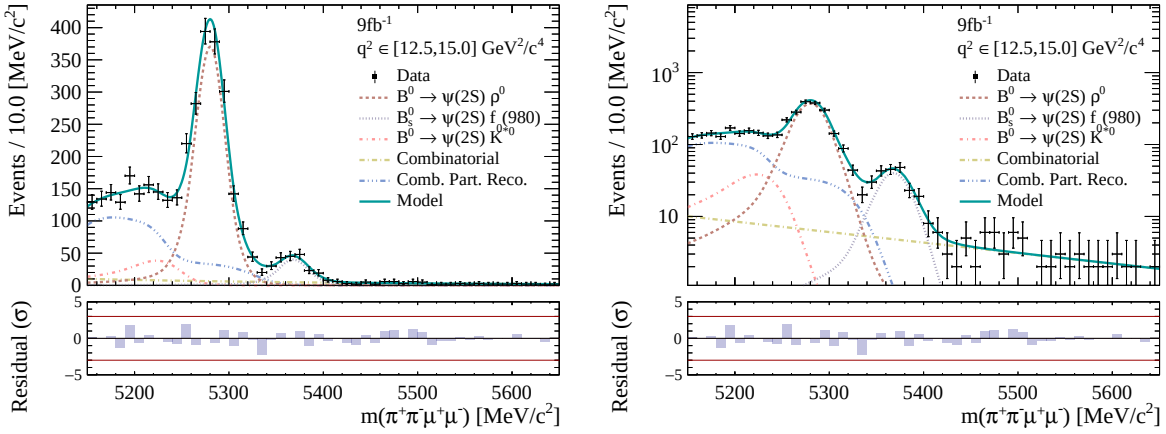
from the fit to LHCb data. Figure 4.1.2 shows the invariant mass fit to the $9\,\text{fb}^{-1}$ dataset, in both linear and logarithmic scales.

Parameter	Value
μ_{offset}	0.64 ± 0.18
σ_{scale}	1.13 ± 0.01
$N_{B^0 \to J/\psi\pi^+\pi^-}$	23819 ± 206
$N_{B^0_s \to J/\psi\pi^+\pi^-}$	5857 ± 102
$N_{B^0 \to J/\psi K^{*0}}$	14785 ± 370
$N_{\text{Part. Reco.}}$	5082 ± 519
λ	$(-3.9 \pm 0.5) \times 10^{-3}$ MeV$^{-1}/c^{-2}$
$N_{\text{Combinatorial}}$	4103 ± 392

Table 4.1: Values of parameters obtained from the mass fit of the $B^0 \to J/\psi\pi^+\pi^-$ decay mode, using the combined Run I and Run II LHCb datasets in the J/ψ-q^2 region.

4.1.2 The $B^0 \to \psi(2S)\pi^+\pi^-$ decay mode

In this section, the invariant mass fit to the $B^0 \to \psi(2S)\pi^+\pi^-$ decay mode, in the $\psi(2S)$-q^2 bin, is detailed. The $c\bar{c}$ resonance lies within the q^2 range of $[12.5, 15.0]$ GeV^2. Given that the similar decay dynamics of the two charmonium states, $\psi(2S)$ and J/ψ, the mass shapes for the $B^0 \to \psi(2S)\pi^+\pi^-$ mode are derived from the simulation samples used for the $B^0 \to J/\psi\pi^+\pi^-$ mode. The model fit to data is:



(a) Linear (b) Logarithmic

Figure 4.1.3: Invariant mass fits to the $B^0 \to \psi(2S)\pi^+\pi^-$ decay mode in the combined Run I and Run II LHCb datasets. The figure is shown in linear and logarithmic scales.

$$\begin{aligned}\mathcal{P}_{\text{total}}(m) = \frac{1}{N_{\text{total}}}\Bigg[& N_{B^0\to\psi(2S)\pi^+\pi^-}\cdot\mathcal{P}_{B^0\to\psi(2S)\pi^+\pi^-}(m) + N_{\text{Combinatorial}}\cdot\mathcal{P}_{\text{Combinatorial}}(m) \\ & + N_{B^0_s\to\psi(2S)\pi^+\pi^-}\cdot\mathcal{P}_{B^0_s\to\psi(2S)\pi^+\pi^-}(m) + N_{B^0\to\psi(2S)K^{*0}}\cdot\mathcal{P}_{B^0\to\psi(2S)K^{*0}}(m) \\ & + N_{\text{Part. Reco.}}\cdot\mathcal{P}_{\text{Part. Reco.}}(m)\Bigg].\end{aligned} \tag{4.15}$$

The fixed parameters are the same for both decay modes. The partially reconstructed backgrounds are also combined in this fit, to aid fit stability, with the same fit fraction as derived in Equation 4.11. The results of the fitting of the LHCb data is shown in Figure 4.1.3 and the parameters allowed to vary in the fit are summarised in Table 4.2.

4.1.3 The $B^0 \to \pi^+\pi^-\mu^+\mu^-$ rare decay mode

The invariant mass is modelled for the rare decay mode in both the low-q^2 and high-q^2 bins. The mass windows of the four body and dihadron systems are the same as for the $c\bar{c}$ resonant modes. The PDF fit to data is the sum of individual components discussed in the following subsections. This total PDF is defined as

Parameter	Value
μ_{offset}	0.83 ± 0.63 MeV
σ_{scale}	1.07 ± 0.04
$N_{B^0 \to \psi(2S)\pi^+\pi^-}$	1691 ± 53
$N_{B_s^0 \to \psi(2S)\pi^+\pi^-}$	191 ± 20
$N_{B^0 \to \psi(2S)K^{*0}}$	298 ± 93
$N_{\text{Part. Reco.}}$	1152 ± 126
λ	-0.003 ± 0.002 MeV$^{-1}/c^{-2}$
$N_{\text{Combinatorial}}$	242 ± 83

Table 4.2: Values of parameters obtained from the mass fit of the $B^0 \to \psi(2S)\pi^+\pi^-$ decay mode, using the combined Run I and Run II LHCb datasets in the $\psi(2S)$ q^2 region.

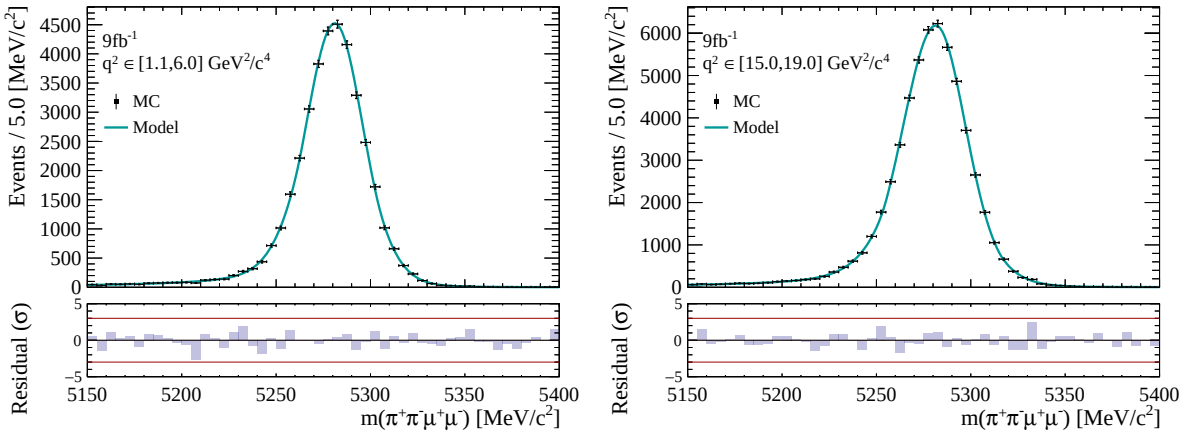
$$\mathcal{P}_{\text{total}}(m) = \frac{1}{N_{\text{total}}} \Bigg[N_{B^0 \to \pi^+\pi^-\mu^+\mu^-} \cdot \mathcal{P}_{B^0 \to \pi^+\pi^-\mu^+\mu^-}(m) + N_{\text{Combinatorial}} \cdot \mathcal{P}_{\text{Combinatorial}}(m) + N_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-} \cdot \mathcal{P}_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-}(m) + N_{B^0 \to K^{*0}\mu^+\mu^-} \cdot \mathcal{P}_{B^0 \to K^{*0}\mu^+\mu^-}(m) \Bigg]. \tag{4.16}$$

The presence of partially reconstructed backgrounds in the rare-q^2 bins was discussed in Section 3.6.2, where it is determined that, given the size of the available dataset, the relative contribution from the $B^0 \to \eta'\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ decay modes are negligible. Therefore, the shapes are not included in the fit.

The models for the individual shapes in the fit to LHCb data are determined using fits to dedicated simulation samples, with the exception of the combinatorial background component, for which there are no simulation samples available. The fits to the simulated samples are shown in Figure 4.1.4.

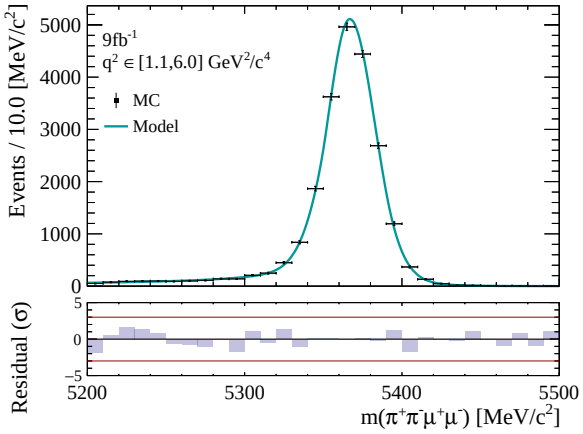
Mass shapes

The same shapes are used to model the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decay modes as for the $B^0 \to J/\psi\pi^+\pi^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ modes: a sum of an asymmetric double-sided Crystal Ball and a Gaussian. The $B^0 \to K^{*0}\mu^+\mu^-$ decay mode contains a misidentified K as a π, producing a shape smeared towards lower mass and sitting in the lower shoulder of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ distribution. This is modelled using a sum of a Gaussian and a Johnson's S_U function. The combinatorial background is modelled with an exponential function and constrained using a wide upper mass window, since no MC is used and this region is dominated by background.

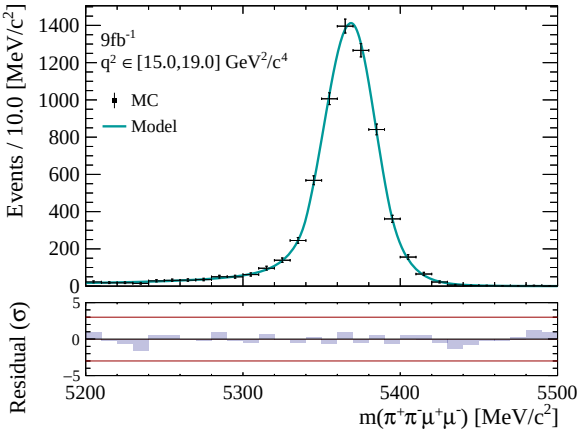


(a) $B^0 \to \pi^+\pi^-\mu^+\mu^-$ (low-q^2)

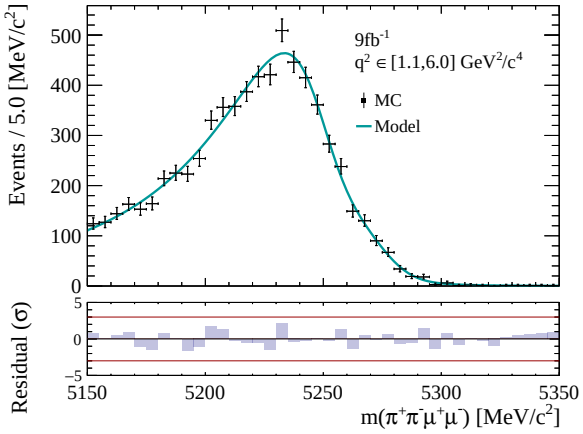
(b) $B^0 \to \pi^+\pi^-\mu^+\mu^-$ (high-q^2)



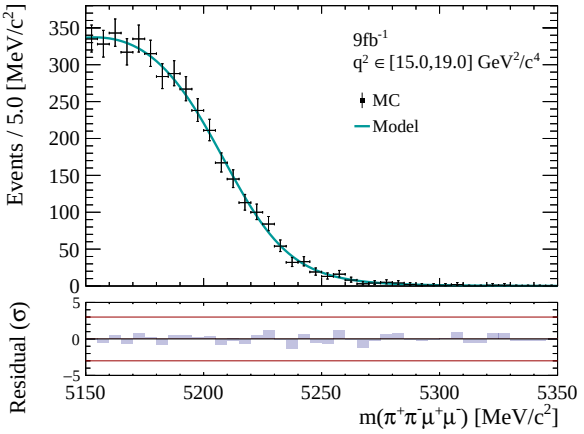
(c) $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ (low-q^2)



(d) $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ (high-q^2)



(e) $B^0 \to K^{*0}\mu^+\mu^-$ (low-q^2)



(f) $B^0 \to K^{*0}\mu^+\mu^-$ (high-q^2)

Figure 4.1.4: Invariant mass fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode in the combined Run I and Run II simulation datasets in the low-q^2 (left column) and high-q^2 (right column) bins..

Fit to LHCb data

The fit parameters which are allowed to vary in the fit to data are the number of candidates for each component and the slope parameter in the combinatorial model. A shared offset of the means of each shape and a shared scaled on their widths are included in the model; these parameters are taken from the fit to the J/ψ sample and fixed to their central values, without incorporating their associated uncertainties. The fitted parameters for the J/ψ mode are given in Table 4.1. The fit to the rare data is shown in Figure 4.1.5, in logarithmic and linear scale. The values of the fit parameters allowed to vary in the fit are given in Table 4.3.

The yields of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal can be compared to the yield obtained using the Run I, $3\,\text{fb}^{-1}$ dataset [89]. This analysis used a wider $m(\pi^+\pi^-)$ window of 500–1300 MeV/c^2, and the $m(\pi^+\pi^-\mu^+\mu^-)$ range of 5190-5990 MeV/c^2. A yield of 40 ± 10 (stat) $\pm$ 3 (syst) was measured in Run I using the whole rare q^2 range, excluding only the $c\bar{c}$ resonances in the $\sqrt{q^2}$ ranges: 1.010–1.030 GeV/c^2, 2.796–3.216 GeV/c^2, and 3.436–3.806 GeV/c^2. With the increase in luminosity between Run I and Run II and an increase of the $b\bar{b}$ production cross section between 7 and 13 TeV, the signal yields are expected to increase by approximately a factor 4 between Run I and Run II. Therefore, if the same selection requirements and fit windows were used, the yield could be expected to be approximately 200. However, since this analysis employs a different q^2 binning strategy, which excludes additional regions near the resonances and kinematic limits, the expected yield is lower. Some further refinement of the selection procedure is needed to extract a yield comparable to the naive expectation.

Parameter	Value	
	Low-q^2	High-q^2
$N_{B^0\to\pi^+\pi^-\mu^+\mu^-}$	29.5 ± 9.7	62.9 ± 9.6
$N_{B^0_s\to\pi^+\pi^-\mu^+\mu^-}$	13.7 ± 5.7	8.1 ± 4.1
$N_{B^0\to K^{*0}\mu^+\mu^-}$	60.7 ± 15.5	38.4 ± 8.3
λ	$(-5.3 \pm 1.7) \times 10^{-3}$ MeV$^{-1}/c^{-2}$	$(-1.3 \pm 5.0) \times 10^{-3}$ MeV$^{-1}/c^{-2}$
$N_{\text{Combinatorial}}$	53.1 ± 19.5	13.6 ± 11.8

Table 4.3: Values of parameters obtained from the mass fit of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, using the combined Run I and Run II LHCb datasets in the low-q^2 and high-q^2 bins.

4.1.4 Closure Tests

To assess potential biases in the mass fitting procedure, closure tests are performed for each of the mass fits presented in this chapter. These tests verify that the fit model can reliably recover input parameters when applied to pseudo-datasets generated from the

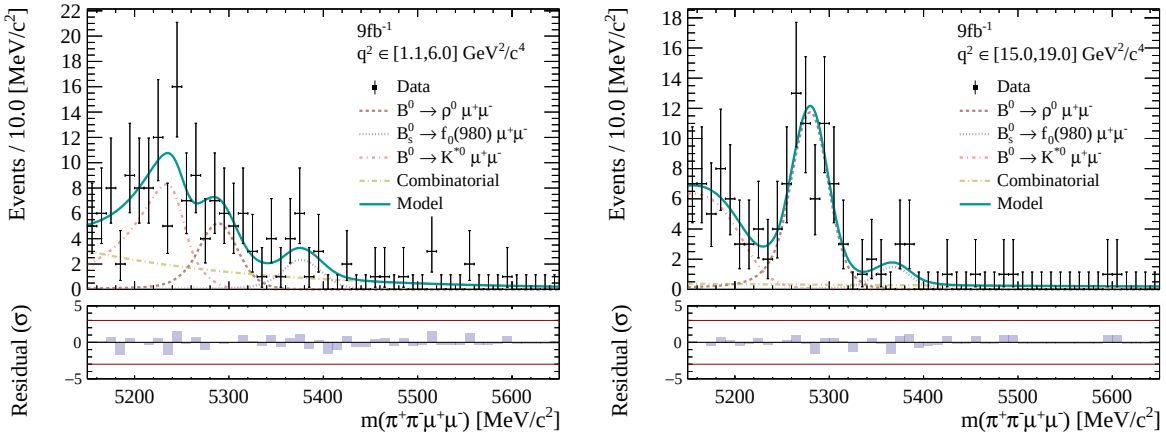


Figure 4.1.5: Invariant mass fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode in the combined Run I and Run II LHCb datasets in the low-q^2 (a) and high-q^2 (b) bins.

fitted model. This provides a validation of the fitting framework and an estimate of any systematic bias introduced by the methodology.

The closure tests specifically assess the stability and accuracy of the yield extraction in data. The method and results of these tests are outlined below.

4.1.4.1 Methodology

For each mass fit to data, the following procedure is applied:

- The nominal fit is performed on the data, yielding a set of best-fit parameters.
- Using these parameters, 1000 pseudo-datasets are generated, each with the same statistical size as the original dataset.
- Each pseudo-dataset is refitted using the same mass model, and the fitted parameters are recorded.
- The distribution of fitted values from the pseudo-experiments is summarised using the mean and standard deviation, which are then compared to the original best-fit parameters to check for potential biases.

The level of agreement is quantified using the pull:

$$\text{Pull} = \frac{x_{\text{observed}} - x_{\text{expected}}}{\sqrt{\sigma^2_{\text{observed}} + \sigma^2_{\text{expected}}}}. \tag{4.17}$$

where θ_{expected} and σ_{expected} are the value and uncertainty obtained from the data fit, and μ_{observed} and σ_{observed} are the mean and standard deviation of the pseudo-experiment

results. Pull values close to zero indicate good agreement, while values significantly greater than one may suggest bias or misestimated uncertainties.

4.1.4.2 Results for $B^0 \to J/\psi\pi^+\pi^-$

The closure test results for the $B^0 \to J/\psi\pi^+\pi^-$ mass fit are summarised in Table 4.1. The extracted parameters from data show agreement with the closure test means, with all pull values well below 2σ. This indicates that the fit reliably recovers the true parameter values and does not introduce significant bias.

Parameter	Fit Value (Data)	$\mu \pm \sigma$ (Closure test)	Pull (σ)
μ_{offset}	0.64 ± 0.18	0.64 ± 0.17	0.00
σ_{scale}	1.13 ± 0.01	1.13 ± 0.01	0.00
$N_{B^0 \to J/\psi\pi^+\pi^-}$	23819 ± 206	23815 ± 208	0.01
$N_{B_s^0 \to J/\psi\pi^+\pi^-}$	5857 ± 102	5855 ± 108	0.01
$N_{B^0 \to J/\psi K^{*0}}$	14785 ± 370	14771 ± 375	0.03
$N_{\text{Part. Reco.}}$	5082 ± 519	5064 ± 544	0.02
λ	$(-3.9 \pm 0.5) \times 10^{-3}$	$(-3.9 \pm 0.5) \times 10^{-3}$	0.01
$N_{\text{Combinatorial}}$	4103 ± 392	4123 ± 428	0.03

Table 4.4: Closure test results for the $B^0 \to J/\psi\pi^+\pi^-$ mass fit. The table compares fitted values obtained from real data to the mean and standard deviation of results from 1000 pseudo-experiments. Uncertainties on the pseudo-experiment values correspond to the standard deviation of the fitted parameters across toys. The parameter λ is expressed in units of MeV^{-1}/c^{-2}.

4.1.4.3 Results for $B^0 \to \psi(2S)\pi^+\pi^-$

The closure test for the $B^0 \to \psi(2S)\pi^+\pi^-$ mass fit similarly demonstrates good agreement between the fitted yields and the closure test results. All deviations are within the expected statistical uncertainties, confirming the robustness of the fitting procedure. The results are presented in Table 4.5.

4.1.4.4 Results for $B^0 \to \pi^+\pi^-\mu^+\mu^-$

Table 4.6 presents the results of the closure tests for the nominal $B^0 \to \pi^+\pi^-\mu^+\mu^-$ mass fit, performed in both the low- and high-q^2 bins. In all cases, the fitted values from data are consistent with the means of the pseudo-experiments. The pull values remain well within statistical expectations, supporting the conclusion that the fitting procedure does not introduce significant biases.

Parameter	Fit Value (Data)	$\mu \pm \sigma$ (Closure test)	Pull (σ)
μ_{offset}	0.83 ± 0.63 MeV	0.84 ± 0.63	0.01
σ_{scale}	1.07 ± 0.04	1.07 ± 0.04	0.00
$N_{B^0 \to \psi(2S)\pi^+\pi^-}$	1691 ± 53	1689 ± 53	0.03
$N_{B_s^0 \to \psi(2S)\pi^+\pi^-}$	191 ± 20	189 ± 20	0.07
$N_{B^0 \to \psi(2S)K^{*0}}$	298 ± 93	302 ± 95	0.03
$N_{\text{Part. Reco.}}$	1152 ± 126	1129 ± 143	0.12
λ	-0.003 ± 0.002	-0.003 ± 0.002	0.00
$N_{\text{Combinatorial}}$	242 ± 83	262 ± 97	0.16

Table 4.5: Closure test results for the $B^0 \to \psi(2S)\pi^+\pi^-$ mass fit. The table compares fitted values obtained from real data to the mean and standard deviation of results from 1000 pseudo-experiments. Uncertainties on the pseudo-experiment values correspond to the standard deviation of the fitted parameters across toys. The parameter λ is expressed in units of MeV^{-1}/c^{-2}.

Parameter	Fit Value (Data)	$\mu \pm \sigma$ (Closure test)	Pull (σ)
	Low-q^2		
$N_{B^0 \to \pi^+\pi^-\mu^+\mu^-}$	29.5 ± 9.7	29.6 ± 7.9	0.01
$N_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-}$	13.7 ± 5.7	13.7 ± 5.6	0.00
$N_{B^0 \to K^{*0}\mu^+\mu^-}$	60.7 ± 15.5	59.7 ± 14.1	0.05
λ	-0.005 ± 0.002	-0.004 ± 0.002	0.35
$N_{\text{Combinatorial}}$	53.1 ± 19.5	54.7 ± 18.0	0.06
	High-q^2		
$N_{B^0 \to \pi^+\pi^-\mu^+\mu^-}$	62.9 ± 9.6	61.2 ± 9.9	0.12
$N_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-}$	8.1 ± 4.1	7.7 ± 3.9	0.07
$N_{B^0 \to K^{*0}\mu^+\mu^-}$	38.4 ± 8.3	31.9 ± 13.6	0.41
λ	-0.001 ± 0.005	-0.006 ± 0.006	0.64
$N_{\text{Combinatorial}}$	13.6 ± 11.8	21.9 ± 18.4	0.38

Table 4.6: Closure test results for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ mass fit in both the low- and high-q^2 bins. The table compares fitted values obtained from real data to the mean and standard deviation of results from 1000 pseudo-experiments. Uncertainties on the pseudo-experiment values correspond to the standard deviation of the fitted parameters across toys. The parameter λ is expressed in units of MeV^{-1}/c^{-2}.

4.1.5 The sPlot technique

The modelling of the angular distributions is performed following the invariant mass fit, where the signal candidates are isolated from the background using the sPlot technique [179]. The sPlot method is a tool for projecting out a signal component from a dataset containing a mixture of components. The result of the unbinned maximum likelihood fit are used to assign statistical weights (sWeights) to each candidate based on the likelihood of the candidate belonging to individual components. The weights are defined as

$$s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)} \tag{4.18}$$

where N_s is the number of components in the fit, y is the set of discriminating variables (in this case the invariant mass), f_j and f_k are the PDF of the discriminating variables for components j and k. The terms $f_j(y_e)$ and $f_k(y_e)$ are the value of the PDF for event e, with y_e the set of values for the discriminating variables for the event. Finally, $\mathbf{V}_{nj}$ is the covariance matrix, given by the second derivatives of $-\mathcal{L}$, from the fit to the mass distribution.

The weights effectively isolate the signal candidates from the background candidates, enabling further study of the angular dimensions. However, a key assumption of the sPlot technique is that the discriminating variable x (the mass) and control variables y (the angles) must be uncorrelated.

To validate the assumption that the distributions factorise, the correlations between x and y are determined. If these correlations are negligible the sWeights can be extracted. In Figure 4.1.6, two-dimensional histograms in the low- and high-q^2 bins, respectively, are presented which include the linear (Pearson) correlation coefficient between the invariant mass and the angular dimensions. Additional comparisons are presented in Appendix C, for the other bins of q^2. The observed correlations are all below 1%, indicating negligible correlation and validating the assumptions of sWeights technique.

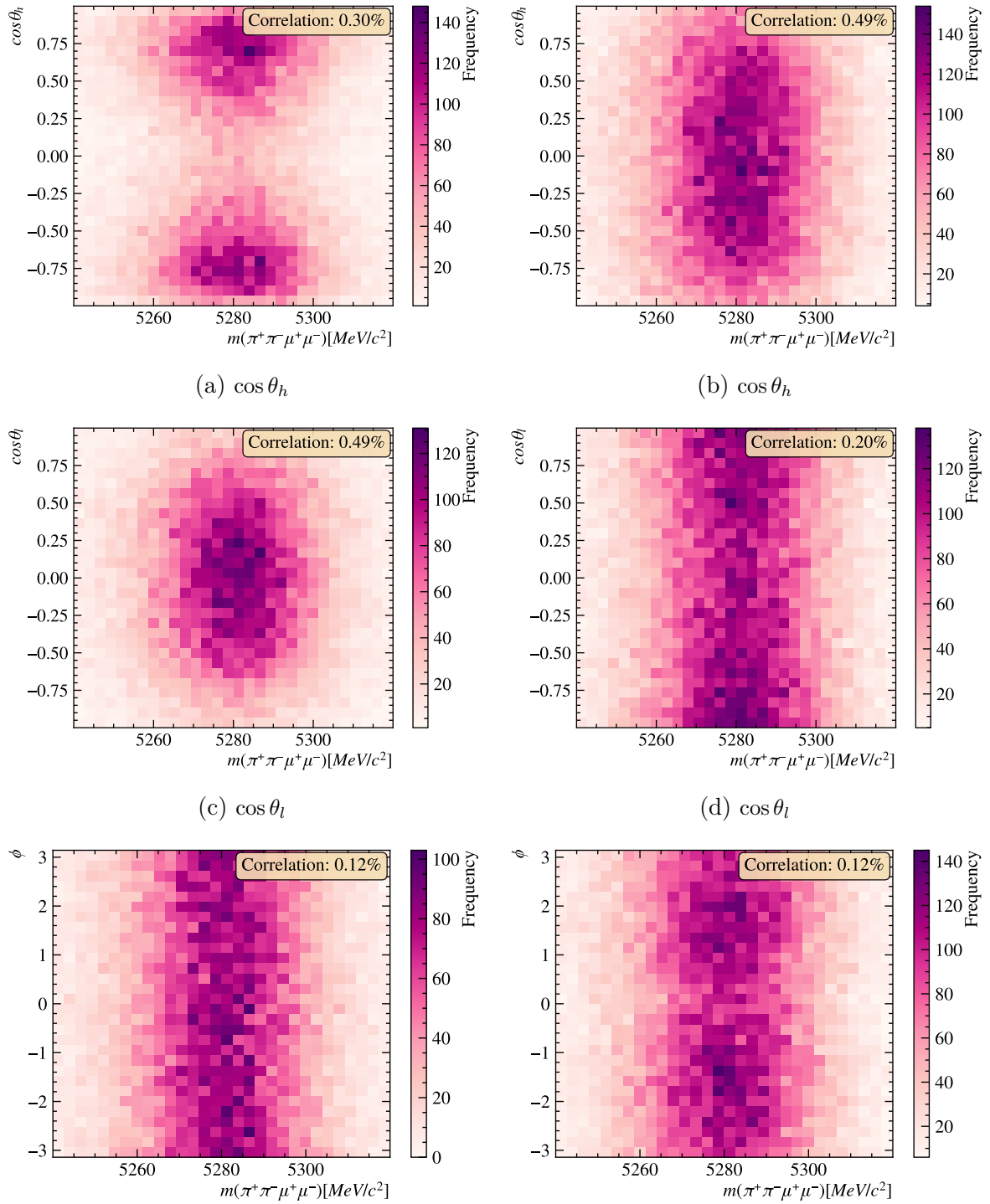
The background components are combined into a single contribution, similar to Equation 4.16, excluding the signal component and incorporating the fitted background yields from the mass fit. The combined background model is given by:

$$\begin{aligned}\mathcal{P}_{\text{background}}(x) = \frac{1}{N_{\text{background}}} \Bigg[& N_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-} \cdot F_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-}(x) \\ & + N_{B^0 \to K^{*0}\mu^+\mu^-} \cdot F_{B^0 \to K^{*0}\mu^+\mu^-}(x) + N_{\text{Combinatorial}} \cdot F_{\text{Combinatorial}}(x) \Bigg],\end{aligned} \tag{4.19}$$

where $N_{\text{background}}$ is the total number of background candidates, obtained by summing the fitted yields of individual background components. The background yields are taken as fixed values from the mass fit, where they were initially floated. To properly account for uncertainties, the covariance matrix from the mass fit is used to propagate these uncertainties into the combined background model.

Fixing the background yields reduces the number of free parameters, leading to a more stable extraction of the sWeights. However, this introduces a potential bias, which is mitigated by assigning a systematic uncertainty. A discussion of this systematic treatment

is provided in Section 5.2.0.0.5.



(a) $\cos\theta_h$ (b) $\cos\theta_h$

(c) $\cos\theta_l$ (d) $\cos\theta_l$

(e) ϕ (f) ϕ

Figure 4.1.6: Correlation between $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass and the three angular distributions in the low-q^2 (left column) high-q^2 (right column) bins. The distributions are from a simulation dataset of the $B^0 \to \rho^0\mu^+\mu^-$ decay mode, therefore the four-body invariant mass window focuses on the B^0 mass.

4.2 Angular fit

This section begins with a focus on the models used to measure the angular observables and identifying which observables can be measured using this method. Next, the distortion of the angular distributions by the reconstruction and selection is examined. Fits are subsequently performed on simulation samples, and the values of the extracted observables are compared to the expected values calculated from the amplitudes used in the generation of the samples.

Unbinned maximum likelihood fits are performed on the LHCb data. After calculating the sWeights, the fits are applied to the background-subtracted dataset. The introduction of weights modifies the likelihood function, previously defined in Equation 4.1, by incorporating a weight term, w_i, for event i. The modified likelihood function becomes:

$$\mathcal{L}_{\text{weighted}} = \prod_{i=1}^{N} f(m_i;\theta)^{w_i}. \tag{4.20}$$

Consequently, the change to the negative log-likelihood (NLL) is given by:

$$-\ln \mathcal{L}_{\text{weighted}} = -\sum_{i=1}^{N} w_i \ln f(m_i;\theta). \tag{4.21}$$

The uncertainties from this fit are derived using the Hessian matrix [185], which is the matrix of the second-order partial derivatives of the NLL, with respect to the parameters. In an unweighted fit, the Hessian provides a good approximation of the covariance matrix, C, which is the inverse of the Hessian. However, in a weighted fit, the contribution of each data point to the NLL is scaled by its weight. The covariance matrix will need to be scaled to account for the weighting.

The scaling of the covariance matrix accounts for the effective sample size, which is reduced by the presence of weights, using the weighted least squares method [186]. This is done by summing the weights, as follows:

$$C \propto \frac{\sum w_i}{\sum w_i^2}. \tag{4.22}$$

The scaling adjusts the uncertainties by reflecting the influence of the weighting, properly accounting for their contributions to the fit.

The fitting is validated using an array of pseudoexperiments. Following the validation, fits are performed on the $B^0 \to J/\psi\pi^+\pi^-$ decay mode, with the results compared to the literature. Further fits are conducted on the $B^0 \to \psi(2S)\pi^+\pi^-$ decay mode. Finally, fits to the rare q^2 bins are performed; these fits are kept blind until the analysis is reviewed within the LHCb collaboration.

4.2.1 Angular Model

In Section 1.2.3, Equation 1.52 introduces the angular function containing the observables which are to be extracted. Due to the limited number of signal candidates in the $9\,\text{fb}^{-1}$ LHCb dataset, a full three-dimensional fit of the angular observables is not feasible. Therefore, the function is simplified by integrating over the specific dimensions, reducing the complexity while capturing the essential physics. This section details this strategy for simplifying the modelling with these reduced functions.

To address the limitations of fitting directly with the full three-dimensional angular distribution the dimensionality of the function is reduced by integrating out certain variables. The three resulting one-dimensional (1D) functions which include both the P-wave and S-wave-interference terms are:

$$\begin{aligned}\frac{1}{d\Gamma/dq^2}\frac{d^2\Gamma}{dq^2 d\cos\theta_h} &= \frac{3}{4}(1-F_S)\left(1-F_L+(-1+3F_L)\cos^2\theta_h\right) \\ &\quad + \frac{1}{2}(F_S + A_{S1}\cos\theta_h) \qquad (4.23)\\ \frac{1}{d\Gamma/dq^2}\frac{d^2\Gamma}{dq^2 d\cos\theta_l} &= -\frac{3}{8}(1-F_S)\left(1+F_L+2A_6\cos\theta_l+(1-3F_L)\cos^2\theta_l\right) \\ &\quad + \frac{3}{4}F_S(1-\cos^2\theta_l) \qquad (4.24)\\ \frac{1}{d\Gamma/dq^2}\frac{d^2\Gamma}{dq^2 d\phi} &= \frac{1}{2\pi}(1-F_S)(1+S_3\cos 2\phi + A_9\sin 2\phi) \\ &\quad + \frac{1}{2\pi}F_S + \frac{3\pi}{64}(S_{S3}\cos\phi + S_{S4}\sin\phi) \qquad (4.25)\end{aligned}$$

By integrating the full three-dimensional function, some information is lost, as certain observables drop out of the model. As a result, only the observables F_L, S_3, A_6, and A_9 can be measured using this reduced framework. The S-wave fraction (F_S) remains present in each of these 1D functions. However, we need to evaluate whether these 1D fits are sufficient for extracting accurate values of the angular observables.

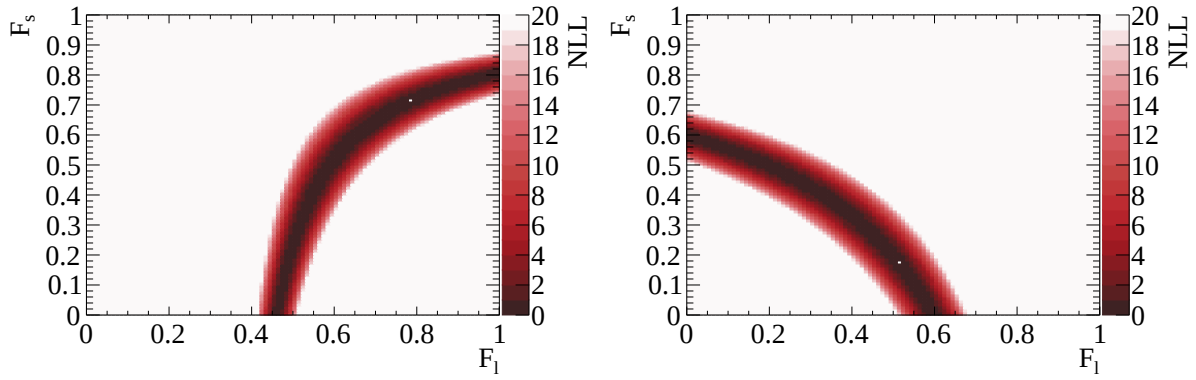
4.2.1.1 Feasibility of one-dimensional fits

The suitability of fitting the observables using different configurations is examined using pseudoexperiments. To conduct this study, samples of signal-only data are generated based on the differential decay rate from Equation 1.52.

A large pseudodataset is generated by randomly sampling events according to the differential decay rate. In these generated samples, the value of F_L is set to 0.5, and the value of F_S is 0.2, while the remaining observables are all generated with a value of 0.

Two observables are scanned over their parameter spaces in a two-dimensional (2D) grid. For each point in the scan, the NLL is computed using Minuit [187], with the observables not being scanned fixed at their generated values.

Figure 4.2.1 illustrates the differences in log-likelihood from the best-fit point for the scans of the F_L and F_S observables. The determination of NLL uses Equations 4.23 and 4.24, the 1D functions in $\cos\theta_h$ and $\cos\theta_l$, respectively. The broad bands observed in the plots indicate that multiple combinations of F_L and F_S can return similar fit results. This suggests that the true value of these observables may not be accurately determined from 1D fits alone.



(a) $\cos\theta_h$ (b) $\cos\theta_l$

Figure 4.2.1: Two-dimensional scans of the NLL in the 1D functions for $\cos\theta_h$ and $\cos\theta_l$, showing the parameter space for F_L and F_S.

4.2.1.2 2D or not 2D

The 1D fits have been shown to have limitations. Given these restrictions a scan is performed using a two-dimensional fit. As with the scan of the 1D functions, the observables not being scanned are fixed at their generated values. Integrating the three-dimensional distribution only over the ϕ leaves a two-dimensional (2D) function in $\cos\theta_h$ and $\cos\theta_l$:

$$\begin{aligned}\frac{1}{d\Gamma/dq^2}\frac{d^2\Gamma}{dq^2 d\cos\theta_h d\cos\theta_l} = \frac{9}{32}(F_S-1)\big((1+2\ A_6\ \cos\theta_l+\cos^2\theta_l)\sin\theta_h^2 \\ + F_L\left(1+\cos^2\theta_l+(-5+3\ \cos^2\theta_l)\right)\big)\cos^2\theta_h \\ -\frac{3}{8}(F_S+S_{S1}\ \cos\theta_h)\sin\theta_l^2\end{aligned} \tag{4.26}$$

The same angular observables survive this integration as in the 1D functions. The results of the 2D NLL scan, shown in Figure 4.2.2, indicate a more constrained parameter space.

This tighter constraint on the values of F_L and F_S suggest the 2D fit is more reliable in determining their true values.

The conclusion of this study is that the fit will be performed using the 2D function in $\cos\theta_h$ and $\cos\theta_l$ (Equation 4.26) and the 1D function in ϕ (Equation 4.25).

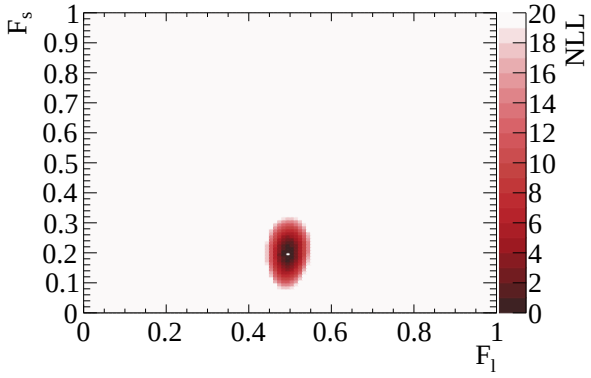


Figure 4.2.2: Two-dimensional scans of the NLL in the 2D function in $\cos\theta_h$ and $\cos\theta_l$, showing the parameter space for F_L and F_S.

4.2.2 Parameter validity regions

Fundamentally, a PDF must be positive-definite. This section examines the viability of the angular distribution of the signal remaining positive across the range of the observables. A scan is conducted where the values of two observables in the function are fixed at 101 equidistant points across their ranges. Two observables are scanned at a time, and the remaining observables, which are not being scanned, are fixed to either 0.5 for the F_L observable, or 0 for the other observables. The respective angular dimensions are then scanned within each bin. If the PDF has a negative value at any point in the scan, the bin is assigned a value of -1; otherwise, it is assigned $+1$. This scanning process is performed for each observable in the fitting models, with the results shown in Figures 4.2.3 and 4.2.4, for the 2D $\cos\theta_h$ and $\cos\theta_l$ and 1D ϕ functions respectively.

The 2D $\cos\theta_h$ and $\cos\theta_l$ function contains the F_L, A_6, and F_S observables. A 2D scan of the A_6 vs F_L parameter space reveals large regions where the PDF becomes negative, which complicates the convergence of the fitting procedure as the PDF approaches this physical boundary. To mitigate this issue, the A_6 parameter is transformed such that the allowed space in the A_6-F_L plane is square, resulting in a new parameter $A_6{}'$, defined as:

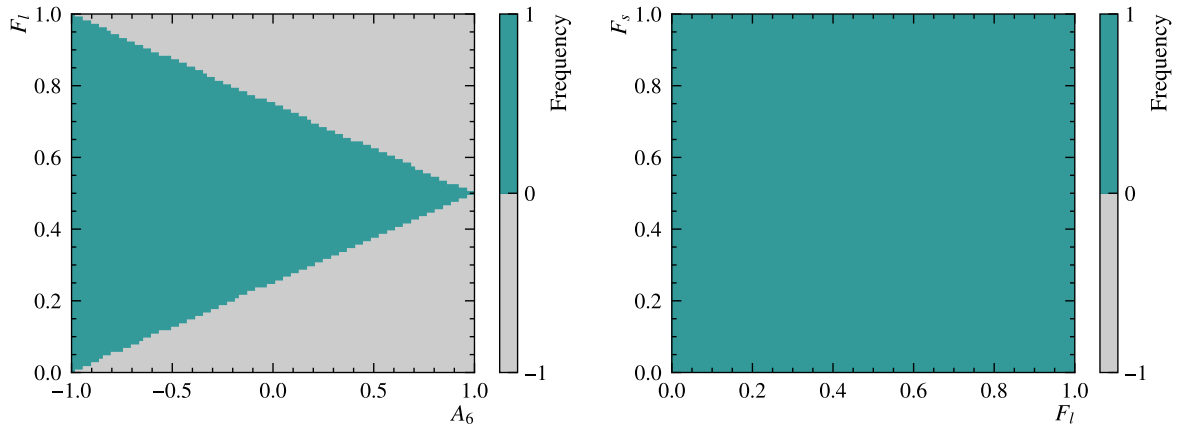
$$A_6{}' = A_6 \frac{3}{4}(1 - F_L). \tag{4.27}$$

This transformation serves to remove unphysical negative values from the region where the A_6-F_L plane was previously poorly constrained, simplifying the parameter space for

fitting. To ensure the PDF remains positive and physically meaningful, parametrisation is employed to transform parameters like A_6 into ${A_6}'$, which helps avoid regions where the PDF becomes negative. This reparametrisation stabilises the fitting process by keeping the model within physically allowed parameter spaces, preventing the fitting procedure from producing unphysical results.

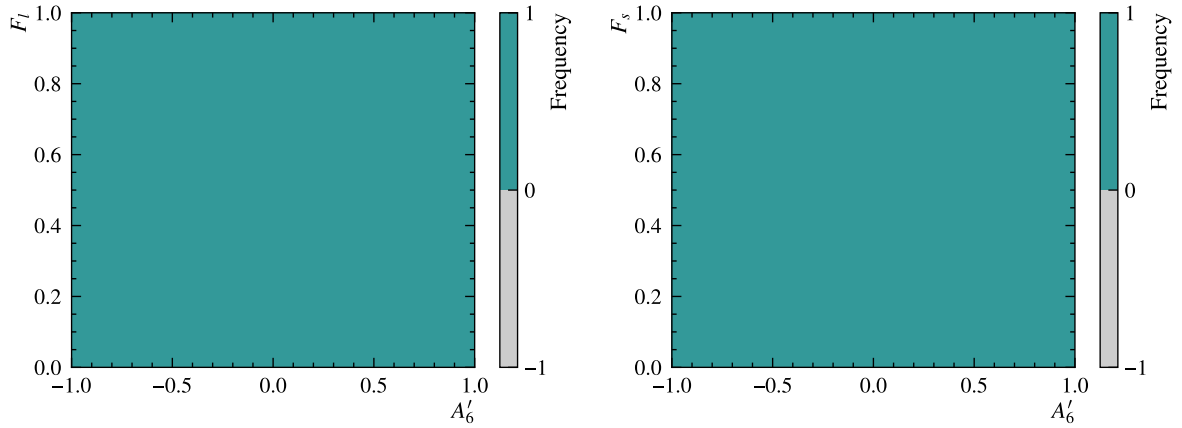
The scans of the remaining observables in the 2D $\cos\theta_h$ and $\cos\theta_l$ function show stable behaviour across their respective parameter spaces. These scans indicate no significant unphysical behaviour outside of the regions where the transformation is applied. The consistency across these spaces supports the robustness of the approach, as the majority of the parameter space does not encounter issues that would require further reparametrisation.

The scan of the observables in the 1D ϕ function show the behaviour of ${S_3}^2 + {A_9}^2 < 1$, which is expected. This means that at the extreme bounds of the two-dimensional space, the function becomes unphysical, consistent with known physical constraints for these observables. As this is only the case in the corners of the space, no transformation is applied to this plane, as the unphysical behaviour does not significantly affect the results in the central regions of the parameter space. The remaining observables are stable across the entire parameter spaces. This supports the conclusion that the transformation is effective, without introducing unnecessary modifications to the other parameter spaces.



(a) F_L vs A_6

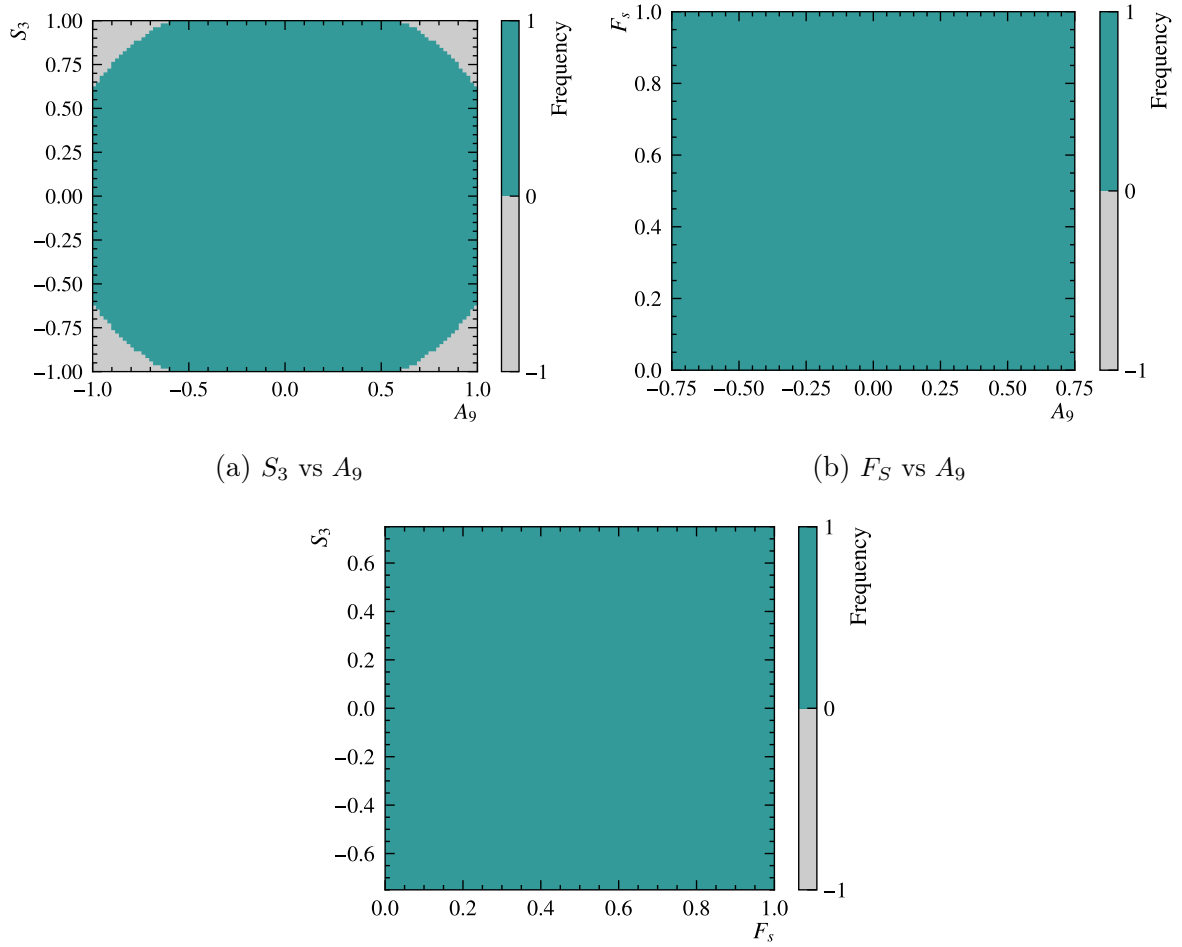
(b) F_S vs F_L



(c) F_L vs $A_6{}'$

(d) F_S vs $A_6{}'$

Figure 4.2.3: 2D scans of Equation 4.26 across the phase space of angular observables. The values of the observables are fixed while the $\cos\theta_h$ and $\cos\theta_l$ dimensions are scanned. If any of the points return a negative PDF the bin has a value of -1.



(a) S_3 vs A_9

(b) F_S vs A_9

(c) S_3 vs F_S

Figure 4.2.4: 2D scans of Equation 4.25 across the phase space of angular observables. The values of the observables are fixed while the ϕ dimension is scanned. If any of the points return a negative PDF the bin has a value of -1.

4.3 Angular Acceptance

The acceptance of the detector geometry, the trigger, and the selection requirements on the data affects the shape of the angular distributions. These effects must be accounted for in the modelling of the angular dimensions to accurately extract the angular observables. The angular acceptance is parametrised as a function of the three angular variables, $\vec{\Omega} = (\cos\theta_h, \cos\theta_l, \phi)$, and the masses of the dipion and dilepton systems, denoted as k and q respectively, across the full q^2 range. This parametrisation is chosen because these five variables fully describe the kinematics of the four-body decay. Any acceptance effects from detector geometry, trigger, and selection criteria must therefore be functions of these variables.

Phase-space simulation samples of the rare $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode are used for each year of data taking. The samples are generated according phasespace in the three helicity angles, and the dihadron and dilepton dimensions. The distributions are generated according to four-body phasespace. The samples share the same detector geometry, triggers, weighting, and selection requirements as the LHCb data, ensuring that the angular distributions of both samples experience the same distortions. However, residual systematic effects may remain due differences in the angular resolution between data and simulation, for example. These effects are studied and accounted for in the systematic uncertainties (see Section 5). The same correction weights discussed in Section 3.7 are applied.

4.3.1 Acceptance model

The acceptance function ε is expanded as a sum of products, with a cosine function for the ϕ dimension and Legendre polynomials $L(x)$ for the remaining dimensions, $\cos\theta_h$, $\cos\theta_l$, k', and q'. While this model does not explicitly model correlations between dimensions, these correlations are addressed through the validation procedure described in Section 4.3.3, ensuring that the final model captures any dependencies between the variables. The Legendre polynomials are defined over the interval $[-1, 1]$. While the angular variables $\cos\theta_h$ and $\cos\theta_l$ naturally fall within this range, the mass distributions k and q do not. Thus, they are reparametrised into the variables k' and q' using:

$$k' = \frac{2(k - \min_k)}{\max_k - \min_k} - 1 \quad \text{and} \quad q' = \frac{2(q - \min_q)}{m(B^0) - k - \min_q} - 1. \tag{4.28}$$

Here, k is scaled to k' using the minimum and maximum values of k in the dataset. For q, the reparametrisation into q' uses the kinematic limit of $m(B^0) - k$ as the maximum value. In the original $k - q$ space, the allowed region is triangular due to the kinematic constraint

$q \leq m(B^0) - k$. This transformation ensures that both the k' and q' distributions are mapped to the square limits of $[-1, 1]$. As q' is defined using this kinematic limit, the triangular region is embedded into the new square space.

The acceptance function is then expressed as:

$$\varepsilon(\vec{\Omega}, k', q') = \sum_{l,m,n,o,p} c_{l,m,n,o,p} L_l(\cos\theta_h) L_m(\cos\theta_l) \cos(n\phi) L_o(k') L_p(q'), \tag{4.29}$$

where the indices l, m, n, o, p refer to the dimensions $\cos\theta_h$, $\cos\theta_l$, ϕ, k', q', respectively. In this expansion, $c_{l,m,n,o,p}$ are the coefficients which will be determined using the Method of Moments [188]. The method projects the acceptance function onto the basis functions, the Legendre polynomials and cosine function, and the orthogonality of the functions is used to extract the coefficients $c_{l,m,n,o,p}$. The Legendre polynomials satisfy the orthogonality relation:

$$\int_{-1}^{1} L_a(x) L_b(x) dx = \frac{2}{2a+1} \delta_{ab}. \tag{4.30}$$

Similarly, the cosine function used for the ϕ dimension satisfies the orthogonality condition:

$$\int_0^{2\pi} \cos(ax)\cos(bx) dx = \begin{cases} 2\pi & \text{if } a = b = 0 \\ \pi\delta_{ab} & \text{otherwise} \end{cases}. \tag{4.31}$$

The following integral is approximated to calculate the moments $M_{l,m,n,o,p}$, which capture the contributions of Equation 4.29 to each dimension:

$$M_{l,m,n,o,p} = \int_{-1}^{1} \varepsilon(\vec{\Omega}, k', q') L_l(\cos\theta_h) L_m(\cos\theta_l) \cos(n\phi) L_o(k') L_p(q') \; d\vec{\Omega} \; dk' \; dq'. \tag{4.32}$$

In practice, the approximation of the integral is performed by summing over the N data points in the sample, where each data point i weighted by w_i. The moment $M_{l,m,n,o,p}$ is then calculated as:

$$
\begin{aligned}
M_{l,m,n,o,p} &= \frac{1}{\sum_i w_i} \sum_i^N w_i L_l(\cos\theta_h) L_m(\cos\theta_l) \cos(n\phi) L_o(k') L_p(q') \\
&= \frac{1}{\sum_i w_i} \sum_i^N w_i c_{l,m,n,o,p} \left(\frac{2}{2l+1}\right)\left(\frac{2}{2m+1}\right)\left(\frac{2}{2o+1}\right)\left(\frac{2}{2p+1}\right) \\
&\times \begin{cases} 2\pi & \text{if } n = 0 \\ \pi & \text{otherwise} \end{cases}.
\end{aligned}
\tag{4.33}
$$

Summing over the N data points empirically estimates the moment integral. Using the orthogonality conditions of the Legendre polynomials and the cosine function, the coefficients $c_{l,m,n,o,p}$ are related to the moments $M_{l,m,n,o,p}$, as follows:

$$
\begin{aligned}
c_{l,m,n,o,p} = \frac{1}{\sum_i w_i} \sum_i^N & w_i L_l(\cos\theta_h) L_m(\cos\theta_l) \cos(n\phi) L_o(k') L_p(q') \\
& \times (2l+1)(2m+1)(2o+1)(2p+1) \begin{cases} \frac{1}{2\pi} & \text{if } n = 0 \\ \frac{1}{\pi} & \text{otherwise} \end{cases}.
\end{aligned}
\tag{4.34}
$$

The Legendre polynomials used for the angular variables $\cos\theta_h$ and $\cos\theta_l$ are even-ordered. This is due to the assumption that there is negligible asymmetry in the detection acceptance between the muons in the dimuon system and the pions in the dipion system. A study of the systematic uncertainty relating to potential asymmetries in the detection acceptance is beyond the scope of this section and is intended be addressed in the complete analysis. The four distributions parametrised using Legendre polynomials are described by high orders, these shapes are complex, particularly at the boundaries. A choice of lower orders was found ineffective in describing the shapes of the distributions, while higher orders introduced increased complexity into the model by increasing the number of $c_{l,m,n,o,p}$ significantly. The cosine function used to describe the ϕ distribution requires a smaller order, increasing this value does not improve the description of the model.

The orders of the cosine and Legendre polynomials used to extract the acceptance from the simulated $B^0 \to \pi^+\pi^-\mu^+\mu^-$ phase-space samples are summarised in Table 4.7. The choice of the polynomial orders in the acceptance model was determined based on the following considerations: Even-order Legendre polynomials for the angular variables $\cos\theta_h$ and $\cos\theta_l$ were selected under the assumption of negligible asymmetry between the muons and pions, as discussed earlier. Higher-order polynomials were used for the mass distributions, k' and q', to accurately describe the complexity of the distribution shapes, especially at the boundaries. The validation of these choices is presented in Section 4.3.3.

Variable	Function	Order
$\cos\theta_h$	Even-order Legendre	4
$\cos\theta_l$	Even-order Legendre	4
ϕ	Cosine	4
k'	Legendre	8
q'	Legendre	8

Table 4.7: Configuration of the acceptance model from Equation 4.29.

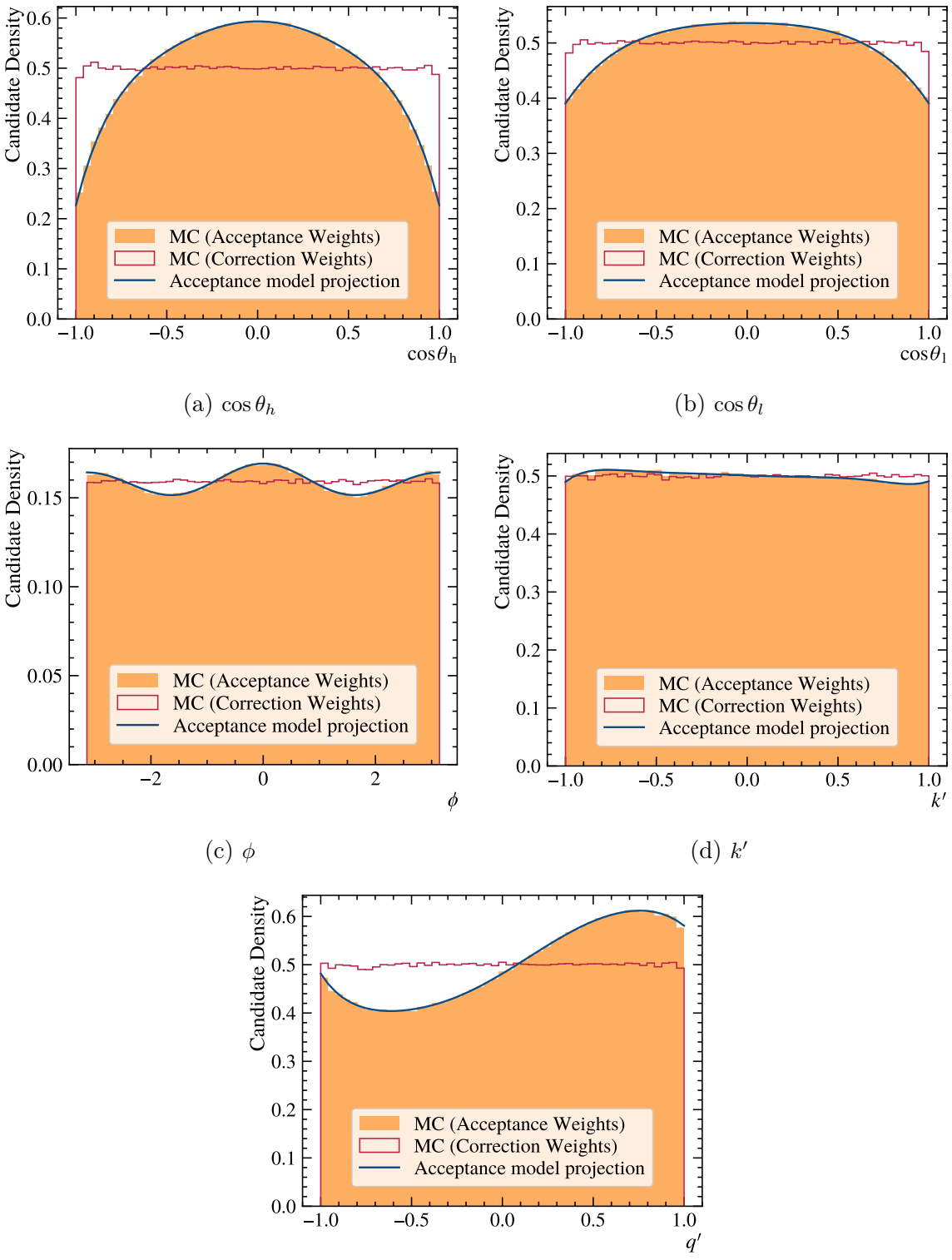
The parametrised acceptance is shown in Figure 4.3.1 for each of the five dimensions, alongside the simulation samples on which the acceptance is modelled. The figure includes the MC samples weighted to remove the acceptance (red), and this distribution is expected to be flat following the removal of acceptance effects. A Kolmogorov-Smirnov test [183] is performed to quantify the agreement between the corrected distributions and a uniform distribution in each dimension. The results, shown in Table 4.8, indicate that we cannot reject the null hypothesis in any of the dimensions. While there is a mild disagreement in the q' distribution (p-value = 0.184), this is well above the usual threshold of 0.05, suggesting that the deviation is not statistically significant. Overall, the results indicate that the weights have successfully corrected for the acceptance distortions when testing in a single dimension.

Distribution	p-value
$\cos\theta_h$	0.998
$\cos\theta_l$	0.995
ϕ	0.998
k'	0.735
q'	0.184

Table 4.8: Results from Kolmogorov-Smirnov tests [183] comparing acceptance-corrected distributions in the simulation data from the combined Run I and Run II datasets with a uniform distribution. A p-value greater than 0.05 indicates that we cannot reject the null hypothesis, suggesting that the distributions are statistically consistent.

4.3.2 Reducing the acceptance into three-dimensions

The angular PDFs used to extract the angular observables were introduced in Section 4.2.1. These functions only concern SM information without making assumptions about the data being fit. With the angular acceptance now parametrised, the detector and reconstruction effects need to be folded into the models.



(a) $\cos\theta_h$

(b) $\cos\theta_l$

(c) ϕ

(d) k'

(e) q'

Figure 4.3.1: Angular acceptance parametrisation projected into each of the dimensions: $\vec{\Omega}$, k', and q'. Plots show the corrected MC sample (orange) and the acceptance-corrected sample (red). The blue line shows the acceptance fitted to the weighted MC. The corrected (red) distribution should appear uniform.

Since the k' or q' distributions are not being fit, the acceptance function is reduced in those dimensions. The value of q' is fixed within the relevant q^2 bin, either at the world average mass of the $c\bar{c}$ resonance (in the case of the $B^0 \to J/\psi\pi^+\pi^-$ and $B^0 \to \psi(2S)\pi^+\pi^-$ samples) or the average of the bin, of the sample used to parametrise the acceptance, for the rare regions. The k' value is taken from the average value of data in the region $m(\pi^+\pi^-) \in [620, 920]\,\text{MeV}/c^2$.

While this approach introduces a dependence on the choice of fixed values, comparisons between different representative choices (including bin centres, averages, and resonance masses) show a mild variation in the resulting angular acceptance function. These comparisons are illustrated in Appendix D. The chosen values are considered reasonable for the current analysis. However, the acceptance function does vary more significantly when evaluated at the bin edges. A more detailed evaluation of the impact of this reduction in dimensions, through a dedicated systematic uncertainty, remains outside the scope of this thesis and is deferred to future work.

The acceptance is then simplified by summing over the now-fixed dimensions, o and p:

$$\varepsilon(\vec{\Omega}) = \sum_{l,m,n} \left(\sum_{p} \sum_{o} c_{l,m,n,o,p} L_o(k') L_p(q') \right) L_l(\cos\theta_h) L_m(\cos\theta_l) \cos(n\phi). \tag{4.35}$$

The acceptance function has now been reduced to three dimensions: $\cos\theta_h$, $\cos\theta_l$, and ϕ. The final step for determining the contribution of the acceptance to the angular PDFs used in the fitting is to integrate the product of the three-dimensional acceptance function and the differential decay rate from Equation 1.52, denoted $f_{SM}(\vec{\Omega})$. Previously, Equation 4.25 was defined as the 1D PDF which is fit to the ϕ distribution. This function was derived by integrating $f_{SM}(\vec{\Omega})$ over the $\cos\theta_h$ and $\cos\theta_l$ dimensions. To incorporate the acceptance into Equation 4.25, the integration is performed for each combination of orders for the indices l, m, and n, as follows:

$$f(\phi) = \sum_{l,m,n} c_{l,m,n} \cos(n\phi) \int_{-1}^{1} \int_{-1}^{1} f_{SM}(\vec{\Omega}) L_l(\cos\theta_h) L_m(\cos\theta_l) d\cos\theta_h d\cos\theta_l. \tag{4.36}$$

This sum introduces extra terms into Equation 4.25. Many of the terms are 0 due to the combination of integrals in the acceptance functions. The same process is carried out for the 2D PDF used to fit the $\cos\theta_h$ and $\cos\theta_l$ dimensions, from Equation 4.26, by integrating over the ϕ.

4.3.3 Validation of the acceptance model

The angular acceptance has been parametrised in five dimensions (5D) using a large sample of simulated phase-space $B^0 \to \pi^+\pi^-\mu^+\mu^-$ candidates. Figure 4.3.1 demonstrates that each of the angular and mass dimensions is individually well-described. However, to ensure that the model also captures potential correlations between dimensions, a more robust multidimensional test is required. This is achieved using a pseudodata validation strategy, described below.

The idea is to compare the original simulation sample, which is subject to the detector and reconstruction effects, with pseudodatasets generated from the fitted acceptance model. If the model accurately captures the full structure of the original sample, including all correlations, then the pseudodatasets should be statistically indistinguishable from the original MC-simulated sample.

To generate these pseudodatasets, an accept-reject Monte Carlo algorithm [189] is used. This method allows for direct sampling from the 5D acceptance model, treating it as a probability density function over the angular and mass variables. The flat phase-space prior ensures that the input distribution in $\vec{\Omega}$, k, and q is uniform, meaning that any structure observed in the generated pseudodata arises solely from the acceptance model. As such, the accept-reject approach preserves and reflects all correlations captured by the model; if the acceptance model includes correlations between variables, they will naturally be present in the pseudodata.

To ensure fair comparison, the size of the pseudodatasets is matched to the effective sample size (ESS) of the original simulation dataset [190]. This accounts for the weights applied to the phase-space MC and ensures equivalent statistical power between the samples:

$$\text{ESS} = \frac{\left(\sum_{i=1}^{N} w_i\right)^2}{\sum_{i=1}^{N} w_i^2}, \tag{4.37}$$

where N is the number of events and w_i is the weight of event i.

Comparison of MC and pseudodata

The aim of the comparison is to determine whether the acceptance model correctly reproduces the structure, including correlations, of the MC-simulated sample. If the acceptance model omits significant correlations, then a classifier should be able to distinguish the pseudodata from the MC sample. Conversely, indistinguishability suggests that the model successfully captures the relevant multidimensional features of the acceptance.

To test this, a total of fifty pseudodatasets are generated from the acceptance model using the accept-reject method. Each pseudodataset is compared against the MC-simulated

sample using a Boosted Decision Tree (BDT) classifier. The BDT is trained to separate the two datasets based on the five input dimensions: $\cos\theta_h$, $\cos\theta_l$, ϕ, k', and q'. Because the BDT can learn non-linear combinations of these variables, it is particularly sensitive to multidimensional correlations, not just marginal differences.

Importantly, both the MC-simulated dataset and the pseudodatasets are generated from a flat phase-space prior; no physics model (e.g., angular interference or resonant structure) is included. Therefore, any differences between them must originate from how the acceptance model represents the detector effects, not from physics assumptions. Furthermore, although the input phase space is flat, the output of the accept-reject generation reflects the non-flat, potentially correlated structure of the acceptance model. Thus, this test is sensitive only to the adequacy of the acceptance model.

To evaluate the outcome of the BDT training, the Area Under the ROC Curve (AUC) is computed for each run. A second set of control trainings is performed, where the BDT is trained to distinguish between pairs of pseudodatasets (MC). These serve as a reference distribution, representing the expected AUC when comparing statistically equivalent samples. Figure 4.3.3 compares the two sets of AUC scores.

A one-sample t-test [191] is performed between the pseudodata-vs-MC AUC scores and the pseudodata-vs-pseudodata baseline. The resulting p-value of 0.1 indicates that the difference in means is not statistically significant, and we cannot reject the null hypothesis that the two sample types are drawn from the same distribution. This provides strong evidence that the acceptance model accurately reproduces the multidimensional structure, including correlations, present in the original MC-simulated sample.

Some BDT runs produce AUC values slightly below 0.5, indicating random statistical fluctuations that are occasionally worse than random guessing. These are expected and do not undermine the overall conclusion.

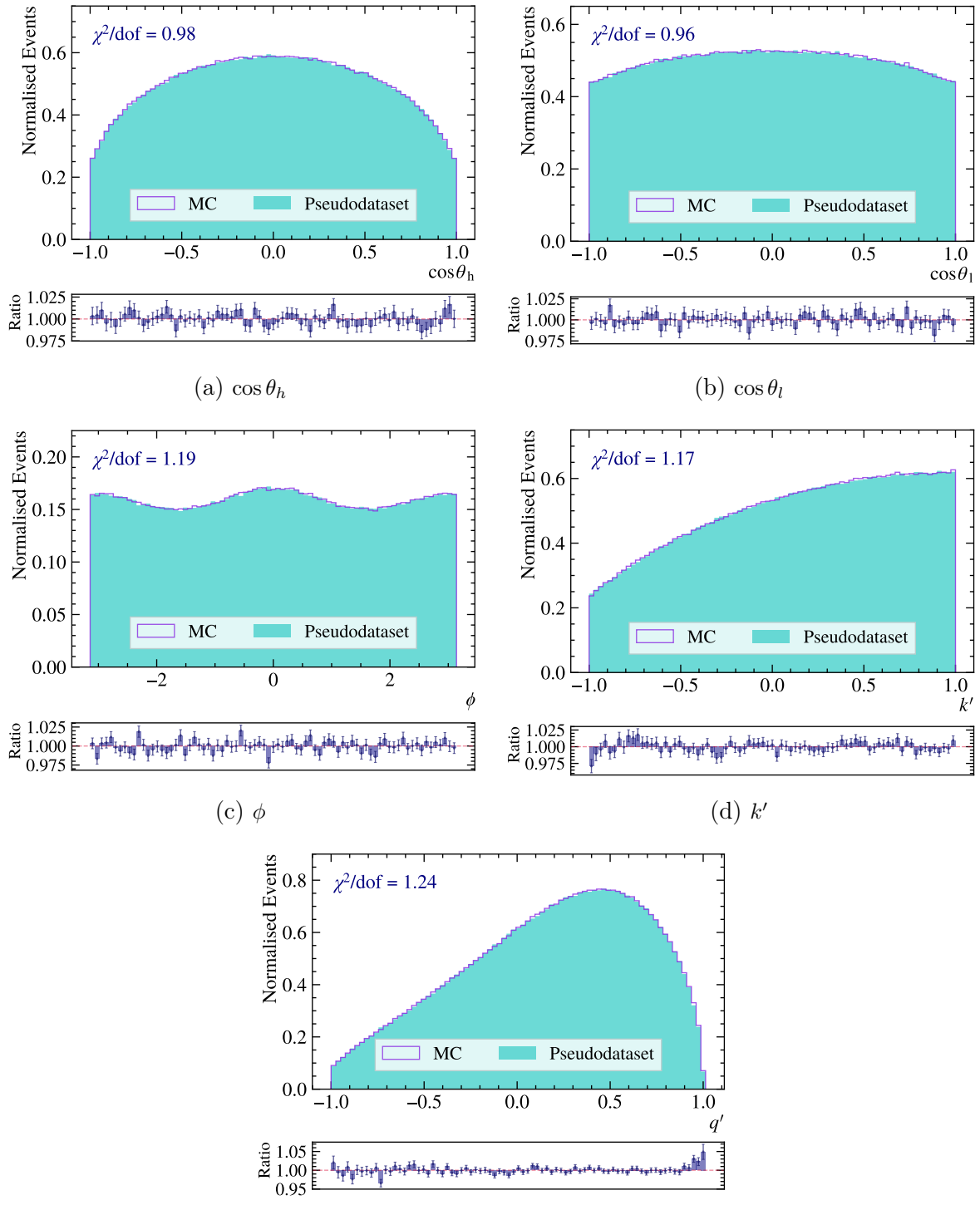


Figure 4.3.2: Comparison of phase-space MC and simulation samples generated from the angular acceptance model in each of the dimensions $\vec{\Omega}$, k', and q'. The plots each show the MC samples weighted to match the LHCb data and an example of a simulated pseudodataset. Below is the ratio of the two samples, which is used as a one-dimensional visualisation of the agreement of the acceptance model and the phase-space MC.

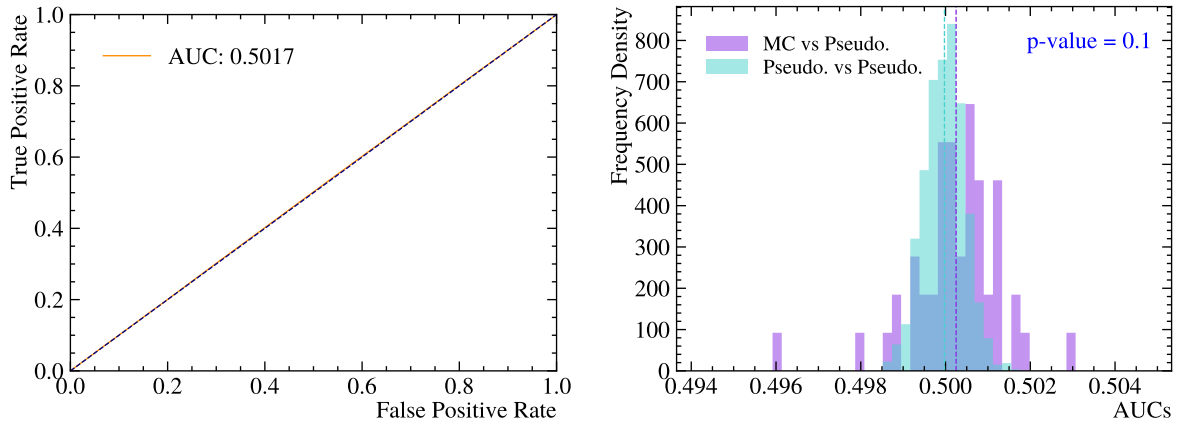


Figure 4.3.3: Comparison of MC and pseudodatasets generated from the efficiency function modelled to the MC sample. The left figure shows an example ROC curve for BDT trained to distinguish between two datasets, a diagonal line rising from the lower left corner, with an area of 0.5, identifies that the BDT is unable to distinguish between them. The right figure shows the distribution of AUCs for "MC vs pseudodataset" and "pseudodataset vs pseudodataset", with the p value determined from the test. The dashed lines are at the mean values of the distribution they share a colour with.

4.4 Fitting simulation datasets

Measuring the angular observables using the dedicated simulation samples provides a test of the modelling method on large samples with known observables values, including the inclusion of the acceptance. This section details the modelling of samples in the J/ψ q^2 region, and in the low-q^2 and high-q^2 regions.

4.4.1 Fits to $B^0 \to J/\psi\pi^+\pi^-$ simulation

The $B^0 \to J/\psi\pi^+\pi^-$ simulation samples are generated so as to replicate the results from the amplitude analysis of the decay mode [53]. After applying selection criteria and correction weights, should closely replicate the characteristics of the real $B^0 \to J/\psi\pi^+\pi^-$ dataset. Validating the modelling on this sample serves as an important test, particularly for assessing how well the model captures the physics of the resonant mode. The fits are performed for the product of the angular distribution and the acceptance. The fits to the angular distributions within the J/ψ q^2 region are shown in Figure 4.4.1.

The expected values of the observables are derived directly from the amplitudes used during the generation of the sample. These expected values are directly compared to those obtained from the fits. Both the observed and expected values are presented in Table 4.9. the largest difference between the expected and observed values is in the measurement of the F_L observable, however this is small compared to the sensitivity of the dataset, discussed in Section 4.4.3, and is likely due to the imperfect nature of the acceptance modelling. The good agreement between these values indicates that the efficiency has been sufficiently incorporated into the modelling.

To assess potential bias in the fits, a coverage test is performed by splitting the high-statistics MC sample into 11 subsamples, each containing an event count comparable to the signal yield in data, as determined in Section 4.1.1, Table 4.1. The results of this test, also shown in Table 4.9, demonstrate that the observed values remain consistent across the subsamples, with no significant deviation from the expected values. The mean and standard deviation of the measured values across the subsamples are reported alongside the expected values.

4.4.2 Fits to $B^0 \to \rho^0\mu^+\mu^-$ simulation in rare q^2 bins

Simulation samples of the $B^0 \to \rho^0\mu^+\mu^-$ decay mode in the rare q^2 regions are reconstructed, have selection requirements applied, and are divided into low-q^2 and high-q^2 regions. The sculpting effects on the angular distributions caused by the selection process are accounted for by incorporating the efficiency functions derived from phase-space

Observable	Expected Value	Observed Value	
		Full Sample	Subsample Mean
F_L	0.600	0.604 ± 0.001	0.589 ± 0.005
S_3	-0.039	-0.042 ± 0.002	-0.039 ± 0.007
A_6	0.000	-0.006 ± 0.005	0.001 ± 0.017
A_9	0.000	0.000 ± 0.002	-0.001 ± 0.008

Table 4.9: Results of the fit validation for $B^0 \to J/\psi\pi^+\pi^-$ simulation samples. The expected values are given alongside the value observed using the full sample and the mean value, with the standard deviation, obtained from a fit to 11 subsamples of realistic size. The expected values are derived from the values of the helicity amplitudes given to the generation software.

simulation samples. The fits to the angular distributions within the rare q^2 regions are shown in Figure 4.4.2.

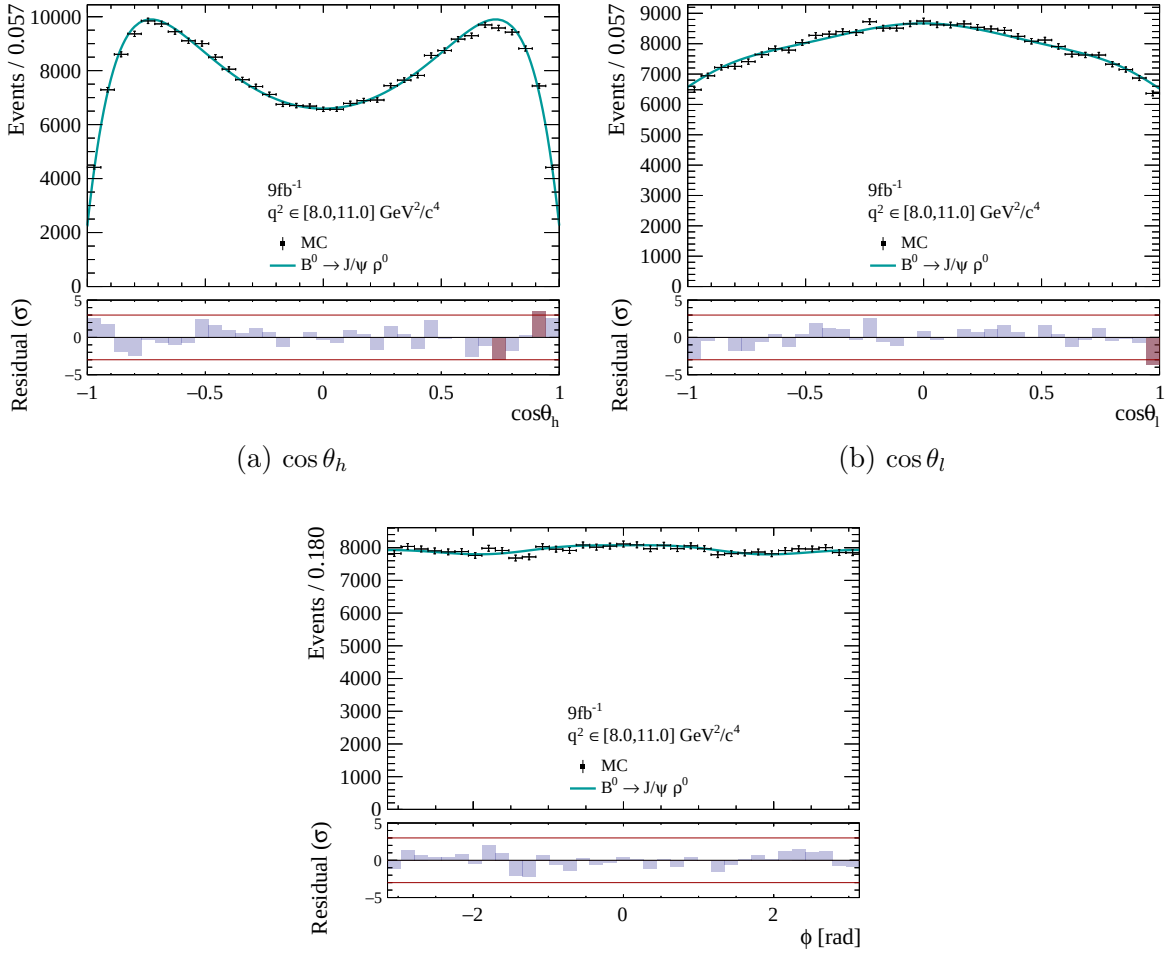
The samples are split into the low-q^2 and high-q^2 bins when modelling, this selection affects the values of the angular observables, which are q^2-dependent. To determine the expected values a large pseudodataset is generated using the same EvtGen [150] model used to generate the simulation sample. The pseudodataset has no sculpting from acceptance and no selection requirements applied, except for cuts separating the sample into the two q^2 bins. Fits are performed on these samples and the measured observables are those that are expected from the fits to the $B^0 \to \rho^0\mu^+\mu^-$ simulation samples.

The expected and observed values are given in Table 4.10. There is a bias in the measurement of the F_L observable between the observed and expected values. This bias likely comes from the treatment of the values of k' and q' with relation to the acceptance model. The values are chosen to be the average value of the distributions in the sample used to parametrise the acceptance. As there is no acceptance included in the pseudodataset used to determine the expected values there is potential for biasing in the choice of the k' and q'. As the observables are dependent on q^2 the choice of the q' value is of particular importance. The other observables are unbiased.

While this bias is observed in the fit to the simulation samples, it could also potentially affect the final fit to the signal data. This issue will need to be further explored and addressed as the analysis progresses, beyond the submission of this thesis, to ensure accurate modelling of the signal.

4.4.3 Fit validation using simple pseudoexperiments

The fitting strategy is validated using pseudoexperiments. In this section, signal-only pseudodatasets are generated, followed by a study including backgrounds in the next section. Signal-only samples are generated to evaluate the model's behaviour in response to angular observables, without the complications introduced by background noise.



(a) $\cos\theta_h$ (b) $\cos\theta_l$

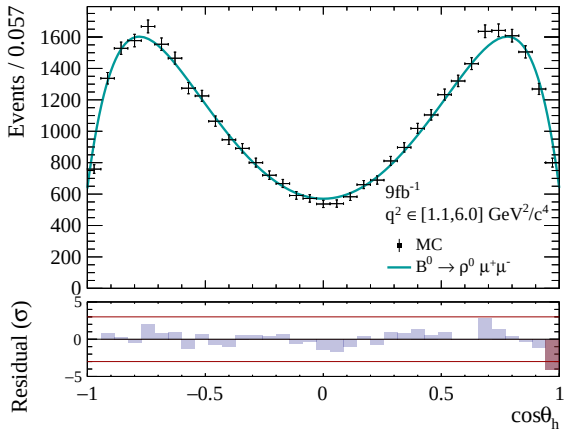
(c) ϕ

Figure 4.4.1: Angular fits to the 9 fb^{-1} simulated dataset of the $B^0 \to J/\psi\pi^+\pi^-$ decay mode. The angular acceptance is folded into the fitting model using moments. The measured observables are given in Table 4.9.

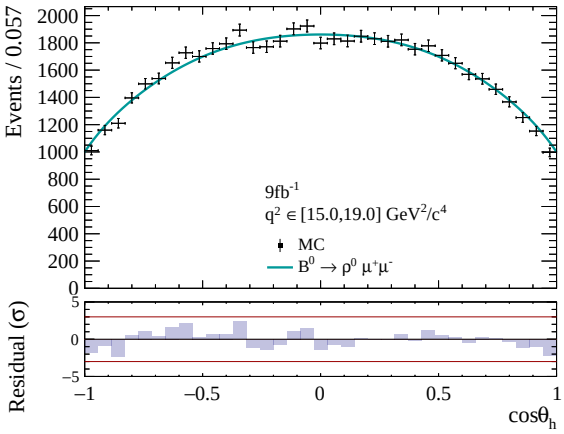
Observable	**Low-q^2 [1.1-6.0 GeV^2]**		**High-q^2 [15.0-19.0 GeV^2]**	
	Expected	Observed	Expected	Observed
F_L	0.762 ± 0.003	0.739 ± 0.005	0.341 ± 0.003	0.326 ± 0.004
S_3	-0.006 ± 0.006	-0.027 ± 0.008	-0.156 ± 0.006	-0.181 ± 0.008
A_6	0.091 ± 0.012	0.094 ± 0.019	-0.018 ± 0.006	-0.018 ± 0.006
A_9	0.000 ± 0.004	0.003 ± 0.005	0.002 ± 0.004	-0.008 ± 0.006

Table 4.10: Results of the fit validation for $B^0 \to \pi^+\pi^-\mu^+\mu^-$ simulations samples, showing expected and observed values for both low-q^2 and high-q^2 bins.

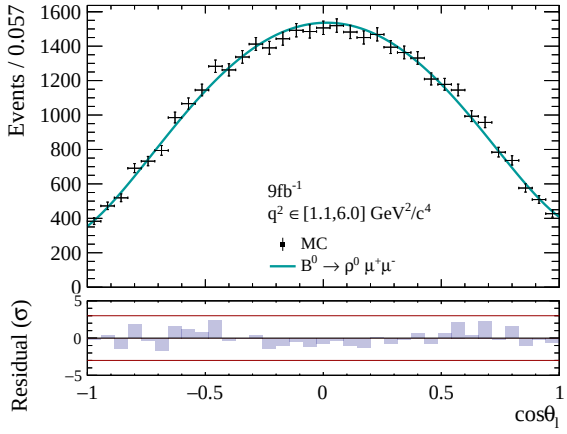
A pseudodataset is generated from the differential decay rate in the three angular dimensions. The randomly-sampled candidates are then accepted or rejected based on the parametrised acceptance, where the q' dimension is fixed to the average of the respective bin in the model, and the k' parameter is not used in the modelling.



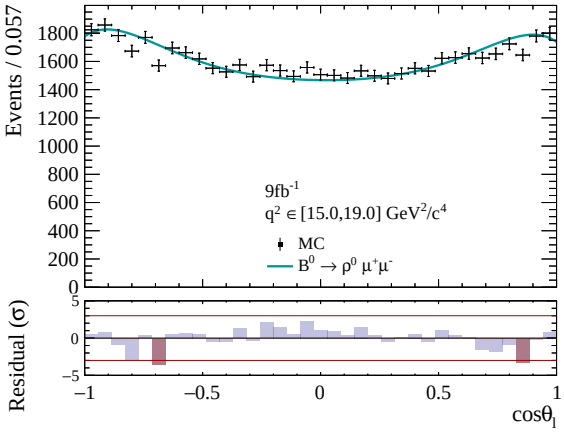
(a) $\cos\theta_h$, low-q^2



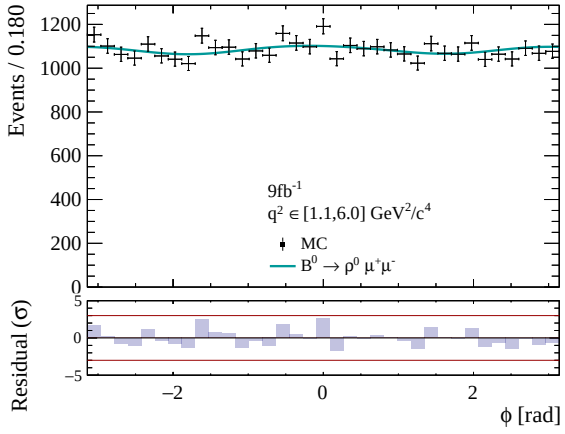
(b) $\cos\theta_h$, high-q^2



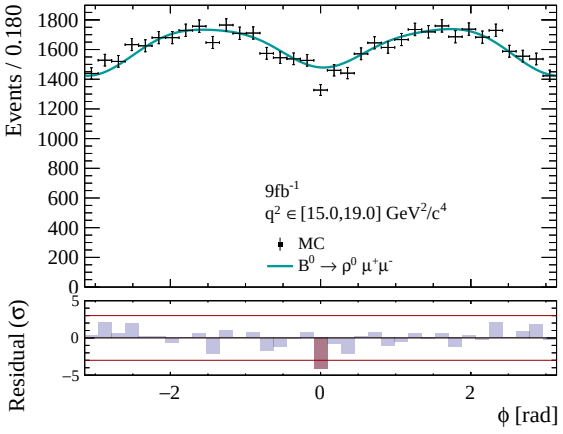
(c) $\cos\theta_l$, low-q^2



(d) $\cos\theta_l$, high-q^2



(e) ϕ, low-q^2



(f) ϕ, high-q^2

Figure 4.4.2: Angular fits to the $9\,\mathrm{fb}^{-1}$ simulated dataset of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode in the low- and high-q^2 bins. The angular acceptance is folded into the fitting model using moments. The measured observables are given in Table 4.10.

Realistic event counts for the components are taken from the yields determined from fitting the invariant mass of the respective samples, as discussed in Section 4.1. The b-hadron mass shapes of each component are derived from fits to their respective simulation samples. The q^2 and $m(\pi^+\pi^-)$ values are fixed to the average of the respective bins in the rare q^2 samples, and the q^2 to the mass of the J/ψ in the resonant samples. While this simplification is not perfect, it corresponds to the fit strategy to data and therefore represents the best approach to modelling the data in the three angular observables. The model itself is q^2-dependent, and since the fitting procedure focuses on the acceptance with respect to angular observables, fixing q^2 to its average value in each bin is a reasonable strategy in this context. In total, a thousand unique datasets are generated.

The modelling strategy is applied using the nominal method described in earlier sections. As only the signal component is kept in the model the invariant mass modelling is not needed to isolate the signal candidates. Therefore, only the angular distributions are modelled, and the values of the angular observables are extracted for each of the pseudodatasets.

The agreement between the observed and expected values is determined using the pull. Over a large number of pseudo-experiments, the distribution of the pull values approach a normal distribution. If the modelling is unbiased and the observable is unconstrained, the pulls will have a mean of zero and a standard deviation of one. A width greater than one indicates that the uncertainty of the measured observable is underestimated, while a width less than one suggests it is overestimated.

A study is performed using the statistics and models of the J/ψ sample. The pull distributions of the observables F_L, S_3, A_6, A_9, and F_S, are shown in Figure 4.4.3. A summary is given in Figure 4.4.4, where the points are the distribution mean and the size of the error bars are the widths.

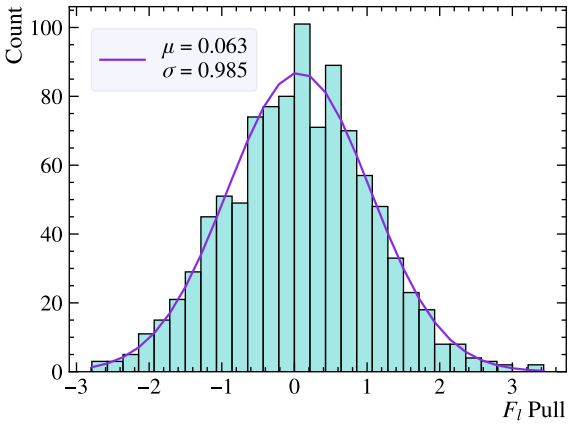
During the study using samples equivalent to the signal in the rare-q^2 regions, it is observed that the observable F_S frequently hit the physical boundary at 0, despite being generated at 0.15. This indicates that, with the current statistical power of the rare samples, the model lacks sensitivity to this observable. Consequently, F_S is fixed to 0.0 in the fits to the rare samples, including in the fits to the LHCb data. The effects of neglecting this observable are evaluated as a systematic uncertainty in Section 5.4.

Figure 4.4.5 shows the summary of the pull distributions for each angular observable across the 1000 pseudo-experiments in the low- and high-q^2 bins.

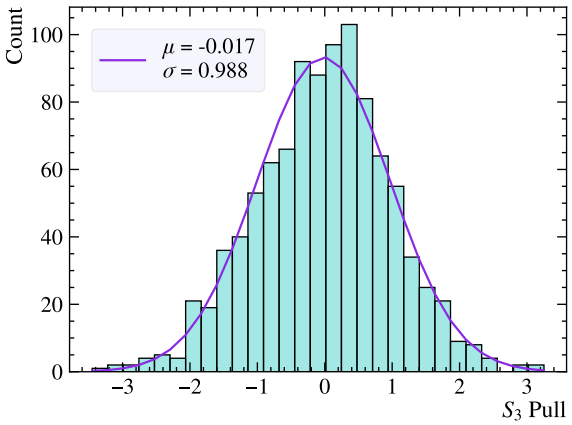
The sensitivity is determined from the adjusted uncertainty [186],

$$\sigma_{x,true} = \langle \sigma_x \rangle \times \sigma_{pull}, \tag{4.38}$$

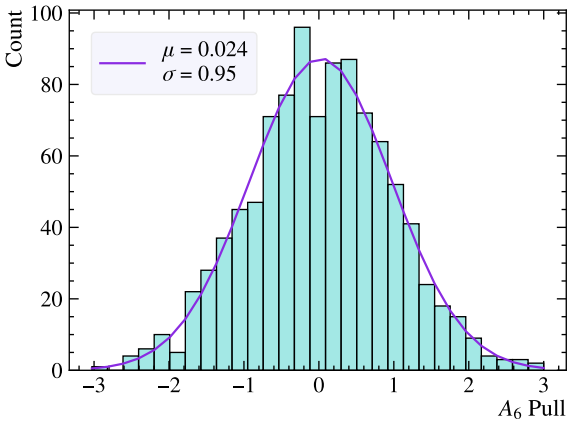
where $\langle \sigma_x \rangle$ is the average of the uncertainties over all pseudoexperiments, and the σ_{pull}



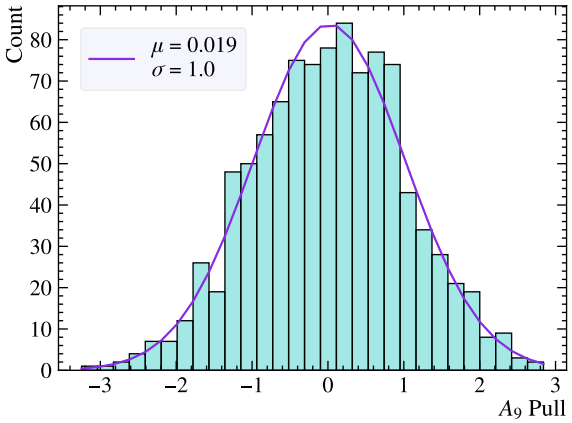
(a) F_L



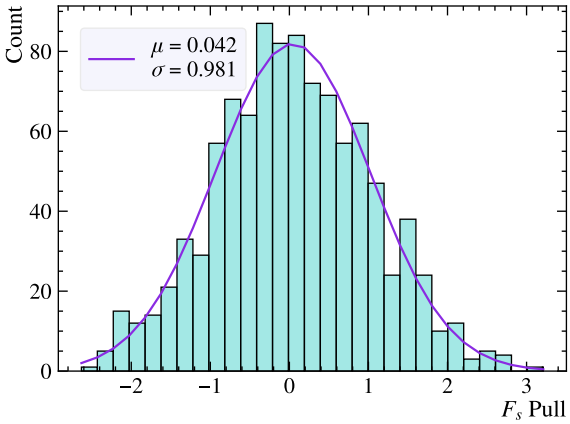
(b) S_3



(c) $A_6{}'$

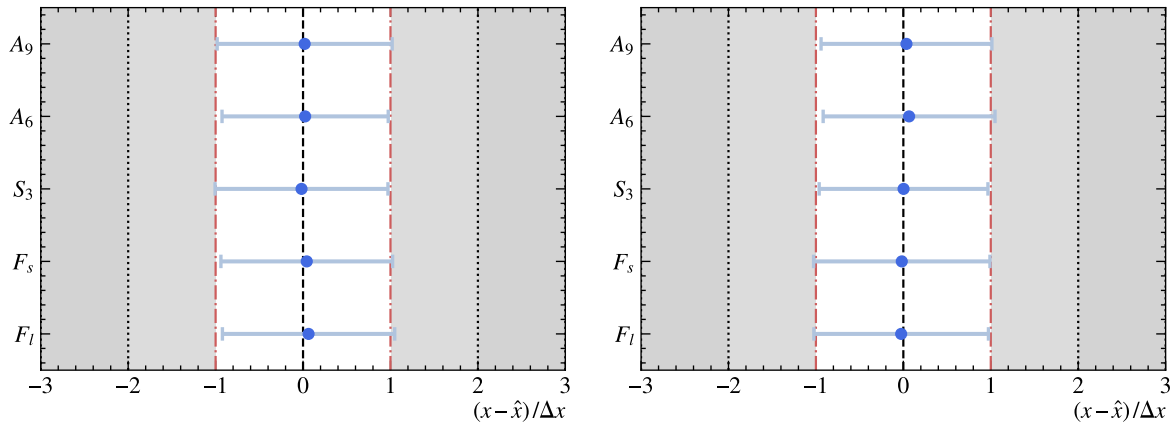


(d) A_9



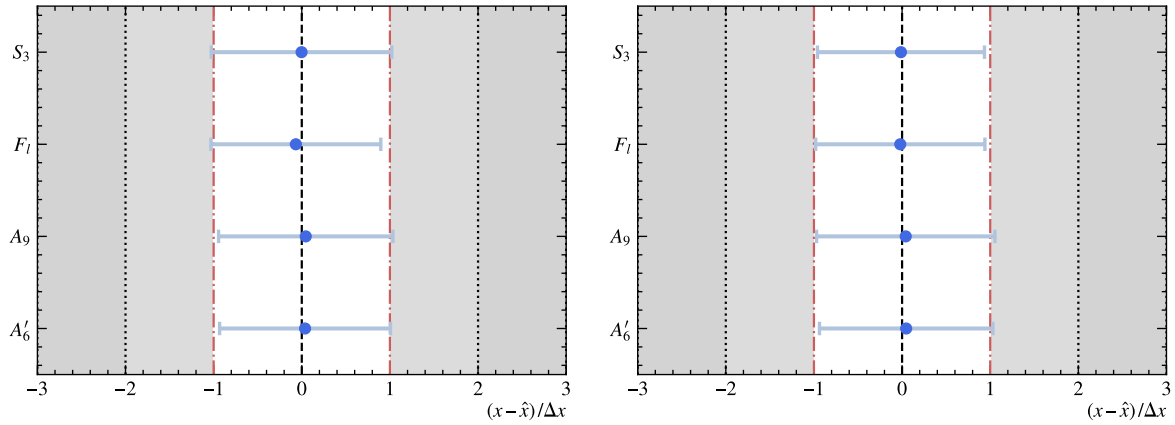
(e) F_S

Figure 4.4.3: Pull distributions of the angular observables measured using 1000 pseudo-experiments. The signal-only samples simulate the J/ψ dataset. The distributions are modelled with a normal Gaussian distribution, the legend shows the mean and width of the fitted shape.



(a) J/ψ q^2 (b) $\psi(2S)$ q^2

Figure 4.4.4: Results of the 1000 pseudoexperiments performed on signal-only sample simulating the (a) J/ψ and (b) $\psi(2S)$ datasets. The points are the mean of the pull distributions and the error bars are the standard deviation.



(a) Low-q^2 (b) High-q^2

Figure 4.4.5: Results of the 1000 pseudo-experiments performed on signal-only sample simulating the dataset in the rare q^2 bins. The points are the mean of the pull distributions and the error bars are the standard deviation.

is the pull width, or the standard deviation of the pull distribution. When the pull distribution has approximately unit width, the model describes the statistical fluctuations accurately, and the sensitivity is $\langle\sigma_x\rangle$ across all of the pseudoexperiments. However, if the pull width deviates from 1, indicating the model over- or underestimates the uncertainties. The adjusted uncertainty corrects for this by scaling $\langle\sigma_x\rangle$, accounting for mis-modelling. This adjusted sensitivity serves as an estimate of the expected uncertainty when applying the model to the real dataset with similar statistical properties. The sensitivity to each angular observable is given in Table 4.11.

4.4.4 Realistic pseudoexperiments

To ensure the robustness of the fitting strategy under realistic conditions, the method is validated using pseudo-experiments that include both signal and background components. Samples are generated to closely mirror the expected distributions from real data.

A pseudodataset is generated by sampling from a distribution intended to realistically match the expected LHCb data. The analysis considers four dimensions: the invariant mass, which is used to separate signal from background decay modes, and the three angular dimensions, which are modelled to extract the angular observables.

The shape of each component in the invariant mass of the b-hadron is determined from fits to the simulation samples, the combinatorial background shape is taken from a fit to the upper mass sideband in data. The angular distributions are the SM distributions, with the specific distribution determined by the spin of the dihadron resonance. The decays which come from a B^0 meson are generated with the P-wave distribution, while the B_s^0 decays are generated with an S-wave distribution. For the combinatorial background, the distribution in each angle is generated from a flat polynomial, representing a random distribution. The value of $m(\pi^+\pi^-)$ is fixed to the average value of the sample used to parametrise the acceptance, in the $m(\pi^+\pi^-)$ fit window. As the $m(\pi^+\pi^-)$ distribution is not fitted in the nominal procedure this has no effect on the outcome. While q^2 is fixed to either the mass of the $c\bar{c}$ resonance or, for the rare q^2 bins, to the bin average from the same sample, matching the fit procedure to data.

Events are generated in four dimensions, using fixed values in k' and q'. The efficiency of each event is determined using the moments derived from the efficiency modelling in Section 4.3. Events are accepted or rejected based on this efficiency. A thousand unique datasets are generated.

For the realistic pseudodatasets, which include background components, a mass fit is performed to separate the signal from the background via the sPlot procedure. The mass fit is conducted following the nominal method previously described in Section

Observable	Low-q^2	High-q^2	J/ψ	$\psi(2S)$
F_L	0.15	0.14	0.007	0.029
S_3	0.26	0.28	0.012	0.050
A_6	0.49	0.48	0.018	0.079
A_9	0.27	0.26	0.012	0.050
F_S	NA	NA	0.013	0.130

Table 4.11: Sensitivity of the model to the angular observables in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset. The sensitivity is determined as the average of the distribution of uncertainties over 1000 pseudo-experiments in each of the low-q^2, high-q^2, J/ψ, and $\psi(2S)$-q^2 bins. The F_S observable is not included in the fits to the rare-q^2 samples.

4.2.1. After the mass fitting, the signal candidates are extracted using sWeights and the angular observables are then extracted. This established procedure is repeated for each pseudodataset.

To assess the accuracy of the fitting strategy, the pull of each measurement is calculated for the realistic samples. Figure 4.4.6 shows the result of fitting a Gaussian to the pull distributions.

In the measurements of the observables in the J/ψ and $\psi(2S)$ q^2 bins, a bias is seen in the fitted value of F_S. A small bias is also present in the J/ψ bin for the F_L observable. The bias in the mean of pull distribution of F_S in the two bins is -2.1σ for the J/ψ bin, and -0.46σ in the $\psi(2S)$ bin, where σ is the statistical uncertainty. Both measurements have a pull width equal to 1, indicating that the uncertainties are well-estimated, but there is a systematic deviation. The origin of the bias is not fully understood and requires further attention, however, while the measurement of the F_S observable is biased, this is a nuisance parameter in this analysis. The performance of the other observables, with negligible bias and stable pull widths, indicates the model is robust to their measurements and they are unbiased. Further improvements to address this issue are expected while the analysis is reviewed by the collaboration.

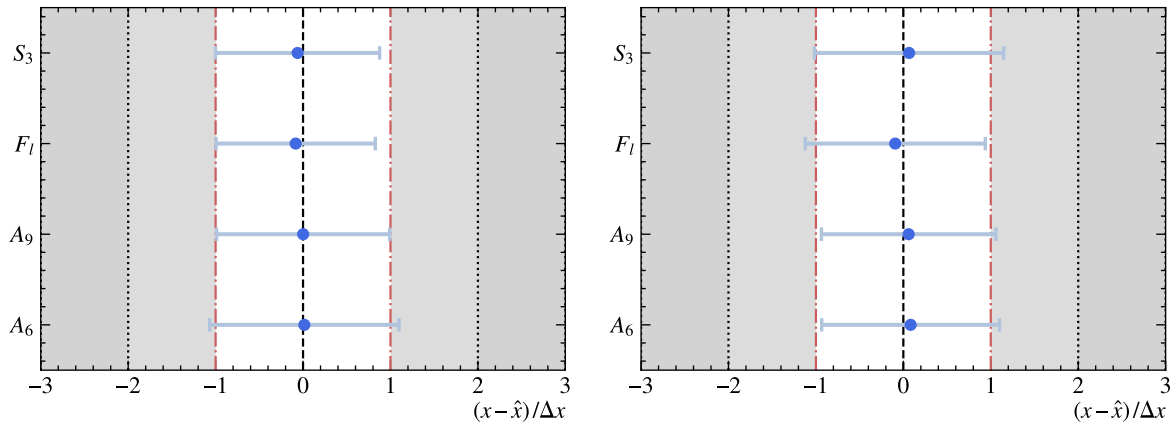
4.4.5 Stress testing the model

While the model has been confirmed to be unbiased on signal-only samples with sensible parameter values, it is crucial to evaluate its performance under more challenging conditions. In this stress test, the pseudodatasets are generated with observables values randomly shifted from those used in the initial generation, where the default values of the observables are set to zero, except for F_L, which is set to 0.5. The fitting is performed with the initial values in the model set to these default values. This test is conducted with 10 different sets of randomly shifted variables, with 200 generated pseudodatasets for each set. The use of 10 sets of variables tests the robustness of the model to handle differing conditions within the dataset.

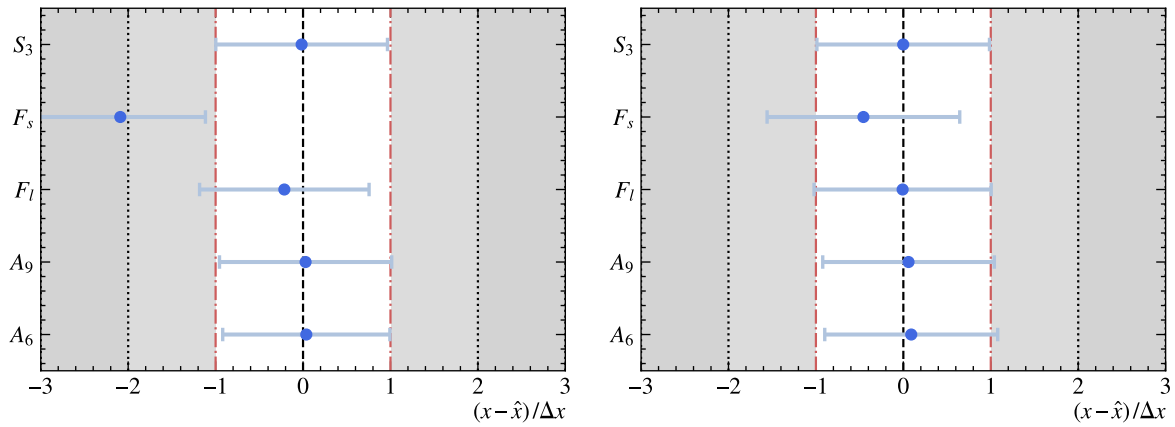
The mean absolute deviation (MAD) [192] is used to assess the bias in the fitted observables across each sets of generated values. It is calculated as

$$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N} \mid x_{fit} - x_{true} \mid, \tag{4.39}$$

where N is the number of sets of observables, x_{fit} is the mean fitted value from each set, and x_{true} is the corresponding generated value. The MAD value provide a measure of the absolute deviation between the true and fitted values. The value is compared to the uncertainty from the nominal studies in the previous section. If the MAD is



(a) Low q^2 (b) High q^2



(c) J/ψ q^2 (d) $\psi(2S)$ q^2

Figure 4.4.6: Summary of the pull distributions of performing the analysis on 1000 realistic pseudodataset simulating the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset in different bins of q^2. The points are the mean of the pull distributions and the error bars are the standard deviation, obtained by fitting the distributions with a Gaussian distribution.

consistent with the uncertainty, this is an indication as to whether the model is robust to randomisation of the true values of the observables, or not. The MAD values for each angular observables, across the ten sets of studies, are given in Table 4.12.

4.5 Modelling LHCb data

The models have been validated using dedicated simulation samples and are now tested on LHCb data samples, in the J/ψ and $\psi(2S)$ q^2 bins. Invariant mass fits have been performed on these samples, in Section 4.1, and sWeights have been calculated and applied, isolating the signal decay mode from the background. Finally, the angular distributions are modelled as a crucial test to determine the validity of the models. The measured

Observable	MAD			
	Low-q^2	High-q^2	J/ψ-q^2	$\psi(2S)$-q^2
F_L	-0.01	-0.01	-0.00	-0.00
S_3	-0.02	0.00	-0.00	-0.00
A_6	-0.02	-0.01	-0.00	0.00
A_9	-0.02	0.01	0.00	-0.00
F_S	NA	NA	-0.03	-0.01

Table 4.12: Mean Absolute Deviation (MAD) values for the angular observables in the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset. MAD values are reported for the low-q^2, high-q^2, and J/ψ-q^2 bins. The F_S observable is not included in the fits to the low-q^2 and high-q^2 bins.

observables are compared to previous measurements [53] for the J/ψ sample. As discussed in Section 4.2.1, the observables which are measured are the F_L, S_3, A_6, and A_9, and the fraction of S-wave, F_S.

The section concludes with the modelling of the LHCb data in the two rare q^2 bins. The measured observables are shifted by a random, blind amount, to hide the true values of the observables. The uncertainties, however, of the measurements are determined.

4.5.1 Modelling $B^0 \to J/\psi\pi^+\pi^-$ data

The amplitudes associated with the $B^0 \to J/\psi\pi^+\pi^-$ decay mode have been determined in previous studies [53]. This allows for a direct comparison with established results, further validating the accuracy of the modelling approach. The modelling process, including the handling of selection efficiency effects, follows the same strategy as discussed for simulation samples.

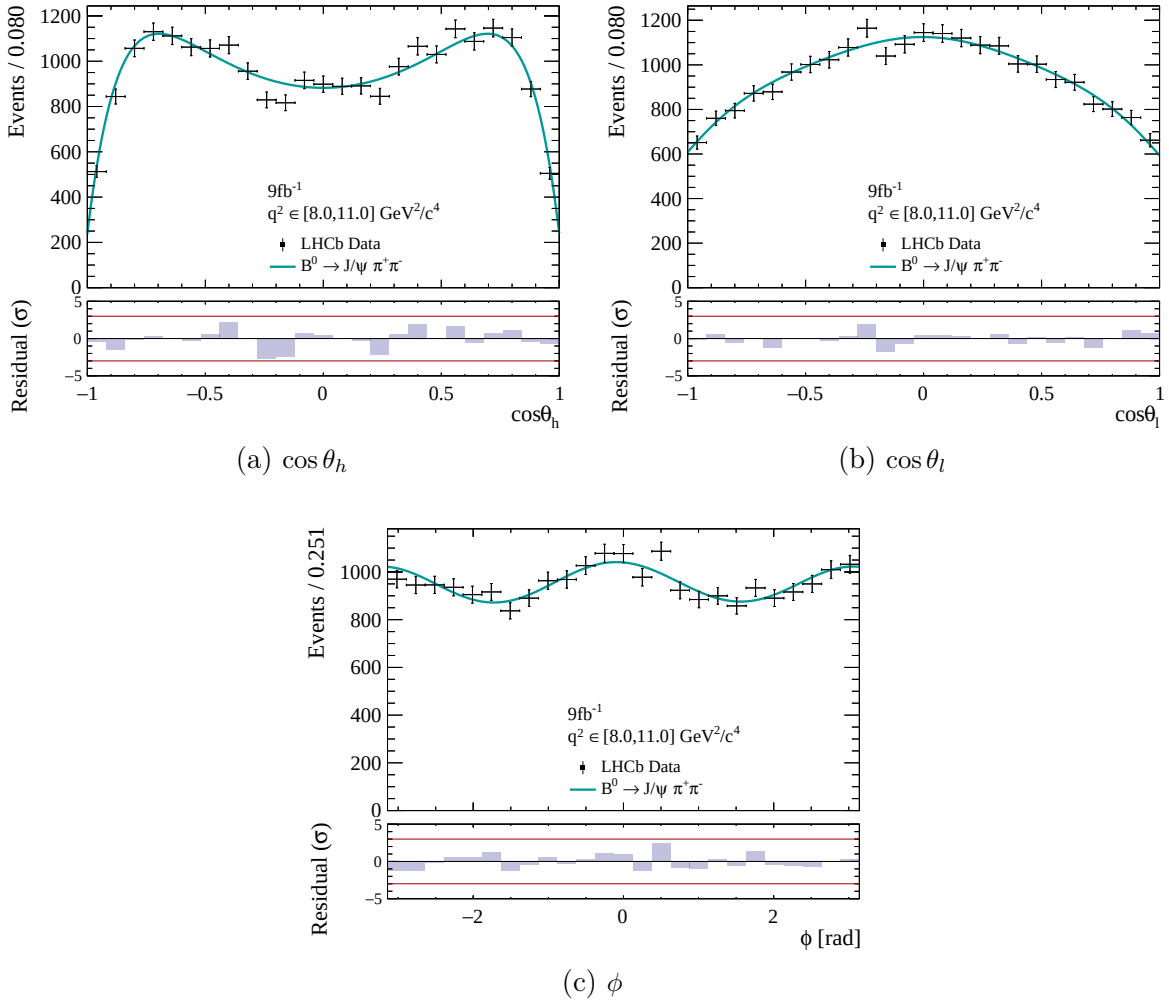
The fits to the angular distribution of the $B^0 \to J/\psi\pi^+\pi^-$ LHCb data are displayed in Figure 4.5.1. The observed values are compared to literature in Table 4.13. The extracted values align with the literature, reinforcing the robustness of the modelling strategy and providing confidence in the fit results to the rare q^2 samples. This validation is important for ensuring the reliability of the analysis and the accuracy of the results.

4.5.2 Modelling $B^0 \to \psi(2S)\pi^+\pi^-$ data

Fitting in the $\psi(2S)$ q^2 bin is a useful test of the robustness of the model. The $\psi(2S)$ has no previous measurements, the observed values of the angular observables are likely similar to those measured in the J/ψ sample, as is the case for the $B^0 \to J/\psi K^{*0}$ [53, 193]. The similarity arises as the $\psi(2S)$ and J/ψ are both $c\bar{c}$ resonances of the $\mu^+\mu^-$ system and share the same quantum numbers, differing only in their respective invariant mass. Therefore, a naive assumption is made in this analysis that the results should be consistent

Observable	Observed	Literature	Pull
F_L	0.588 ± 0.010	0.574 ± 0.029	0.45
S_3	0.026 ± 0.015	-0.021 ± 0.009	2.68
A_6	-0.016 ± 0.025	0.000 ± 0.000	-0.63
A_9	-0.015 ± 0.015	-0.036 ± 0.019	0.92
F_S	0.238 ± 0.029	0.222 ± 0.012	0.50

Table 4.13: Results of the fitting of $B^0 \to J/\psi\pi^+\pi^-$ data samples, showing observed values and the values derived from amplitudes in literature [53]. The literature value of F_S is not derived from amplitudes but is the reported fraction of $f_0(500)$ mesons in the dataset. The pull values compared the observed value with the literature value using their combined uncertainty.



(a) $\cos\theta_h$ (b) $\cos\theta_l$

(c) ϕ

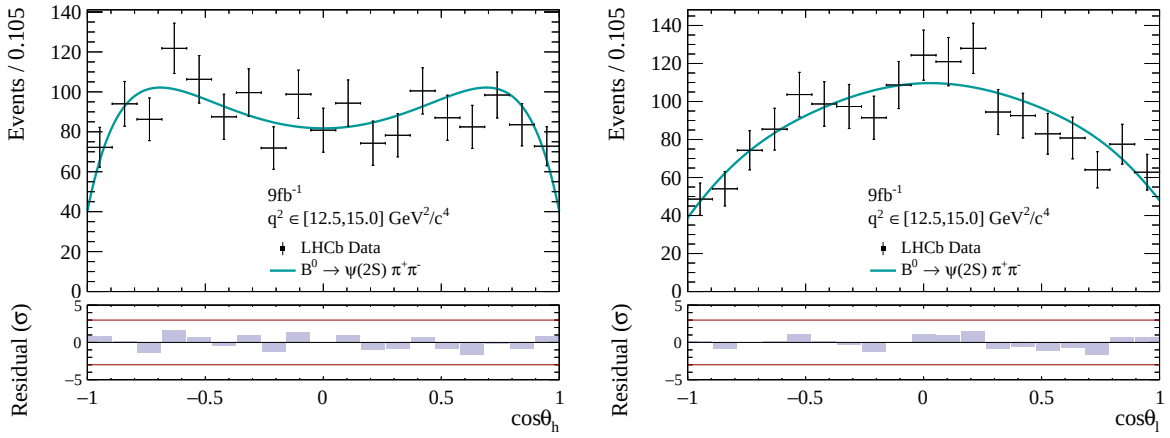
Figure 4.5.1: Angular fits to the background subtracted $B^0 \to J/\psi\pi^+\pi^-$ LHCb dataset.

with the observed values in the J/ψ q^2 bin.

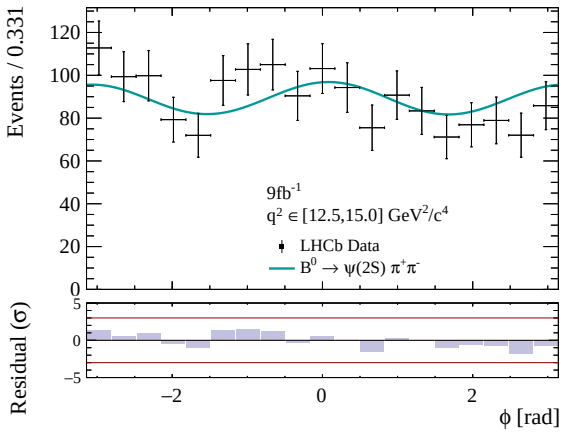
The fit to the angular distributions of the $B^0 \to \psi(2S)\pi^+\pi^-$ LHCb data are displayed in Figure 4.5.2. The observed values are given in Table 4.14. The results have a reasonable level of agreement with the measurements in the J/ψ bin.

Observable	Observed	J/ψ observation	Pull (in σ)
F_L	0.640 ± 0.037	0.588 ± 0.010	1.36
S_3	0.004 ± 0.059	0.026 ± 0.015	-0.36
A_6	0.117 ± 0.106	-0.016 ± 0.025	1.22
A_9	0.001 ± 0.058	-0.015 ± 0.015	0.26
F_S	0.339 ± 0.089	0.238 ± 0.029	1.08

Table 4.14: Measured angular observables from the fitting of $B^0 \to \psi(2S)\pi^+\pi^-$ data samples. The final column is the pull with the results from the measurement in the J/ψ q^2 bin, with the values given in Table 4.13.



(a) $\cos\theta_h$ (b) $\cos\theta_l$



(c) ϕ

Figure 4.5.2: Angular fits to the background subtracted $B^0 \to \psi(2S)\pi^+\pi^-$ LHCb dataset.

4.5.3 Modelling rare data

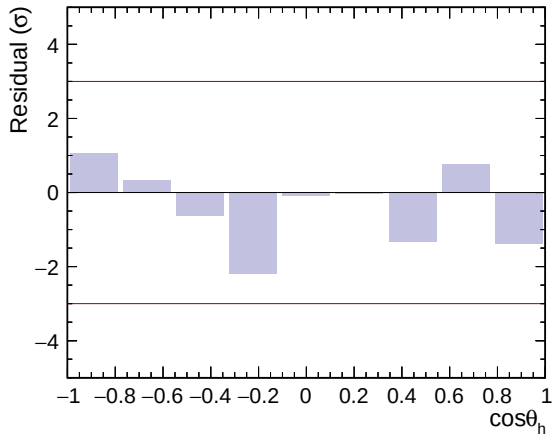
A blinded fit is performed on the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ LHCb data in the low-q^2 and high-q^2 regions. The modelling is performed according to the same strategy employed for the fits to the J/ψ and $\psi(2S)$ datasets. The pulls of the fits to the angular distribution of the

$B^0 \to \pi^+\pi^-\mu^+\mu^-$ LHCb data are displayed in Figure 4.5.3 for both the low-q^2 and high-q^2 regions. The distribution of the pulls indicate the model is an appropriate description of the data. The fits are not shown to retain the blinding of the measurement. The uncertainties of the blinded values are presented in Table 4.15.

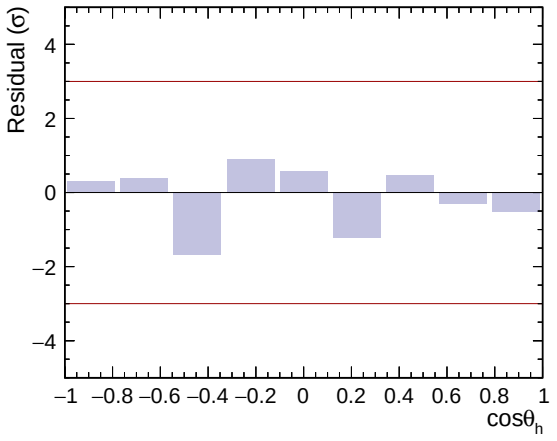
Only uncertainties are quoted, as the central values remain blinded. These uncertainties differ slightly from the sensitivities derived from the realistic pseudoexperiment studies; the origin of this difference is not yet fully understood and is under investigation as the analysis continues to be refined. An evaluation of the systematic uncertainties is presented in the next chapter, followed by a summary of the results in Chapter 6.

Observable	Low-q^2	High-q^2
F_L	0.171	0.097
S_3	0.332	0.186
A_6	0.278	0.362
A_9	0.535	0.199

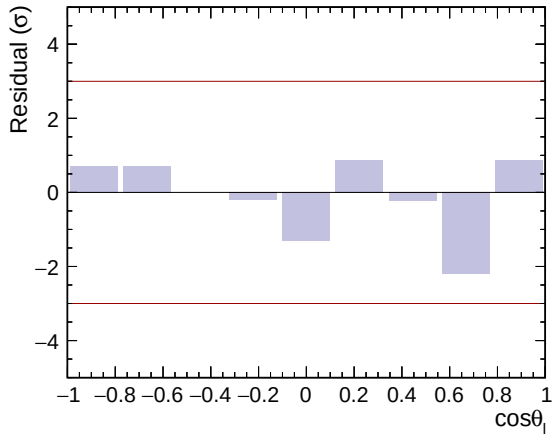
Table 4.15: Uncertainties of the angular observables from the blind fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ dataset, in both low-q^2 and high-q^2 bins.



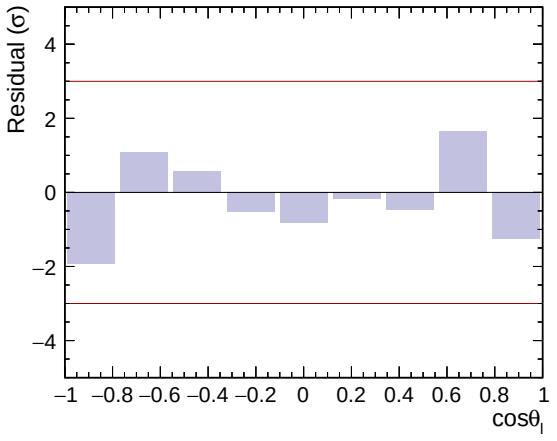
(a) $\cos\theta_h$, low-q^2



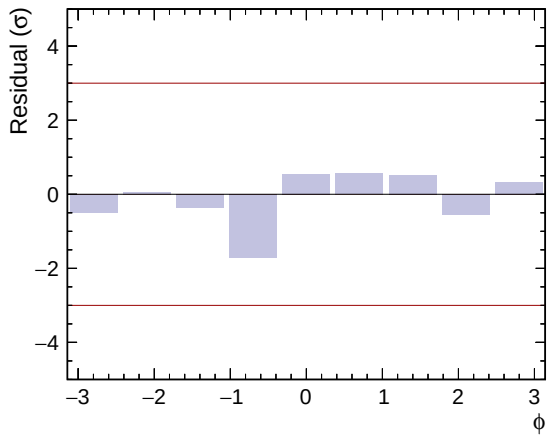
(b) $\cos\theta_h$, high-q^2



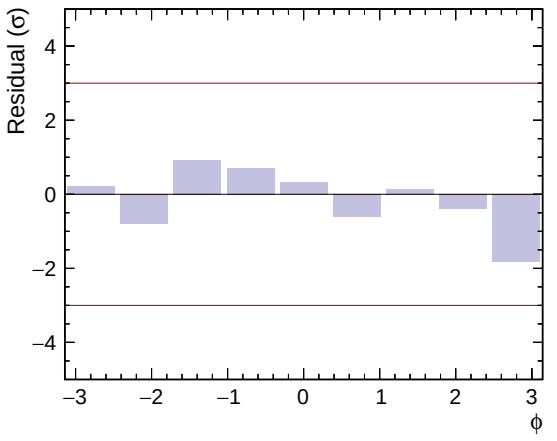
(c) $\cos\theta_l$, low-q^2



(d) $\cos\theta_l$, high-q^2



(e) ϕ, low-q^2



(f) ϕ, high-q^2

Figure 4.5.3: Residuals from fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ 9 fb^{-1} data samples in the low-q^2 and high-q^2 bins.

Chapter 5

Systematic Uncertainties

Thus far statistical uncertainties have been quoted. In addition to the statistical uncertainties one considers systematic uncertainties which arise due to inaccuracies in the measurement methods and assumptions which are made over the course of the physics analysis, which can introduce biases. While care is taken to minimise their effect they can lead to consistent over- or under-estimation of the true value of the observables and so are discussed in the following sections.

The measurement of the angular observables is statistically-limited and therefore the statistical uncertainty is expected to be the larger source of uncertainty, as the size of the sample increases one would see a reduction in the magnitude of the statistical but not necessarily the systematic uncertainty. It is crucial to handle both sources of uncertainty to ensure a reliable measurement. The main sources of uncertainty which are discussed result from the modelling of the invariant mass distribution of the B^0, the angular acceptance and the angular modelling.

Unless otherwise stated, pseudodatasets are generated using an alternative model, such as an alternative mass shape. The analysis is then performed on each of the pseudodatasets using the nominal strategy. To evaluate the size of the systematic uncertainty, the average shift of the measured value of the observable from the generation value is calculated using the following equation:

$$\sigma_{\text{syst}} = \langle x_{\text{fit}} \rangle - x_{\text{generation}}, \tag{5.1}$$

where $\langle x_{\text{fit}} \rangle$ represents the mean fitted value, and $x_{\text{generation}}$ is the value used to generate the pseudodatasets. This difference quantifies the bias introduced to the measurement when altering the mass shape.

The following sections detail the evaluation of systematic uncertainties, categorised according to the specific parts of the analysis to which they relate. The results of the evaluation of the systematic uncertainties is given in Section 5.5.1.

5.1 MC corrections

Multiple assumptions are made when correcting the MC simulation samples to the LHCb data. As the corrected samples are used to determine the mass shapes and, importantly, the angular acceptance, any choices have repercussions which pervade through to the final measurement. Here, two sources of systematic uncertainty relating to these corrections are discussed.

PID corrections

The PID corrections are derived from templates modelled using a Kernel Density Estimation (KDE). There is a potential for significant uncertainties in the parametrisation of the PID control sample, which could propagate through the analysis. To evaluate the impact of these uncertainties, an alternative template with wider kernels is utilised. This alternative template serves as a tool to assess the sensitivity of the analysis to the KDE parametrisation procedure.

A second source of uncertainty from the method of correcting the PID distributions comes from the statistical size of the calibrations samples. A total of five bootstrapped samples are used to evaluate alternative PID corrections.

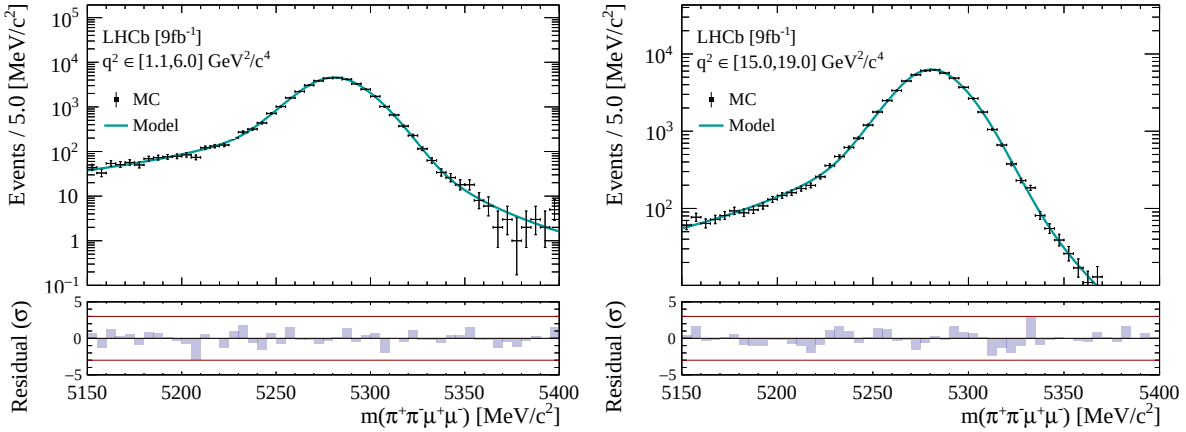
The PID corrections based on these alternative samples are then propagated to the simulation samples used for determining the efficiency. From the efficiency model, one large pseudo-dataset is generated with one million events. Each of the large samples is modelled and the bias of the observables is determined. the total bias from the five bootstrapped samples is the average of the five results. This average is summed in quadrature with the bias from the study using the wider kernel template to obtain the systematic uncertainty.

5.2 Mass modelling

The modelling of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ invariant mass distribution involves several assumptions that could introduce biases into the analysis. These assumptions primarily concern the shapes of the mass models and the treatment of the background decay modes. It is important to assess the impact of these assumptions on the final results.

Mass resolution

The resolution of the mass shapes is determined through fits to simulation datasets. A scaling factor is applied to the resolution widths during the fitting of the data to account for potential discrepancies between the simulated and real data samples. This scaling factor is determined from the fit to the $B^0 \to J/\psi\pi^+\pi^-$ dataset. To evaluate the impact



(a) Low-q^2 (b) High-q^2

Figure 5.2.1: Fit of the sum of a two-sided Hypatia distribution and a Gaussian to a simulation sample of the $B^0 \to \rho^0 \mu^+ \mu^-$ decay mode in the low-q^2 (left) and high-q^2 (right) bins.

of these discrepancies the resolution scale is varied, according to the uncertainty from the $B^0 \to J/\psi \pi^+ \pi^-$ fit, when generating the pseudo-datasets. These samples are then fit using the nominal mass model to assess the effects on the results.

Alternative signal mass shapes

The shape used in the nominal mass modelling to fit the signal component is a sum of an asymmetric double-sided Crystal Ball function (DCB) and a Gaussian function (G). An alternative to the nominal signal mass model is the sum of a two-sided version of a Hypatia distribution [194] and a Gaussian function. The Hypatia distribution is a robust choice for modelling peaks with asymmetric tails, featuring a hyperbolic core, similar to the Crystal Ball, and two flexible tails which which can be adjusted independently.

To evaluate the systematic uncertainty, the signal simulation samples are fitted using the Hypatia distribution. The parameters obtained form these fits are then used to generate pseudo-datasets. These samples are subsequently modelled using the nominal method. An example of a Hypatia fit to a simulation sample of the $B^0 \to \rho^0 \mu^+ \mu^-$ decay mode is shown in Figure 5.2.1.

Inclusion of background decay modes

In the nominal mass fit to the rare dataset, the partially reconstructed backgrounds, $B_s^0 \to (\phi \to \pi^+ \pi^- \pi^0) \mu^+ \mu^-$ and $B^0 \to (\eta' \to \rho^0 \gamma) \mu^+ \mu^-$, are not included because, as discussed in Section 3.6.2, the expected number of events for both decays is determined to be negligible. However, to evaluate the impact of neglecting these backgrounds on the analysis, the largest component, the $B^0 \to \eta' \mu^+ \mu^-$ decay mode, is included in the mass

model when generating the pseudo-datasets.

The number of events for the $B^0 \to \eta' \mu^+ \mu^-$ decay is extrapolated from the yield of the $B_s^0 \to J/\psi \eta'$ component in the fit to the $B^0 \to J/\psi \pi^+ \pi^-$ data. The model describing this decay in the fit is combined with the $B_s^0 \to J/\psi \phi$ component into the "Part. Reco." component, using Equation 4.11. The $B_s^0 \to J/\psi \eta'$ yield is determined using the relative fit fraction.

The samples are generated with the $B^0 \to \eta' \mu^+ \mu^-$ component included in the mass model and are then modelled using the nominal mass model, which does not include the $B^0 \to \eta' \mu^+ \mu^-$ component. The systematic uncertainty for each observable is determined as the standard deviation of the fitted values across the pseudodatasets.

sPlot fit fractions

After the mass modelling, signal candidates are isolated from the background using sWeights. In this process, the background components are summed into a single contribution to aid stability, as described in Section 4.1.5 and Equation 4.19. To evaluate the impact of this choice, the yields of each background component are varied according to the covariance matrix from the invariant mass fit.

A total of 1000 pseudodatasets are generated, each with a unique variation of the background yields. For each sample, the fitting is performed using the nominal method.

Factorisation of dimensions in sPlotting

The sPlot method requires the factorisation of the mass dimension from the angular dimensions. However, this assumption may not hold perfectly in practice, leading to the introduction of potential biases in the results. To evaluate the systematic uncertainty associated with this factorisation assumption, the peaking-background components in the generated pseudo-datasets are replaced with those derived from dedicated MC-simulated samples.

This is particularly important for the backgrounds which come from the misidentification or particles or partially reconstructed background, as the incorrectly identified particle will have a biased p_{T} and mass. If the mass and angles do not completely factorise, then biases may be introduced in the angular observables. The combinatorial background is not replaced as it is assumed that the variables in the combinatorial background, which consists of random tracks, factorise. Therefore, these samples capture more complex correlations between the mass and angular dimensions than the simplified pseudodatasets provide. By comparing the results from the more realistic MC-based samples to those from the nominal analysis, the impact of this assumption can be estimated.

5.3 Acceptance modelling

Bootstrapping

The MC simulation samples used to model the efficiency has a finite sample size, which could introduce uncertainties into the acceptance modelling. To quantify this systematic uncertainty, a bootstrap technique is employed. Specifically, the original simulation dataset is resampled with replacement to create a new sample. This resampling allows for some events to be duplicated, creating a random distribution of possible efficiency models which reflect inherent statistical variability.

For each resampled dataset, the efficiency is reparameterised, and a new pseudo-dataset is generated, with one million events, using this parametrisation. The default analysis method is then applied to each sample to extract the angular observables. This process is repeated 100 times to produce a distribution of results.

The systematic uncertainty evaluates the variation in the fitted uncertainties across each of the bootstrap samples. This uncertainty, therefore, captures the impact of the acceptance parameterisation. It is estimated as the variation in the fitted uncertainties across each of the samples:

$$\sigma_{\text{syst}} = \sqrt{\langle \sigma_{\text{fit}}^2 \rangle - \langle \sigma_{\text{fit}} \rangle^2}, \tag{5.2}$$

Orders in acceptance model

The choice of polynomial orders in the 5D efficiency model can introduce biases into the analysis. To assess the uncertainty associated with this choice, the orders of the polynomials used in the acceptance modelling are all increased by one. This increase can potentially capture features which are overlooked by the nominal model. However, increasing the orders excessively can lead to overfitting, where statistical fluctuations in the sample are incorrectly modelled. Conversely, decreasing the orders may result in insufficient modelling of the features.

For this altered configuration, the efficiency is reparameterised, and pseudodatasets are generated from this new efficiency model. The nominal analysis procedure, using the original polynomial orders, is then applied to each of the new datasets. By analysing the results obtained from this altered sets of orders, the bias is taken as the systematic uncertainty.

5.4 Angular modelling

Inclusion of S-wave interference

In Section 4.4.3 it was determined that with the statistics in the two rare q^2 bins there is no sensitivity to the F_S observable, therefore in the pseudoexperiments this observable is omitted from the simulations and from the modelling. To evaluate this assumption as a source of uncertainty a study is performed where the samples are generated with a value of F_S fixed to 0.222. This value is obtained from the amplitudes determined for the J/ψ decay mode [53]. The systematic uncertainty for each observable is determined as the standard deviation of the fitted values across the pseudodatasets.

Angular resolution

The angular resolution for each dimension is determined using simulation samples of the signal decay mode. The resolution is quantified as the difference between the generated angular distribution and the reconstructed angular distribution. Figure 5.4.1 shows example distributions of the resolution for the three angular dimensions.

The pseudodatasets are generated using the nominal method. A per-event smearing is applied to each angle independently, where the smearing values are sampled from the respective resolution distributions. This approach catches the uncertainties inherent in the resolution of the angular dimensions.

5.5 Consideration of B^0-$\overline{B}^0$ symmetry

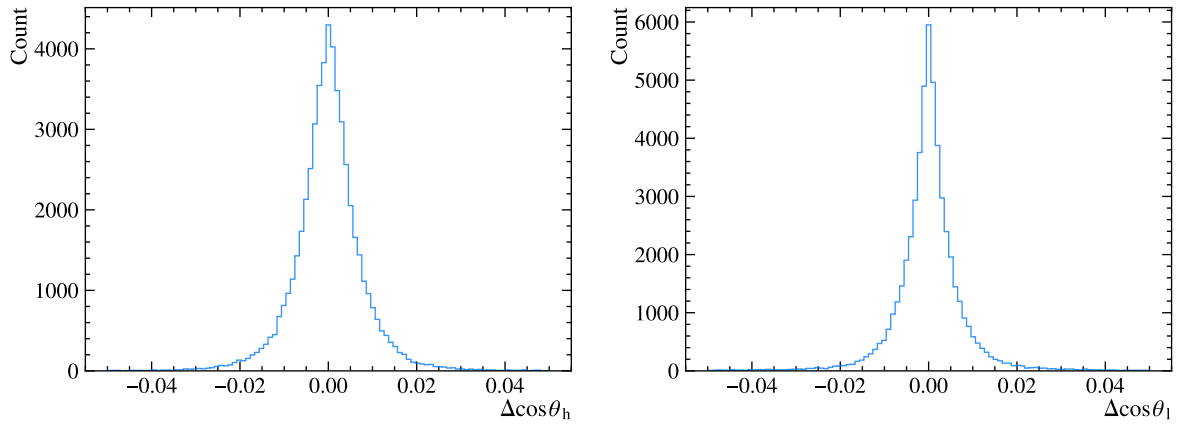
If the number of B^0 and $\overline{B}^0$ decays are not equal in the data samples then the terms in the angular distribution no longer correspond to pure CP averages or asymmetries. Through various physical processes the observables experience dilutions which are discussed in the following sections.

Production asymmetry

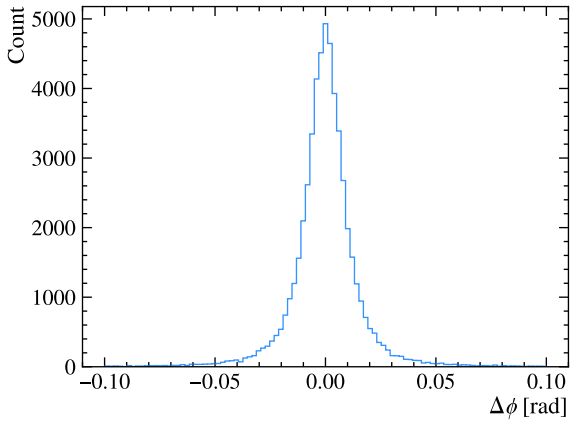
At the LHC, the production of B^0 and $\overline{B}^0$ mesons is known to be asymmetric due to the non-charge symmetric initial state in pp-collisions. This production asymmetry, A_{prod}, is defined as:

$$A_{prod} = \frac{N(\overline{B}^0) - N(B^0)}{N(\overline{B}^0) + N(B^0)}, \tag{5.3}$$

where $N(\overline{B}^0)$ and $N(B^0)$ represent the numbers of $\overline{B}^0$ and B^0 mesons produced, respectively. The value of A_{prod} at the LHCb detector has been measured to be



(a) $\cos\theta_h$ (b) $\cos\theta_l$



(c) ϕ

Figure 5.4.1: Distributions of the resolution of the three angular dimensions. The resolution is the difference between value with which the event in the sample was generated and the value following the application of selection requirements. The distributions are from simulation samples of the $B^0 \to \rho^0 \mu^+ \mu^-$ decay mode, in the high-q^2 bin.

$(-0.35 \pm 0.76 \pm 0.28)\%$ [195]. Although this asymmetry is small, it can impact the measured *CP*-asymmetric and *CP*-averaged observables.

The relationship between the measured observables ($I_i^{meas.}$) and the true observables (I_i) is given by:

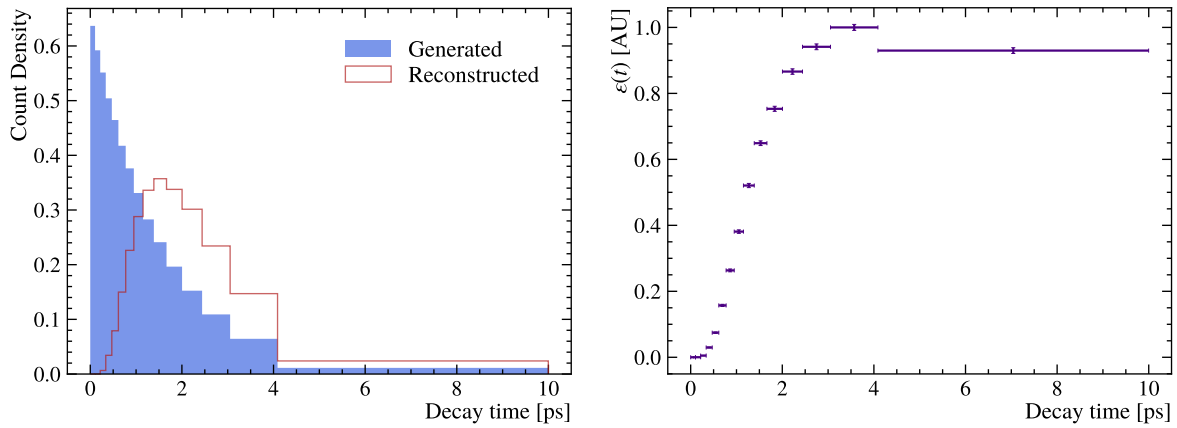
$$A_i^{meas.} = A_i - S_i(\kappa A_{prod}), \quad S_i^{meas.} = S_i - A_i(\kappa A_{prod}). \tag{5.4}$$

where κ is a dilution factor that accounts for B^0-$\overline{B}^0$ mixing. The *CP*-asymmetric and *CP*-averaged observables are bounded by $\mid A_i \mid \leq 1$ and $\mid S_i \mid \leq 1$. The dilution factor κ is calculated as:

$$\kappa = \frac{\int_0^\infty \epsilon(t) e^{-\Gamma t} \cos(\Delta m_d t) dt}{\int_0^\infty \epsilon(t) e^{-\Gamma t} dt}, \tag{5.5}$$

where $\Gamma = 1/\tau_d = 1/1.519\,\text{ps}^{-1}$ is the decay rate, $\Delta m_d = 0.510\,\text{ps}^{-1}$ is the B^0-$\overline{B}^0$ mixing frequency, and $\epsilon(t)$ is the decay time-dependent efficiency. This efficiency is determined using rare phase-space MC samples, the distribution of $\epsilon(t)$ is given in Figure 5.5.1.

From this analysis, the value of κ is determined to be 0.619 ± 0.001. Consequently, the resulting scale of the dilution effect on the observables is $\kappa A_{prod} = (-0.217 \pm 0.001)\%$. Given that this dilution is much smaller than the statistical uncertainties of the analysis, it can be considered negligible.



(a) Selected and generated simulation samples

(b) $\epsilon(t)$ from ratio of samples

Figure 5.5.1: Figure (a) shows the rare phase-space simulation samples of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, at generation and after reconstruction and the selection is applied. Figure (b) shows the decay time dependent efficiency ($\epsilon(t)$). The efficiency is determined by taking a ratio of the reconstructed sample, with selection requirements applied, to the generated sample. The online selection procedure introduces a significant lifetime bias. The behaviour of the efficiency demonstrates that the selection criteria is biased towards B-mesons with longer lifetimes. The binning is chosen to contain equal statistics for the generated sample.

Decay width difference

The final state does not allow to determine the decay flavour of the B^0 meson. The oscillation of B^0-$\overline{B}^0$ mesons leads to the introduction of time dependent decay amplitudes, which could produce a bias on the measured observables. In the presence of a non-zero difference between the decay widths of the B^0 and $\overline{B}^0$ mesons, the sum of the time-dependent decay rates can be expressed as [50]:

$$J_i(t) + \tilde{J}_i(t) = e^{-\Gamma t}[(J_i + \tilde{J}_i)\cosh(y\Gamma t) - h_i \sinh(y\Gamma t)], \tag{5.6}$$

where $J_i(t)$ and $\tilde{J}_i(t)$ represent the time-dependent decay rates for the B^0 and $\overline{B}^0$ mesons, respectively. The term $y = \Delta\Gamma/2\Gamma$ quantifies the relative difference between the decay widths, where $\Delta\Gamma = \Gamma_L - \Gamma_H$ [51] is the difference between the decay widths of the *CP*-even (Γ_L) and *CP*-odd (Γ_H) components. For the B^0 mesons, $2y = \Delta\Gamma_d/\Gamma_d = 0.001 \pm 0.010$ [196].

The coefficient h_i is a new angular coefficient of the $\sinh\Gamma t$ component in Equation 5.6, which emerges in relation to the time-dependent angular distribution. This coefficient leads to an effect on the measured angular observables. Although the h_i coefficients are very difficult to measure directly, as y is very small, the contribution can be expressed as:

$$S_i^{meas.} = S_i - h_i D. \tag{5.7}$$

Where S_i represents the angular observables and D is the dilution factor, defined as:

$$D = \frac{\int_0^\infty \epsilon(t) e^{-\Gamma t} \cosh(y\Gamma t) dt}{\int_0^\infty \epsilon(t) e^{-\Gamma t} \sinh(y\Gamma t) dt}. \tag{5.8}$$

Here, $\epsilon(t)$ is the decay time-dependent efficiency, as discussed in the previous section. As discussed in Section 1.2.3, the decay rates $J_i(t)$ and $\tilde{J}_i$ are related to the transversity amplitudes $\mathcal{A}_0$, $\mathcal{A}_\parallel$, and $\mathcal{A}_\perp$. The amplitude terms $\mathcal{A}_0$ and $\mathcal{A}_\parallel$ are *CP* even, while $\mathcal{A}_\perp$ is *CP* odd. These states are modulated by the decay widths Γ_L and Γ_H, respectively.

In this analysis, the value of the dilution factor D is determined to be -0.005 ± 0.004. Given the small magnitude of the time dilution factor, the effect of the decay width asymmetry between B^0 and $\overline{B}^0$ can be considered negligible for the purposes of this measurement.

5.5.1 Summary of the systematic uncertainties

The evaluated systematic uncertainties for each of the angular observables is given in Table 5.1 for the low-q^2 measurements and in Table 5.2 for the high-q^2 measurements. The uncertainties are given as shifts of the average fitted value from the generated value of each observable, unless otherwise stated in the previous section. The absolute (total) uncertainty is determined by summing the individual estimated uncertainties in quadrature.

No source of systematic uncertainty is considered to contribute considerably more to the absolute systematic uncertainty, than any other. The systematic uncertainties are consistently smaller than the statistical uncertainties, given in Table 4.15. The systematic uncertainties for the F_L observable are determined to be 59% and 52% the size of the statistical uncertainties, in the low-q^2 and high-q^2 bins, respectively. The systematic uncertainties for the S_3, A_6, and A_9 observables are all below 29%.

The sizes of the systematic uncertainties relative to the statistical uncertainties demon-

strates that the analysis is dominated by the statistical uncertainty. Therefore, in future measurements of these observables using the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode, the precision of the measurements will be improved initially with increased sample sizes. With increased sample sizes some of the systematic uncertainty will be reduced, for instance the "Orders in acceptance model" is necessary as statistical fluctuations affect the accuracy of the modelling of the acceptance affects. Other sources will not be applicable, such as "Inclusion of background shapes" and "Inclusion of S-wave interference". As the sample sizes increase the features which are not included in the nominal modelling can be introduced. New sources of systematic uncertainty will be pertinent to evaluate as the challenges in the analysis evolves. By the conclusion of Run III of the LHC, there should be a large enough sample size that the statistical and systematic uncertainties are of a similar size.

Source of systematic uncertainty	F_L	S_3	A_6	A_9
Weighting				
PID	0.02	0.00	0.01	0.00
Mass modelling				
Mass resolution	0.03	-0.01	-0.00	-0.00
Alternative signal mass shapes	0.04	-0.01	-0.02	0.00
Inclusion of background shapes	0.01	-0.01	0.04	0.01
sPlot fit fractions	0.05	-0.01	0.01	0.01
Factorisation of dimensions in sPlotting	0.00	-0.00	0.06	0.01
Acceptance parameterisation				
Bootstrapping	0.02	-0.00	0.01	0.00
Orders in acceptance model	0.03	-0.01	-0.01	0.00
Angular modelling				
Inclusion of S-wave interference	0.03	-0.01	-0.00	-0.02
Angular resolution	0.05	0.01	0.01	0.01
Total				
$\sum$ in quadrature	0.09	0.03	0.08	0.03

Table 5.1: Summary of the systematic uncertainties evaluated for the measurement of the angular observable in the low-q^2 range. The final row shows the absolute systematic uncertainty, which is the sum, in quadrature, of the individual uncertainties.

Source of systematic uncertainty	F_L	S_3	A_6	A_9
Weighting				
PID	0.00	0.00	0.01	0.00
Mass modelling				
Mass resolution	0.01	0.00	0.01	-0.01
Alternative signal mass shapes	0.01	0.01	0.01	0.00
Inclusion of background shapes	0.02	0.01	0.00	0.02
sPlot fit fractions	0.01	0.00	-0.00	0.00
Factorisation of dimensions in sPlotting	0.01	0.01	0.01	0.01
Acceptance parameterisation				
Bootstrapping	0.00	-0.01	0.01	0.00
Orders in acceptance model	0.03	-0.04	0.01	0.01
Angular modelling				
Inclusion of S-wave interference	0.02	0.03	-0.02	-0.00
Angular resolution	0.02	0.00	0.02	0.03
Total				
$\sum$ in quadrature	0.05	0.05	0.04	0.04

Table 5.2: Summary of the systematic uncertainties evaluated for the measurement of the angular observable in the high-q^2 range. The final row shows the absolute systematic uncertainty, which is the sum, in quadrature, of the individual uncertainties.

Chapter 6

Summary and prospects

The Standard Model is a robust theory which successfully explains much of the current understanding of particle physics. However, there are phenomena which the SM has proved inadequate to explain, such as the imbalance of matter and anti-matter in the universe, or the origin of neutrino mass. These shortcomings indicate that the SM is an approximation of a more complete theory and either an expansion of the theory or a completely new approach is required. Precision measurements, such as the determination of angular observables in the $b \to d\ell^+\ell^-$ FCNC processes, could indicate potential New Physics contributions which give rise to deviations.

The analysis presented in this thesis is the angular analysis of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode in the low-q^2 bin of $[1.1, 6.0]\,\mathrm{GeV}^2/c^4$, and the high-$q^2$ bin of $[15.0, 19.0]\,\mathrm{GeV}^2/c^4$. A method for measuring the angular observables F_L, S_3, A_6, and A_9, has been set out, using the Run I and Run II datasets collected by the LHCb experiment. Requirements on the dataset selecting decays with a $\pi^+\pi^-\mu^+\mu^-$ final state have been detailed, followed by the modelling of the invariant mass distributions in several bins of q^2. Further, the sPlot method has been used to statistically separate the signal from the background in the dataset, the acceptance has been parameterised, and the feasibility of performing this measurement has been demonstrated.

The angular observables have been measured in simulation samples of the $B^0 \to J/\psi\pi^+\pi^-$ and $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay modes, where the observables were mostly in agreement with the expected values. Any tensions between the observed values and expected values are understood. Further, pseudoexperiments have been performed which stress-test the measurements by varying the generated values of the observables, showing the model is robust. The fit strategy has been developed and validated on data in the J/ψ and $\psi(2S)$ q^2 bins, where the results were in agreement with the literature.

The analysis remains blind, pending review by the LHCb collaboration. However, the blind fits have been performed to the data in the rare-q^2 bins and the statistical and systematic uncertainties have been evaluated. The statistical uncertainties have been

shown to dominate the overall uncertainties. The uncertainties of the observables, for the low-q^2 bin, in data are:

$$
\begin{aligned}
F_L &= \pm 0.17\ (\text{stat}) \pm 0.09\ (\text{syst}) \\
S_3 &= \pm 0.33\ (\text{stat}) \pm 0.03\ (\text{syst}) \\
A_6 &= \pm 0.28\ (\text{stat}) \pm 0.08\ (\text{syst}) \\
A_9 &= \pm 0.54\ (\text{stat}) \pm 0.03\ (\text{syst}),
\end{aligned}
$$

where "stat" is the statistical uncertainty and "syst" is the systematic uncertainty. The uncertainties for the observables in the high-q^2 bin are:

$$
\begin{aligned}
F_L &= \pm 0.10\ (\text{stat}) \pm 0.05\ (\text{syst}) \\
S_3 &= \pm 0.19\ (\text{stat}) \pm 0.05\ (\text{syst}) \\
A_6 &= \pm 0.36\ (\text{stat}) \pm 0.04\ (\text{syst}) \\
A_9 &= \pm 0.20\ (\text{stat}) \pm 0.04\ (\text{syst}).
\end{aligned}
$$

At the time of writing, the LHC accelerator is up and running for the third period of data-taking. Major upgrades have been undertaken to the LHCb detector [154] in preparation for Run III, preparing the detector for to take $50\,\text{fb}^{-1}$ of data. This totals a factor five increase in instantaneous luminosity over the combined Run I and Run II. The sensitivity of the angular observables measured using fits to the rare q^2 bins are limited due to the small samples sizes available. The significant increase in data collected during Run III and beyond will increase the precision of the measurements of these observables, which will provide greater insight into the presence of New Physics effects from $b \to d\ell^+\ell^-$ transitions. The increased statistical power will allow the measurement of observables to which the current samples have no sensitivity, such as the observables which dropped out following the integration of the differential decay rate. Additionally, with the increased sample sizes, the observables related to the S-wave spin states, such as F_S, can be determined. Further, sensitivity to the higher spin states will eventually make it possible to consider measuring the D-wave observables.

Appendix A

Angular Basis

The decay mode $B^0 \to \pi^+\pi^-\mu^+\mu^-$ is a pseudo-scalar to vector-vector decay. The angular distribution is a combination of observables which are combinations of transversity amplitudes. The physics content of the decay is thus expressed in three angles, $\vec{\Omega} = (\theta_h, \theta_l, \phi)$, which are derived as follows.

The θ angles

The $\pi^+\pi^-$ and $\mu^+\mu^-$ pairs are both associated with a polar angle. This polar angle is the direction between the particle in the rest frame of the diparticle system and the direction of the diparticle system in the rest frame of the b-hadron. The polar angles are denoted by θ_h for the dipion system, and θ_l for the muons.

The momentum vector of particle x in the rest frame of particle a is $\vec{p}_x^{\,a}$. The sum and difference of the momentum of particles x and y is thus $\vec{p}_{xy}^{\,a} = \vec{p}_x^{\,a} + \vec{p}_y^{\,a}$ and $\vec{q}_{xy}^{\,a} = \vec{p}_x^{\,a} - \vec{p}_y^{\,a}$, respectively. The angle between the vector which defines the direction of the positive muon in the dimuon rest frame and the direction of the dimuon in the B^0 rest frame is the angle θ_l. It is defined as

$$\cos\theta_l = \frac{\vec{p}_{\mu^+}^{\,\mu^+\mu^-} \cdot \vec{p}_{\mu^+\mu^-}^{\,B^0}}{|\,\vec{p}_{\mu^+}^{\,\mu^+\mu^-}\,||\,\vec{p}_{\mu^+\mu^-}^{\,B^0}\,|} = -\frac{\vec{p}_{\mu^-}^{\,\mu^+\mu^-} \cdot \vec{p}_{\mu^+\mu^-}^{\,B^0}}{|\,\vec{p}_{\mu^-}^{\,\mu^+\mu^-}\,||\,\vec{p}_{\mu^+\mu^-}^{\,B^0}\,|}. \tag{A.1}$$

To transform between the B^0 and $\overline{B}^0$ one swaps a particle for its anti-particle by reversing the momentum vectors. The definition of the equivalent angle in the dipion system is the same when the momentum vectors are swapped for those of the pions.

The ϕ angles

The azimuthal angle is a rotation of a plane containing two particles around the axis of the direction of the diparticle system in the b-hadron rest frame. The angle ϕ is related

to the azimuthal angles of the dipion and dimuon systems in their own planes, it is the angles between the two planes in the b-hadron rest frame.

The unit normal vector to the plane containing x and y in the rest frame of particle a is

$$\hat{n}^a_{xy} = \frac{\vec{p}_x^{\,a} \times \vec{p}_y^{\,a}}{|\ \vec{p}_x^{\,a} \times \vec{p}_y^{\,a}\ |}.$$

The angle ϕ is defined, in the B^0 case, as

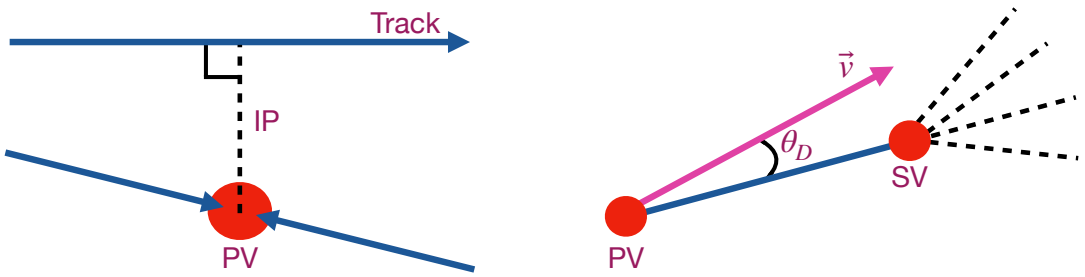
$$\cos\phi = \hat{n}^{B^0}_{\mu^+\mu^-} \cdot \hat{n}^{B^0}_{\pi^+\pi^-}, \quad \sin\phi = (\hat{n}^{B^0}_{\mu^+\mu^-} \times \hat{n}^{B^0}_{\pi^+\pi^-}) \cdot \frac{\vec{p}^{\,B^0}_{\pi^+\pi^-}}{|\ \vec{p}^{\,B^0}_{\pi^+\pi^-}\ |}. \tag{A.2}$$

In the case of $\overline{B}^0$ decays the momentum vectors are reversed which changes the signs of both the cosine and the sin, thus the ϕ is unchanged.

Appendix B

Variable definitions

Several variables are referred to throughout this thesis which are used to describe the geometry of a decay in the LHCb detector and the particle identification of the constituents of said decays. The definitions of these variables are provided in Tables B.1 and B.2.



(a) Impact Parameter

(b) Direction angle

Figure B.0.1: Illustrations of variable definitions. Figure (a) shows the Impact parameter (IP) between a track and a vertex. The value is given by the length of the black dashed line. Figure (b) shows the direction angle (θ_D). The angle is given between the line connecting the PV and SV, and the momentum vector. The momentum vector corresponds to the tracks which contribute to the SV.

Variables	**Description**
Mass, $m(X)$	The invariant mass of X. Where X represents a particle, or combination of particles.
Momentum, p	The magnitude of the total momentum of a particle.
Transverse momentum, p_{T}	The component of the total momentum perpendicular to the proton beam (z) axis.
Pseudo-rapidity, η	The position of a particle in the LHCb detector relative to the proton beam axis.
Decay time, τ	The time taken for a particle to travel from the production vertex to the decay vertex.
Primary vertex, PV	Location of the pp-collision vertex.
Secondary vertex, SV	Location at which the parent particle in a decay chain decays.
Vertex quality, χ^2_{vtx}/ndf	The minimised χ^2 per degree of freedom, χ^2/N_{DOF}, determined in the fit to the location of the vertex.
Impact Parameter, IP	The shortest distance between a track and a vertex location. Illustrated in Figure B.0.1a.
χ^2_{IP}	The difference of a vertex's χ^2/N_{DOF}, with and without a track included in the fit.
Flight distance, FD	Distance between the origin and decay vertex of a particle.
χ^2_{FD}	Significance of the FD to a combination of particles. The χ^2 associated with $\vec{d} = \vec{v}_1 - \vec{v}_2$, where $\vec{d}$ is the difference in the position of two vertices, and $\vec{v}_1$ and $\vec{v}_2$ are the postilions of the first and second vertices, respectively.
Track quality, χ^2_{trk}	The minimised χ^2/N_{DOF} from the fit to the hits of the tracks.
Ghost track probability, TRGHP	The probability that a track is a "ghost" track, an incorrect combination of tracking hits. Determined using an MVA trained to distinguish between signal tracks from ghost tracks.
Isolation	The probability derived from the minimum vertex χ^2 from replacing a track of the signal candidate with all the tracks which fall in a cone around the signal track.

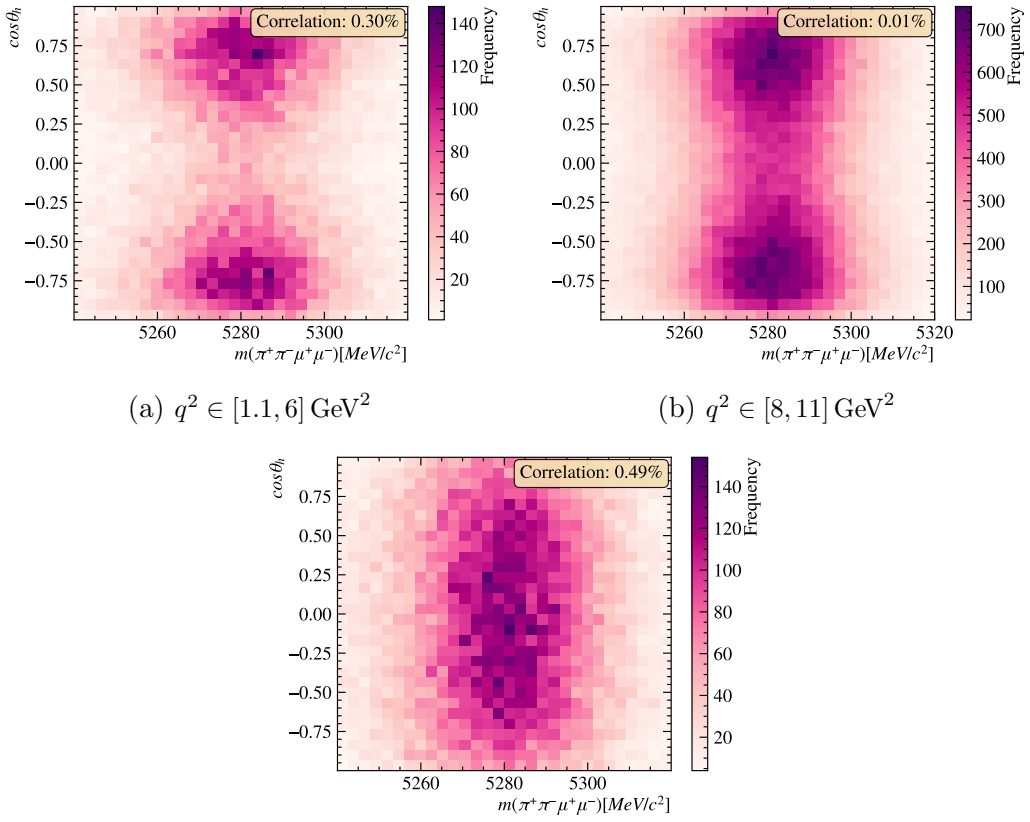
Table B.1: Geometric and PID variable definitions (set1).

Variables	**Description**
Distance of closest approach, DOCA	The shortest distance between two tracks.
Direction angle, θ_D	The pointing angle between a particle's momentum vector and the vector between its decay vertex and the PV. Signal decays point to the PV, giving a small angle and $\cos\theta_D$ close to 1; backgrounds may not point to the PV and have smaller $\cos\theta_D$. Illustrated in Figure B.0.1b.
nTracks	The number of reconstructed tracks.
nSPDHits	The number of hits in the SPD detector.
ProbNN	The probability from the output of an MVA which identifies particle species using global PID and tracking information.
$P(x \mid y)$	Probability of particle x being of type y
hasMuon	Boolean variable identifying if the track is a muon by requiring hits in the muon subdetectors match reconstructed muon tracks [155].
$min(p_{\mathrm{T}})_{hadron}$	The minimum p_{T} of the two final-state hadrons. A similar variable is defined for the dilepton system.
$min(\chi^2_{\mathrm{IP}})_{hadron}$	The minimum χ^2_{IP} of the two final-state hadrons. A similar variable is defined for the dilepton system.

Table B.2: Geometric and PID variable definitions (set2).

Appendix C

Correlations between angles and $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass

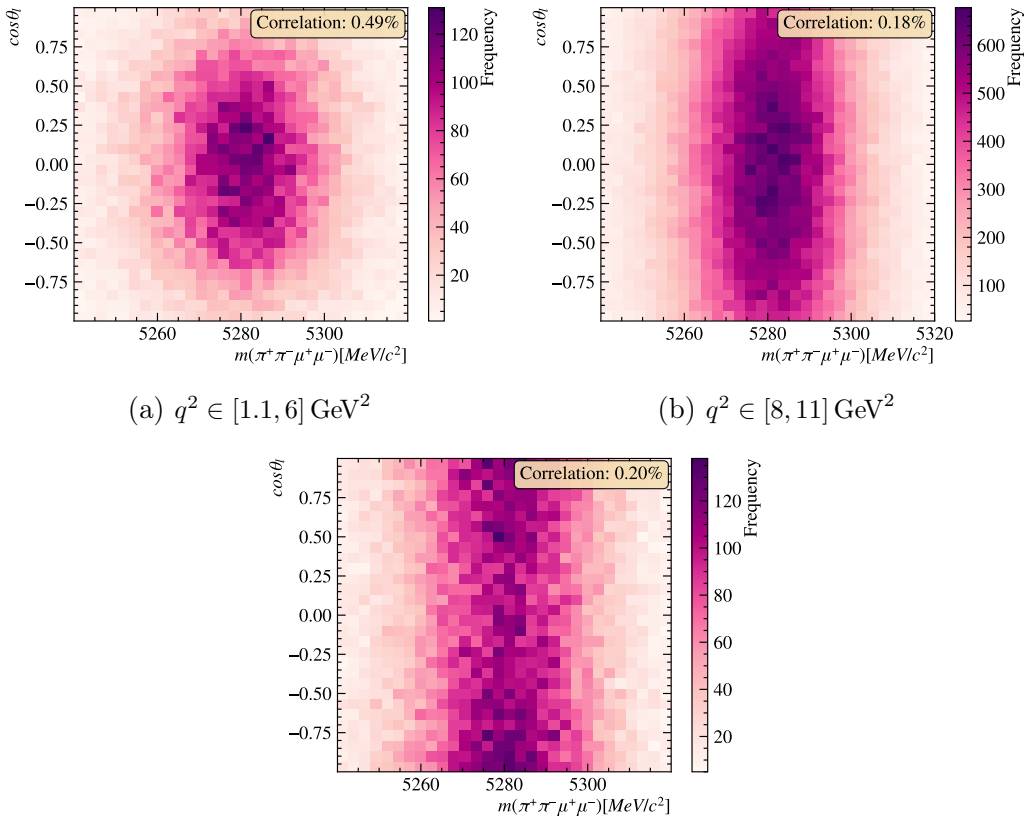


(a) $q^2 \in [1.1, 6]\,\text{GeV}^2$

(b) $q^2 \in [8, 11]\,\text{GeV}^2$

(c) $q^2 \in [15, 19]\,\text{GeV}^2$

Figure C.0.1: Correlation between $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass and $\cos\theta_h$ in bins of q^2.

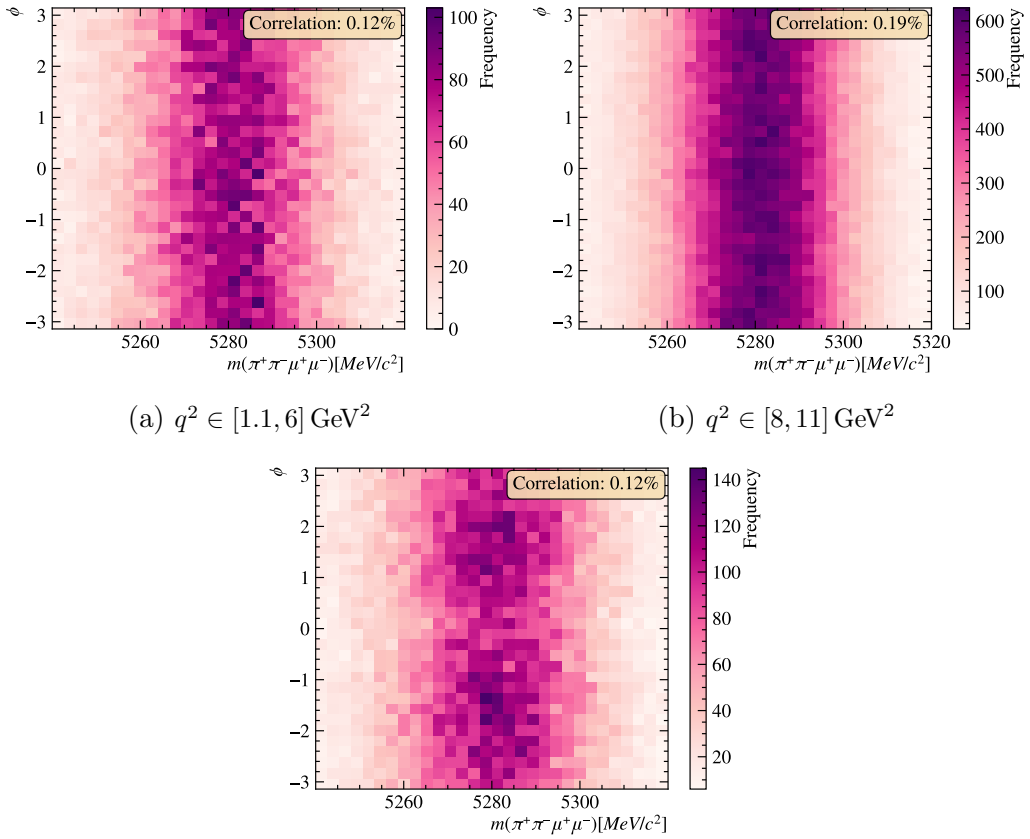


(a) $q^2 \in [1.1, 6]\,\mathrm{GeV}^2$

(b) $q^2 \in [8, 11]\,\mathrm{GeV}^2$

(c) $q^2 \in [15, 19]\,\mathrm{GeV}^2$

Figure C.0.2: Correlation between $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass and $\cos\theta_l$ in bins of q^2.

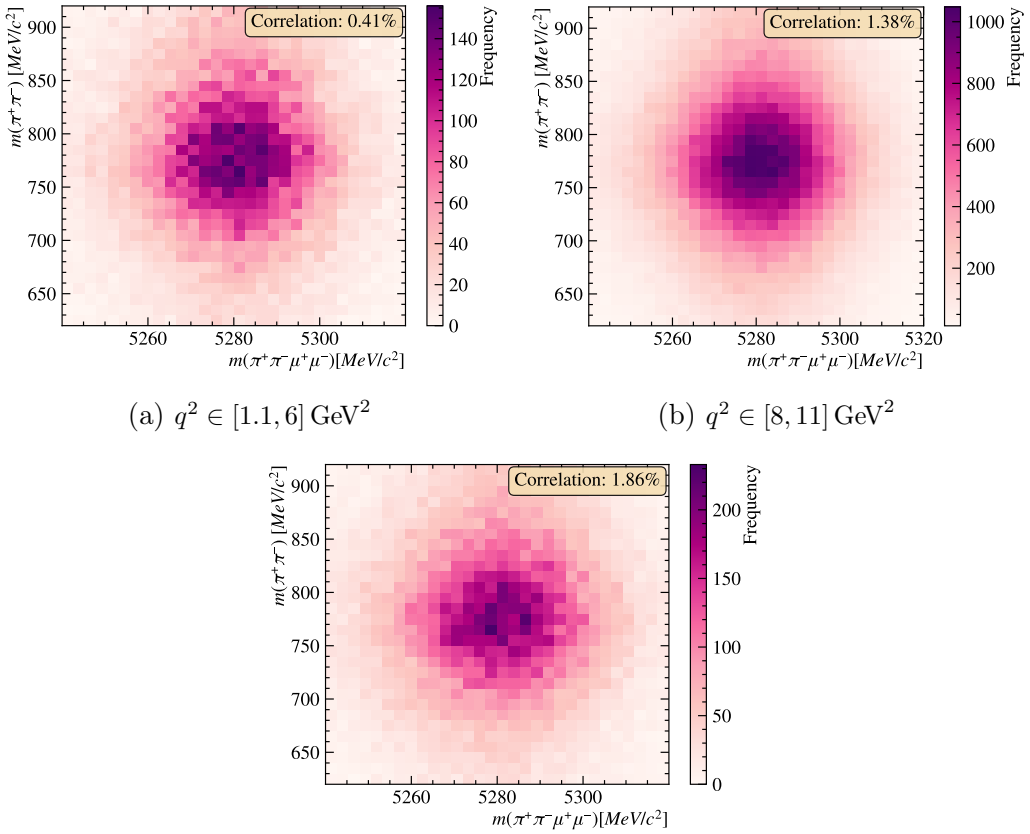


(a) $q^2 \in [1.1, 6]\,\mathrm{GeV}^2$

(b) $q^2 \in [8, 11]\,\mathrm{GeV}^2$

(c) $q^2 \in [15, 19]\,\mathrm{GeV}^2$

Figure C.0.3: Correlation between $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass and ϕ in bins of q^2.



(a) $q^2 \in [1.1, 6]\,\mathrm{GeV}^2$

(b) $q^2 \in [8, 11]\,\mathrm{GeV}^2$

(c) $q^2 \in [15, 19]\,\mathrm{GeV}^2$

Figure C.0.4: Correlation between $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass and $m(\pi^+\pi^-)$ in bins of q^2.

Appendix D

Acceptance: choice of k' and q' values

In constructing the angular acceptance functions described in Section 4.3, the full five-dimensional acceptance is reduced to three dimensions by fixing the values of k' and q'. This appendix provides comparisons of the angular acceptance function when different fixed values of k' and q' are used in the reduction. These include:

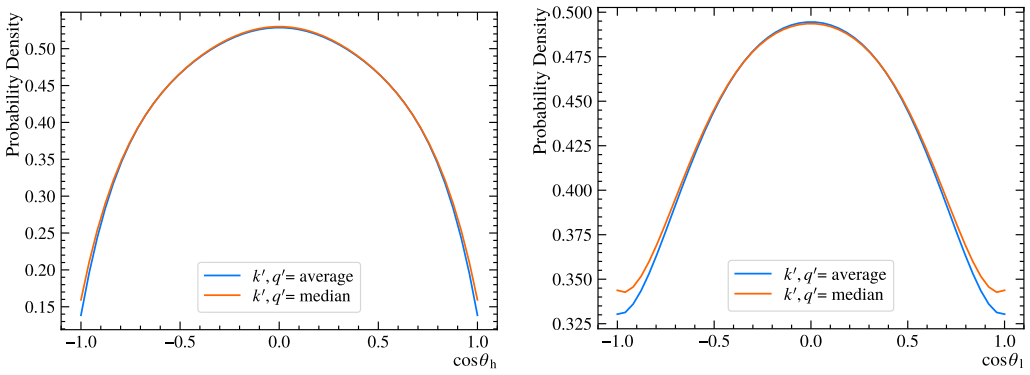
- The average values from the sample used for acceptance modelling,
- The median k' and q' value in the sample,
- The resonance mass in the case of J/ψ- and $\psi(2S)$-q^2 bins,
- And values corresponding to the edges of the k' bin.

D.1 Comparison of acceptance with different fixed values

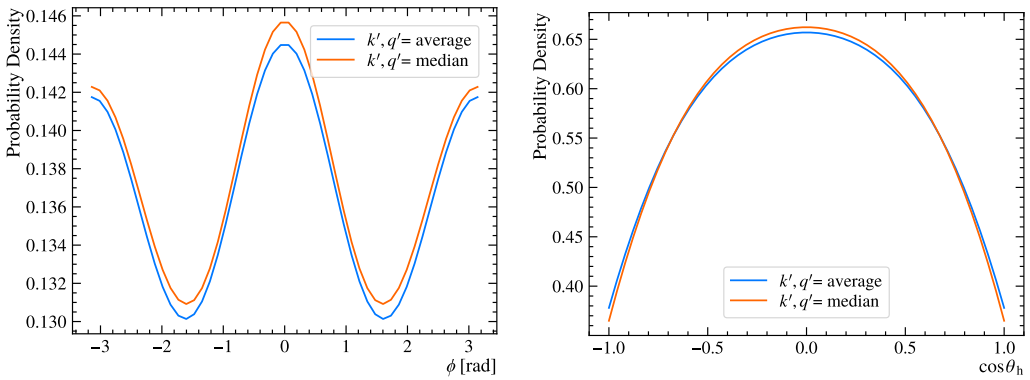
Figures D.2.1 and D.2.2 illustrate how the angular acceptance varies with different fixed k' and q' values. The scenarios shown include the sample mean and median, as well as the invariant mass corresponding to the $c\bar{c}$ resonances in specific q^2 bins.

D.2 Acceptance at k' bin edges

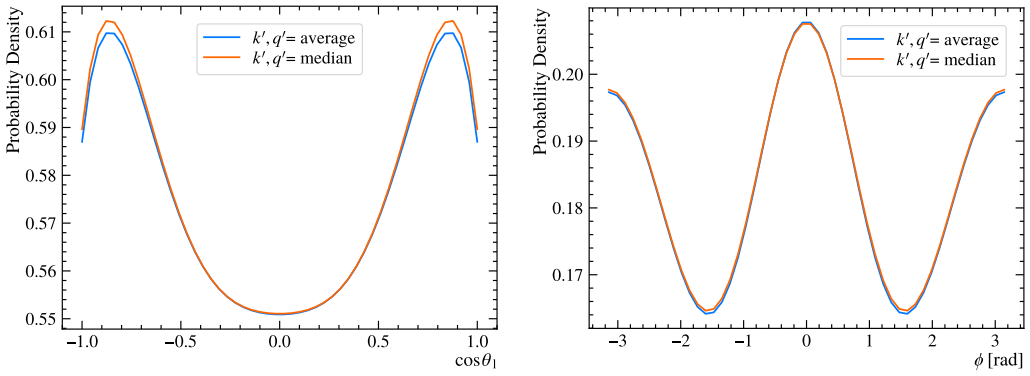
Figures D.2.3 and D.2.4 show the acceptance function when k' is fixed at extreme values, including the bin edges. A more pronounced variation is observed in this case, supporting the use of average values to mitigate edge effects in the current fit. This suggests that while the edge behaviour is non-negligible, the central values provide a more stable choice. A more detailed treatment of this variation may be warranted in future studies to incorporate it as a systematic uncertainty.



(a) $\cos\theta_h$, low-q^2 (b) $\cos\theta_l$, low-q^2

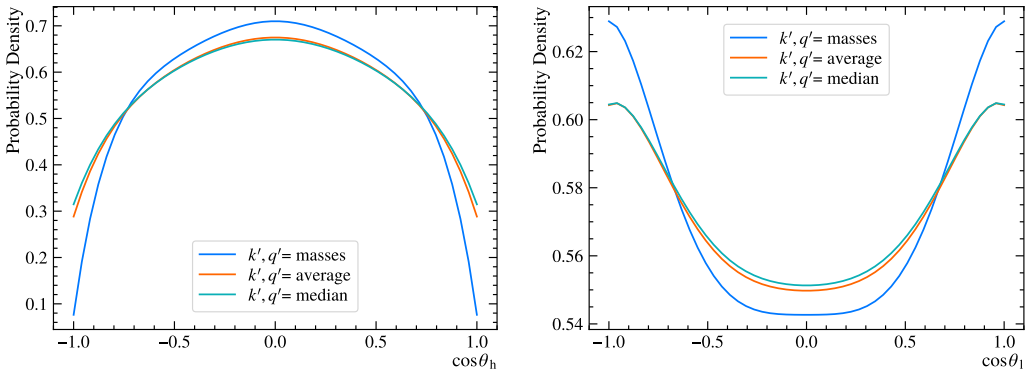


(c) ϕ, low-q^2 (d) $\cos\theta_h$, high-q^2

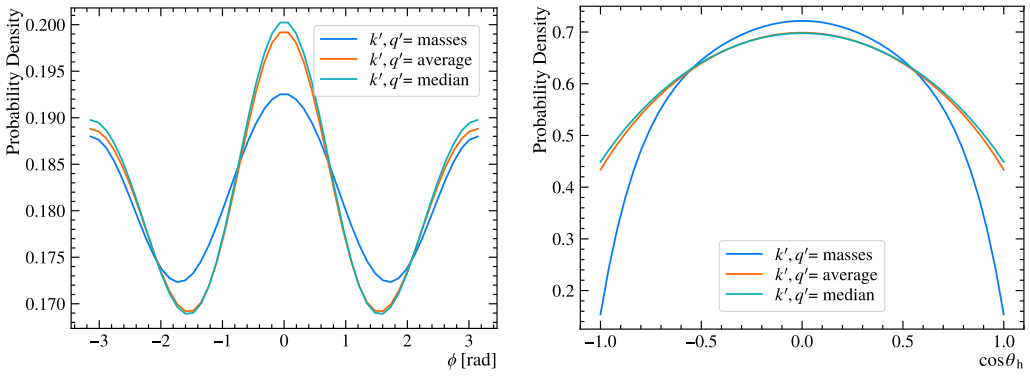


(e) $\cos\theta_l$, high-q^2 (f) ϕ, high-q^2

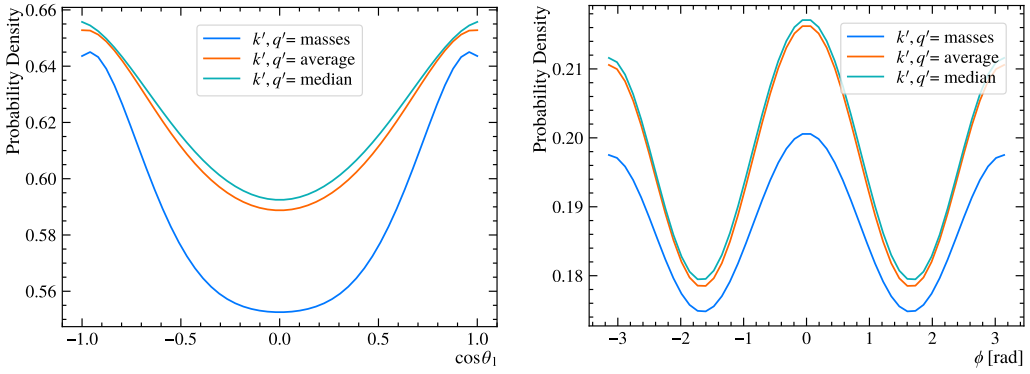
Figure D.2.1: Comparison of angular acceptance functions for different choices of fixed k' values in the $m(\pi^+\pi^-)$ range $[620, 920]$ MeV/c^2 and in the low- and high- q^2 bins.



(a) $\cos\theta_h$, J/ψ-q^2 (b) $\cos\theta_l$, J/ψ-q^2

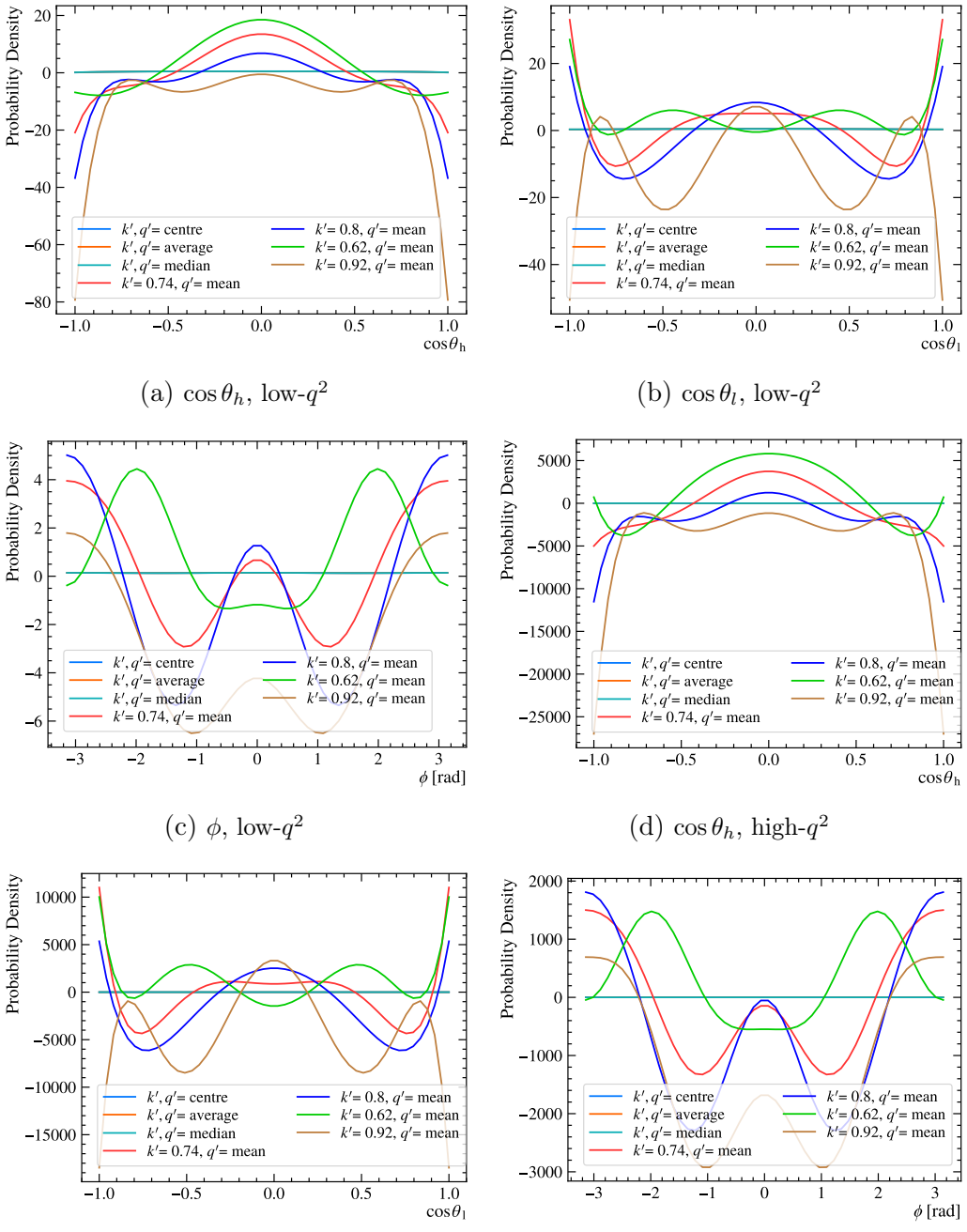


(c) ϕ, J/ψ-q^2 (d) $\cos\theta_h$, $\psi(2S)$-q^2



(e) $\cos\theta_l$, $\psi(2S)$-q^2 (f) ϕ, $\psi(2S)$-q^2

Figure D.2.2: Comparison of angular acceptance functions for different choices of fixed k' values in the $m(\pi^+\pi^-)$ range $[620, 920]$ MeV/c^2 and in the $c\bar{c}$ resonance q^2 bins.



(a) $\cos\theta_h$, low-q^2

(b) $\cos\theta_l$, low-q^2

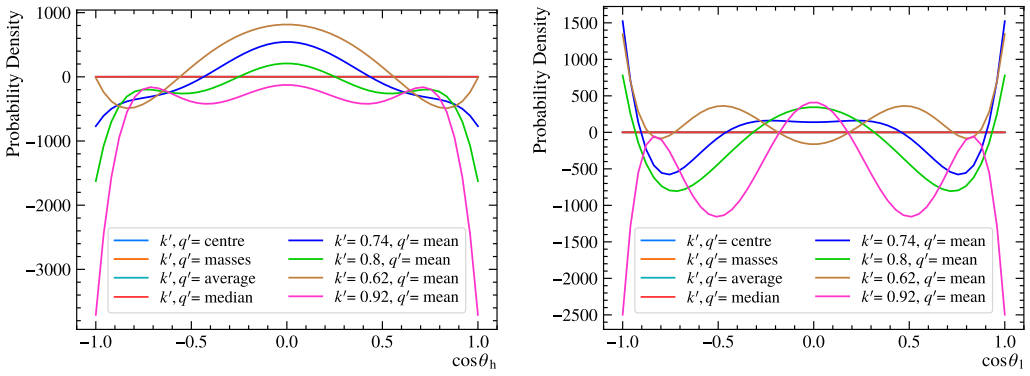
(c) ϕ, low-q^2

(d) $\cos\theta_h$, high-q^2

(e) $\cos\theta_l$, high-q^2

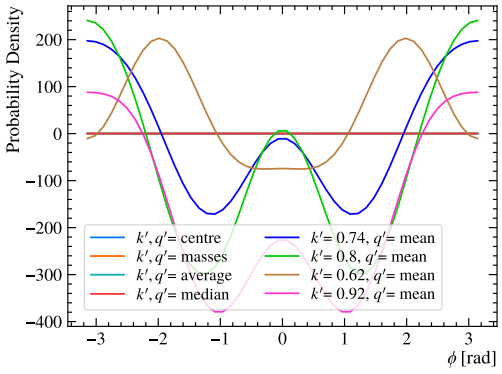
(f) ϕ, high-q^2

Figure D.2.3: Comparison of angular acceptance functions for different choices of fixed k' values in the $m(\pi^+\pi^-)$ range $[620, 920]$ MeV/c^2 and in the low- and high- q^2 bins.

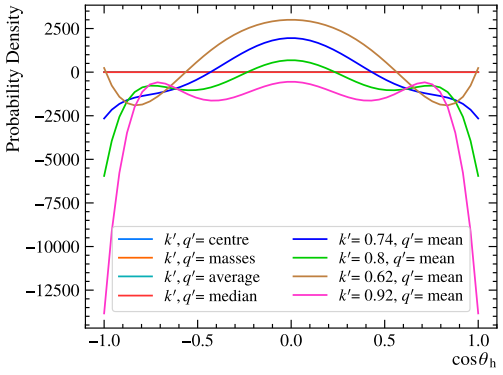


(a) $\cos\theta_h$, J/ψ-q^2

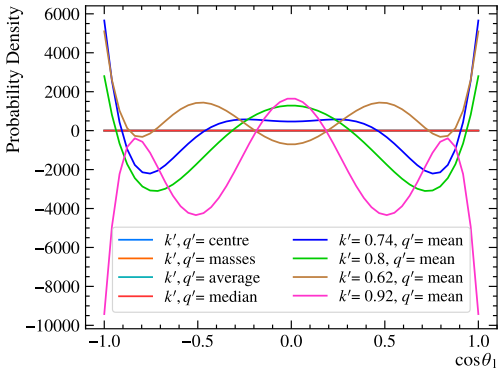
(b) $\cos\theta_l$, J/ψ-q^2



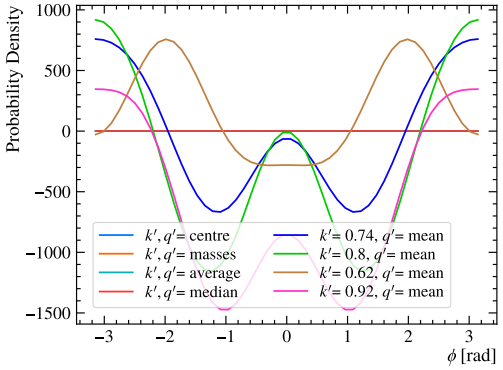
(c) ϕ, J/ψ-q^2



(d) $\cos\theta_h$, $\psi(2S)$-q^2



(e) $\cos\theta_l$, $\psi(2S)$-q^2



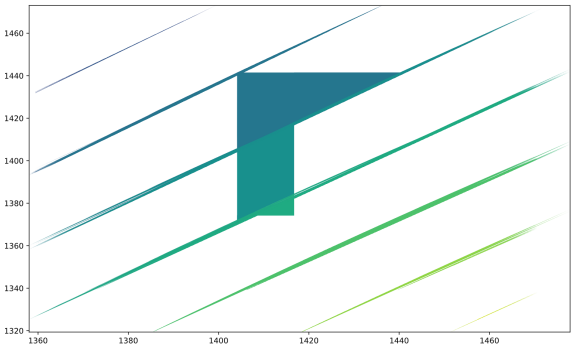
(f) ϕ, $\psi(2S)$-q^2

Figure D.2.4: Comparison of angular acceptance functions for different choices of fixed k' values in the $m(\pi^+\pi^-)$ range $[620, 920]$ MeV/c^2 and in the $c\bar{c}$ resonance q^2 bins.

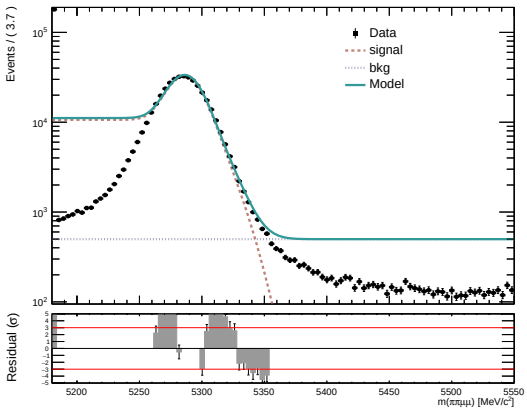
Appendix E

The wall of bad plots

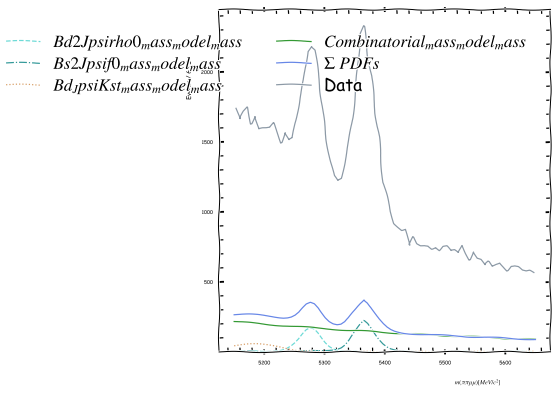
Over the course of this PhD, mistakes were made in the creation of figures. From typos in applying weights to using incorrect histogramming functions, from poor initial fit parameters to applying the wrong selection requirements, this is a selection of "bad plots that will undoubtedly cement my legacy. Figures E.0.1 and E.0.2 contain a curated selection of figures, with references to the equivalent figure which made it into this thesis.



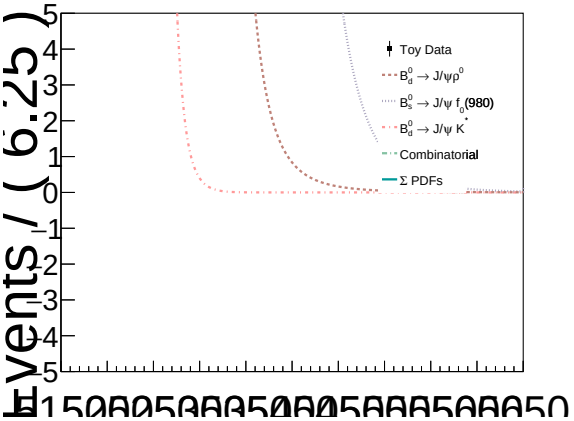
(a) 3.6.4



(b) 3.7.2

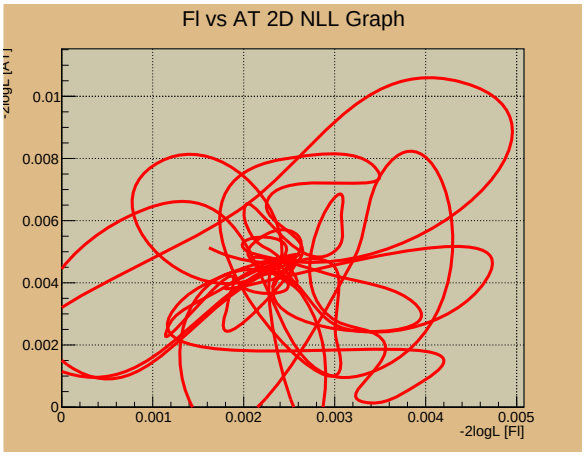


(c) 4.1.2

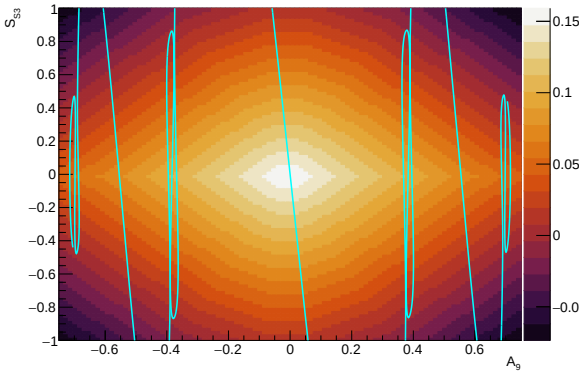


(d) 4.1.2

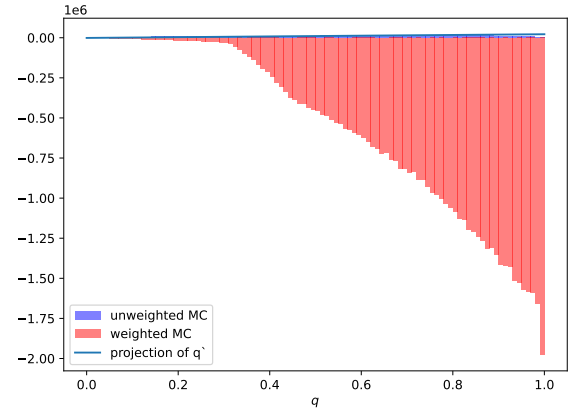
Figure E.0.1: "Bad Plot collage (1). The references point to the version of the plot which made it into this thesis.



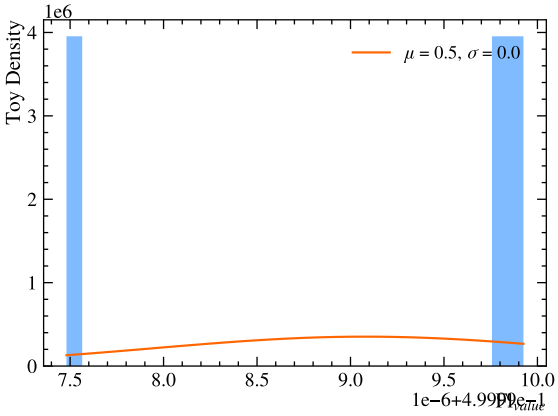
(a) 4.2.1



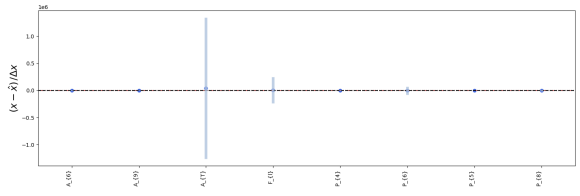
(b) 4.2.4a



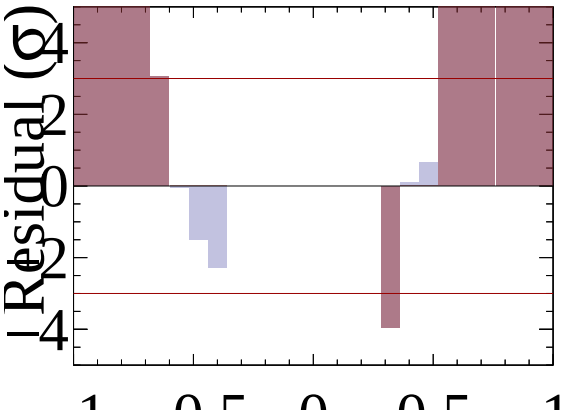
(c) 4.3.1



(d) 4.4.3



(e) 4.4.4



(f) 4.5.3

Figure E.0.2: "Bad Plot collage (2). The references point to the version of the plot which made it into this thesis.

Bibliography

[1] The LHCb Collaboration, *Angular analysis of the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay using* $3fb^{-1}$ *of integrated luminosity*, Journal of High Energy Physics **2016** (2016) .

[2] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) 1.

[3] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) 30.

[4] M. Neubert, *Heavy-Quark Effective Theory*, 1996.

[5] The LHCb Collaboration, *Measurement of the* Ω_c^0 *Baryon Lifetime*, Physical Review Letters **121** (2018) .

[6] The LHCb Collaboration, *Precision measurement of the* Λ_c^+, Ξ_c^+ *and* Ξ_c^0 *baryon lifetimes*, Physical Review D **100** (2019) .

[7] Particle Data Group, M. Tanabashi *et* al., *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[8] M. Thomson, *Modern particle physics*, Cambridge University Press, New York, 2013.

[9] W. N. Cottingham and D. A. Greenwood, *An Introduction to the Standard Model of Particle Physics*, Cambridge University Press, 2023.

[10] A. J. Buras, *Weak Hamiltonian, CP Violation and Rare Decays*, 1998.

[11] T. Blake, G. Lanfranchi, and D. M. Straub, *Rare B Decays as Tests of the Standard Model*, Prog. Part. Nucl. Phys. **92** (2017) 50, arXiv:1606.00916.

[12] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.

[13] Belle Collaboration, *Observation of a narrow charmonium-like state in exclusive $B^+ \to K^+\pi^+\pi^- J/\psi$ decays*, Physical Review Letters **91** (2003) .

[14] The LHCb Collaboration, *Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Physical Review D **95** (2017) .

[15] The LHCb Collaboration, *Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Physical Review Letters **118** (2017) .

[16] The BABAR Collaboration, The Belle Collaboration, *Precise measurement of the $e^+e^- \to \pi^+\pi^- J/\psi$ cross section at center-of-mass energies from 3.77 to 4.60 GeV*, Physical Review Letters **118** (2017) .

[17] The BESIII Collaboration, *Evidence of two resonant structures in $e^+e^- \to \pi^+\pi^- h_c$*, Physical Review Letters **118** (2017) .

[18] The Belle Collaboration, *Observation of an alternative $\Xi_c^0(2P)$ candidate in $e^+e^- \to J/\psi DD^-$*, Physical Review D **95** (2017) .

[19] The BESIII Collaboration, *Measurement of $e^+e^- \to \pi^+\pi^-\psi(3686)$ from 4.008 to 4.600 GeV and observation of a charged structure in the $\pi^\pm\psi(3686)$ mass spectrum*, Physical Review D **96** (2017) .

[20] The LHCb Collaboration, *Evidence for an $\eta_c(1S)\pi^-$ resonance in $B^0 \to \eta_c(1S)K^+\pi^-$ decays*, The European Physical Journal C **78** (2018) .

[21] The LHCb Collaboration, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Physical Review Letters **115** (2015) .

[22] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[23] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[24] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

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

[26] B. W. Lee, C. Quigg, and H. B. Thacker, *The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass*, Phys. Rev. Lett. **38** (1977) 883.

[27] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[28] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[29] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

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

[31] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[32] J. Charles *et* al., *CP violation and the CKM matrix: assessing the impact of the asymmetric B factories*, The European Physical Journal C **41** (2005) 1–.

[33] The CDF Collaboration, *Measurement of* $B(t \to Wb)/B(t \to Wq)$ *at the Collider Detector at Fermilab*, Phys. Rev. Lett. **95** (2005) 102002, arXiv:hep-ex/0505091.

[34] Particle Data Group, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[35] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

[36] A. Buras, *Gauge Theory of Weak Decays*, Cambridge University Press, 2020.

[37] A. Bharucha, D. M. Straub, and R. Zwicky, $b \to vl^+l^-$ *in the standard model from light-cone sum rules*, Journal of High Energy Physics **2016** (2016) .

[38] The LHCb Collaboration, *Measurement of CP Averaged Observables in the* $B^0 \to K^{*0}\mu^+\mu^-$ *Decay*, Physical Review Letters **125** (2020) .

[39] The LHCb Collaboration, *Differential branching fractions and isospin asymmetries of* $B \to K(*)\mu^+\mu^-$ *decays*, Journal of High Energy Physics **2014** (2014) .

[40] The LHCb Collaboration, *Measurements of the S-wave fraction in* $B^0 \to K^+\pi^-\mu^+\mu^-$ *decays and the* $B^0 \to K^*(892)^0\mu^+\mu^-$ *differential branching fraction*, JHEP **11** (2016) 047, arXiv:1606.04731, [Erratum: JHEP 04, 142 (2017)].

[41] The LHCb Collaboration, *Branching Fraction Measurements of the Rare* $B_s^0 \to \phi\mu^+\mu^-$ *and* $B_s^0 \to f_2'(1525)\mu^+\mu^-$ *Decays*, Physical Review Letters **127** (2021) .

[42] The LHCb Collaboration, The LHCb Collaboration, *Measurement of the* $\Lambda_b^0 \to \Lambda(1520)\mu^+\mu^-$ *Differential Branching Fraction*, Phys. Rev. Lett. **131** (2023) 151801, arXiv:2302.08262.

[43] W. Wang and S. Zhao, *Implications of the R_K and R_{K^*} anomalies*, Chin. Phys. C **42** (2018) 013105, arXiv:1704.08168.

[44] S.-L. Chen et al., *Signatures of a flavor changing Z' boson in $B \to \gamma Z'$*, Nuclear Physics B **962** (2021) 115237.

[45] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Physics Reports **516** (2012) 1–102.

[46] T. Blake, T. Gershon, and G. Hiller, *Rare b Hadron Decays at the LHC*, Annual Review of Nuclear and Particle Science **65** (2015) 113–143.

[47] K. G. Wilson, *Nonlagrangian models of current algebra*, Phys. Rev. **179** (1969) 1499.

[48] K. G. Wilson and W. Zimmermann, *Operator product expansions and composite field operators in the general framework of quantum field theory*, Commun. Math. Phys. **24** (1972) 87.

[49] R. Bause, H. Gisbert, M. Golz, and G. Hiller, *Model-independent analysis of $b \to d$ processes*, The European Physical Journal C **83** (2023) .

[50] S. Descotes-Genon and J. Virto, *Time dependence in $B \to V l^+ l^-$ decays*, Journal of High Energy Physics **2015** (2015) .

[51] The LHCb Collaboration, *Measurements of the B^+, B^0, B_s^0 meson and Λ_b^0 baryon lifetimes*, Journal of High Energy Physics **2014** (2014) .

[52] F. Krüger and J. Matias, *Probing new physics via the transverse amplitudes of $B^0 \to K^{*0}(\to K^-\pi^+)l^+l^-$ at large recoil*, Physical Review D **71** (2005) .

[53] The LHCb Collaboration, *Measurement of the resonant and $\overline{B}^0 \to J/\psi\pi^+\pi^-$ decays*, Physical Review D **90** (2014) .

[54] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the $B \to K^* l^+ l^-$ decays in the large-recoil region: SM challenges and BSM opportunities*, Physical Review D **93** (2016) .

[55] C. Bobeth, G. Hiller, and D. van Dyk, *More benefits of semileptonic rare B decays at low recoil: CP violation*, Journal of High Energy Physics **2011** (2011) .

[56] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, *The decay $\overline{B} \to \overline{K} l^+ l^-$ at low hadronic recoil and model-independent ΔB=1 constraints*, Journal of High Energy Physics **2012** (2012) .

[57] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Lattice QCD calculation of form factors describing the rare decays $B \to K^* l^+ l^-$ and $B_s \to \phi l^+ l^-$*, Physical Review D **89** (2014) .

[58] W. Altmannshofer and D. M. Straub, *New physics in $b \to s$ transitions after lhc run 1*, The European Physical Journal C **75** (2015) .

[59] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ Observables Using Form Factors from Lattice QCD*, Physical Review Letters **112** (2014) .

[60] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Rare B decays using lattice QCD form factors*, 2015.

[61] S. Descotes-Genon and M. N. Brunet, *Angular analysis of the rare decay $\Lambda_b \to \Lambda(1520)\left(\to N\overline{K}\right) l^+l^-$*, Journal of High Energy Physics **2019** (2019) .

[62] Y.-S. Li, S.-P. Jin, J. Gao, and X. Liu, *Transition form factors and angular distributions of the $\Lambda_b \to \Lambda(1520)(\to N\bar{K})l^+l^-$ decay supported by baryon spectroscopy*, Physical Review D **107** (2023) .

[63] Y. Amhis, M. Bordone, and M. Reboud, *Dispersive analysis of $\Lambda_b \to \Lambda(1520)$ local form factors*, Journal of High Energy Physics **2023** (2023) .

[64] S. Meinel and G. Rendon, *$\Lambda_c \to \Lambda^*(1520)$ form factors from lattice QCD and improved analysis of the $\Lambda_b \to \Lambda^*(1520)$ and $\Lambda_b \to \Lambda_c^*(2595, 2625)$ form factors*, Physical Review D **105** (2022) .

[65] M. e. a. Aaboud, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Journal of High Energy Physics **2018** (2018) .

[66] The CMS Collaboration, *Angular analysis of the $B^0 \to K^{*0}(892)\mu^+\mu^-$ decay at $\sqrt{s}$ = 13 TeV* , CERN, Geneva, 2024.

[67] Belle Collaboration, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, 2016.

[68] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of $B \to K^*\mu^+\mu^-$ at large recoil*, Journal of High Energy Physics **2013** (2013) .

[69] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *On the impact of power corrections in the prediction of $B \to K^* \mu^+ \mu^-$ observables*, Journal of High Energy Physics **2014** (2014) .

[70] The LHCb Collaboration, *Angular analysis of the rare decay $B_s^0 \to \phi \mu^+ \mu^-$*, Journal of High Energy Physics **2021** (2021) .

[71] The LHCb Collaboration, *Angular Analysis of the $B^+ \to K^{*+} \mu^+ \mu^-$ Decay*, Physical Review Letters **126** (2021) .

[72] N. Gubernari, M. Reboud, D. van Dyk, and J. Virto, *Improved theory predictions and global analysis of exclusive $b \to s \mu^+ \mu^-$ processes*, Journal of High Energy Physics **2022** (2022) .

[73] B. Capdevila *et* al., *Patterns of New Physics in $b \to s l^+ l^-$ transitions in the light of recent data*, Journal of High Energy Physics **2018** (2018) .

[74] The LHCb Collaboration, *Test of lepton universality in $b \to s l^+ l^-$ decays*, Phys. Rev. Lett. **131** (2023) 051803, arXiv:2212.09152.

[75] M. Bordone, G. Isidori, and A. Pattori, *On the standard model predictions for R_K and R_{K^*}*, The European Physical Journal C **76** (2016) .

[76] The LHCb Collaboration, *Test of Lepton Universality Using $B^+ \to K^+ l^+ l^-$ Decays*, Physical Review Letters **113** (2014) .

[77] The LHCb Collaboration, *Test of lepton universality with $B^0 \to K^{*0} l^+ l^-$ decays*, Journal of High Energy Physics **2017** (2017) .

[78] The LHCb Collaboration, *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.

[79] The LHCb Collaboration, *Test of lepton universality with $\Lambda_b^0 \to p K^- \ell^+ \ell^-$ decays*, Journal of High Energy Physics **2020** (2020) .

[80] J.-T. e. a. Wei, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \to K^{(*)} l^+ l^-$*, Physical Review Letters **103** (2009) .

[81] S. e. a. Wehle, *Test of Lepton-Flavor Universality in $B \to K^* l^+ l^-$ Decays at Belle*, Physical Review Letters **126** (2021) .

[82] S. e. a. Choudhury, *Test of lepton flavor universality and search for lepton flavor violation in $B \to K l^+ l^-$ decays*, Journal of High Energy Physics **2021** (2021) .

[83] The BABAR Collaboration, *Measurement of branching fractions and rate asymmetries in the rare decays* $B \to K^{(*)} l^+ l^-$, Physical Review D **86** (2012) .

[84] LHCb collaboration, R. Aaij *et* al., *First observation of the decay* $B^+ \to \pi^+ \mu^+ \mu^-$, JHEP **12** (2012) 125, arXiv:1210.2645.

[85] The LHCb Collaboration, *First measurement of the differential branching fraction and CP asymmetry of the* $B^\pm \to \pi^\pm \mu^+ \mu^-$ *decay*, Journal of High Energy Physics **2015** (2015) .

[86] The LHCb Collaboration, *Measurement of the* $B_s^0 \to \mu^+ \mu^-$ *decay properties and search for the* $B^0 \to \mu^+ \mu^-$ *and B0s* $\to \mu^+ \mu^- \gamma$ *decays*, Physical Review D **105** (2022) .

[87] The LHCb Collaboration, *Evidence for the decay* $B_S^0 \to \overline{K}^{*0} \mu^+ \mu^-$, Journal of High Energy Physics **2018** (2018) .

[88] The LHCb Collaboration, *Observation of the suppressed decay* $\Lambda_b^0 \to p \pi^- \mu^+ \mu^-$, Journal of High Energy Physics **2017** (2017) .

[89] The LHCb Collaboration, *Study of the rare* B_s^0 *and* B^0 *decays into the* $\pi^+ \pi^- \mu^+ \mu^-$ *final state*, Physics Letters B **743** (2015) 46–55.

[90] D. Melikhov, N. Nikitin, and S. Simula, *Rare exclusive semileptonic* $b \to s$ *transitions in the standard model*, Physical Review D **57** (1998) 6814–6828.

[91] M. Beneke, T. Feldmann, and D. Seidel, *Exclusive radiative and electroweak* $b \to d$ *and* $b \to s$ *penguin decays at NLO*, The European Physical Journal C **41** (2005) 173–188.

[92] A. Ali, A. Y. Parkhomenko, and A. V. Rusov, *Precise calculation of the dilepton invariant-mass spectrum and the decay rate in* $B^\pm \to \pi^\pm \mu^+ \mu^-$ *in the SM*, Physical Review D **89** (2014) .

[93] C. Hambrock, A. Khodjamirian, and A. Rusov, *Hadronic effects and observables in* $B \to \pi l^+ l^-$ *decay at large recoil*, Physical Review D **92** (2015) .

[94] J. A. Bailey *et* al., $B \to \pi l^+ l^-$ *Form Factors for New Physics Searches from Lattice QCD*, Physical Review Letters **115** (2015) .

[95] T. Mannel, *Effective field theories in flavour physics*, Springer tracts in modern physics, Springer, Berlin, 2004.

[96] A. Greljo, J. Salko, A. Smolkovič, and P. Stangl, *Rare b decays meet high-mass Drell-Yan*, Journal of High Energy Physics **2023** (2023) .

[97] M. Algueró *et* al., *To (b)e or not to (b)e: no electrons at LHCb*, The European Physical Journal C **83** (2023) .

[98] M. Ciuchini *et* al., *Constraints on lepton universality violation from rare B decays*, Physical Review D **107** (2023) .

[99] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

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

[101] The LHCb Collaboration, *LHCb detector performance*, International Journal of Modern Physics A **30** (2015) 1530022.

[102] T. Fazzini *et* al., *Electron Decay of the Pion*, Phys. Rev. Lett. **1** (1958) 247.

[103] G. Fidecaro, *The Discoveries of Rare Pion Decays at the CERN Synchrocyclotron*, Adv. Ser. Dir. High Energy Phys. **23** (2015) 397.

[104] A. Lopes and M. L. Perrey, *FAQ-LHC The guide*, Official LHC FAQ, 2022.

[105] E. Lopienska, *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, , General Photo.

[106] M. Thomson, *M*odern Particle Physics, Cambridge University Press, 2013.

[107] Chao, Alexander Wu and Mess, Karl Hubert and Tigner, Maury and Zimmermann, Frank, ed., *Handbook of accelerator physics and engineering*: 2nd Edition, World Scientific, Hackensack, USA, 2 ed., 2013.

[108] ATLAS collaboration, G. Aad *et* al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[109] CMS collaboration, S. Chatrchyan *et* al., *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[110] ALICE collaboration, K. Aamodt *et* al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[111] LHC Study Group, T. S. Pettersson and P. Lefèvre, *The Large Hadron Collider: Conceptual design* , CERN, 1995.

[112] TOTEM collaboration, G. Anelli *et* al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[113] LHCf collaboration, O. Adriani *et* al., *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.

[114] J. Pinfold, *Moedal-mapp, an lhc dedicated detector search facility*, 2023.

[115] E. A. D. Wolf, *Diffractive scattering*, Journal of Physics G: Nuclear and Particle Physics **28** (2002) 1023–1044.

[116] F. Collaboration, *The faser detector*, 2022.

[117] T. S. Collaboration, *Snd@lhc: The scattering and neutrino detector at the lhc*, 2023.

[118] R. e. a. Aaij, *Performance of the lhcb vertex locator*, Journal of Instrumentation **9** (2014) P09007–P09007.

[119] R. Lindner, *LHCb layout 2. LHCb schema 2*, LHCb Collection., 2008.

[120] M. Stahl, *Machine learning and parallelism in the reconstruction of lhcb and its upgrade*, Journal of Physics: Conference Series **898** (2017).

[121] LHCb collaboration, e. a. Barbosa-Marinho, *LHCb VELO (VErtex LOcator): Technical Design Report*, Technical design report. LHCb, CERN, Geneva, 2001.

[122] R. Aaij and et al, *Performance of the LHCb vertex locator*, Journal of Instrumentation **9** (2014) P09007.

[123] R. A. et al, *Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*, Physical Review Letters **108** (2012) .

[124] R. A. et al, *Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^-\pi^+$*, New Journal of Physics **15** (2013) 053021.

[125] J. Gassner, F. Lehner, and S. Steiner, *The mechanical Design of the LHCb Silicon Trigger Tracker* , CERN, Geneva, 2010.

[126] C. A. Beteta *et* al., *Monitoring radiation damage in the LHCb tracker turicensis*, Journal of Instrumentation **15** (2020) P08016.

[127] The LHCb Collaboration, *LHCb magnet: Technical Design Report*, CERN, Geneva, 2000.

[128] The LHCb Collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN, Geneva, 2003.

[129] The LHCb Collaboration, *Performance of the LHCb Outer Tracker*, Journal of Instrumentation **9** (2014) P01002–P01002.

[130] The LHCb Collaboration, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, Journal of Instrumentation **12** (2017) P11016–P11016.

[131] M. e. a. Adinolfi, *Performance of the lhcb rich detector at the lhc*, The European Physical Journal C **73** (2013) .

[132] LHCbRICH collaboration, N. e. a. Brook, *LHCb RICH1 Engineering Design Review Report* , CERN, Geneva, 2005.

[133] M. e. a. Adinolfi, *LHCb RICH 2 engineering design review report* , CERN, Geneva, 2002. revised version number 1 submitted on 2002-05-21 14:24:22.

[134] LHCb collaboration, R. Forty, *RICH pattern recognition for LHCb*, Nucl. Instrum. Meth. A **433** (1999) 257.

[135] R. Calabrese *et* al., *Performance of the LHCb RICH detectors during LHC Run 2*, Journal of Instrumentation **17** (2022) P07013.

[136] LHCb collaboration, S. e. a. Amato, *LHCb calorimeters: Technical Design Report*, Technical design report. LHCb, CERN, Geneva, 2000.

[137] Y. Guz, *LHCb Calorimeter Upgrade*, .

[138] G. F. Knoll, *Radiation Detection and Measurement, 3rd ed.*, John Wiley and Sons, New York, 3rd edition ed., 2000.

[139] The LHCb Collaboration, *LHCb technical proposal: A Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays* , CERN, 1998.

[140] A. A. Alves, Jr. *et* al., *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[141] The LHCb Collaboration, *Performance of the muon identification at lhcb*, Journal of Instrumentation **8** (2013) P10020–P10020.

[142] The LHCb Collaboration, *The lhcb trigger and its performance in 2011*, Journal of Instrumentation **8** (2013) P04022–P04022.

[143] The LHCb Collaboration, *LHCb Trigger and Online Upgrade Technical Design Report* , CERN, 2014.

[144] The LHCb Collaboration, *Trigger schemes*, .

[145] G. Corti *et* al., *Software for the LHCb experiment*, IEEE Trans. Nucl. Sci. **53** (2006) 1323.

[146] The LHCb Collaboration, *Design and performance of the lhcb trigger and full real-time reconstruction in run 2 of the lhc*, Journal of Instrumentation **14** (2019) P04013–P04013.

[147] LHCb collaboration, M. Clemencic *et* al., *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[148] C. Bierlich *et* al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, SciPost Phys. Codeb. **2022** (2022) 8, arXiv:2203.11601.

[149] I. Belyaev *et* al., *Handling of the generation of primary events in Gauss, the LHCb simulation framework* , CERN, Geneva, 2010.

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

[151] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and physics documentation*, Comp. Phys. Comm. **199** (2016) 86, arXiv:1011.0937.

[152] GEANT4 collaboration, S. Agostinelli *et* al., *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[153] J. Allison *et* al., *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[154] A. Piucci, *The LHCb Upgrade*, J. Phys. Conf. Ser. **878** (2017) 012012.

[155] M. Williams *et* al., *The HLT2 Topological Lines* , CERN, Geneva, 2011.

[156] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.

[157] K. Do, T. Tran, and S. Venkatesh, *Learning deep matrix representations*, 2018.

[158] T. Likhomanenko *et* al., *LHCb topological trigger reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025.

[159] The LHCb Collaboration, *Measurement of the resonant and CP components in $B^0 \to J/\psi\pi^+\pi^-$ decays*, Physical Review D **90** (2014) .

[160] T. Chen and C. Guestrin, *Xgboost: A scalable tree boosting system*, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016) .

[161] A. P. Bradley, *The use of the area under the roc curve in the evaluation of machine learning algorithms*, Pattern Recognition **30** (1997) 1145.

[162] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, in *Learning From Data*, ch. 4, AMLBook, 2012.

[163] M. Stone, *Cross-validatory choice and assessment of statistical predictions*, Journal of the royal statistical society. Series B (Methodological) (1974) 111.

[164] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, *Gene selection for cancer classification using support vector machines*, Machine Learning **46** (2002) 389.

[165] J. Bergstra and Y. Bengio, *Random search for hyper-parameter optimization*, Journal of Machine Learning Research **13** (2012) 281.

[166] LHCb collaboration, D. Derkach, M. Hushchyn, and N. Kazeev, *Machine Learning based Global Particle Identification Algorithms at the LHCb Experiment*, EPJ Web Conf. **214** (2019) 06011.

[167] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, The Annals of Mathematical Statistics **9** (1938) 60 .

[168] L. Anderlini *et* al., *The PIDCalib package*, LHCb-PUB-2016-021, 2016.

[169] LHCb collaboration, The LHCb Collaboration, *Observation of $B^0_{(s)} \to J/\psi p\overline{p}$ decays and precision measurements of the $B^0_{(s)}$ masses*, Phys. Rev. Lett. **122** (2019) 191804, arXiv:1902.05588.

[170] The LHCb Collaboration, *Observation of the $\Lambda^0_b \to J/\psi p\pi^-$ decay*, Journal of High Energy Physics **2014** (2014) .

[171] 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.

[172] The LHCb Collaboration, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B^0_s decay branching fractions*, Physical Review D **104** (2021) .

[173] C. O. Dib, C. S. Kim, and N. A. Neill, *The b quark fragmentation fractions at LHCb and meson decays with heavy quark spectators*, The European Physical Journal C **83** (2023) .

[174] The LHCb Collaboration, *Precision measurement of the B_c^+ meson mass*, JHEP **07** (2020) 123, arXiv:2004.08163.

[175] R. Aaij et al., *Selection and processing of calibration samples to measure the particle identification performance of the lhcb experiment in run 2*, EPJ Techniques and Instrumentation **6** (2019) .

[176] The LHCb collaboration, *Measurement of the track reconstruction efficiency at lhcb*, Journal of Instrumentation **10** (2015) P02007–P02007.

[177] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb* , CERN, Geneva, 2014.

[178] M. Oreglia et al., *A Study of the Reaction $\psi' \to \gamma\gamma J/\psi$*, Phys. Rev. D **25** (1982) 2259.

[179] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[180] H. Dembinski, M. Kenzie, C. Langenbruch, and M. Schmelling, *Custom orthogonal weight functions (cows) for event classification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1040** (2022) 167270.

[181] A. Rogozhnikov, *Reweighting with boosted decision trees*, Journal of Physics: Conference Series **762** (2016) 012036.

[182] A. Rogozhnikov, *Hep ml documentation*, https://arogozhnikov.github.io/hepml.

[183] I. M. Chakravarti, R. G. Laha, J. Roy, and J. Roy, *H*andbook of Methods of Applied Statistics, no. v. 1 in Handbook of Methods of Applied Statistics, Wiley, 1967.

[184] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949) 149.

[185] G. Upton and I. Cook, *A* Dictionary of Statistics, Oxford Paperback Reference, OUP Oxford, 2008.

[186] G. Cowan, *Statistical data analysis*, Oxford University Press, 1998.

[187] F. James and M. Roos, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, Computer Physics Communications **10** (1975) 343.

[188] F. Beaujean, M. Chrząszcz, N. Serra, and D. van Dyk, *Extracting angular observables without a likelihood and applications to rare decays*, Physical Review D **91** (2015) .

[189] G. E. Forsythe, *Von neumann's comparison method for random sampling from the normal and other distributions*, Mathematics of Computation **26** (1972) 817.

[190] D. Mccaffrey, G. Ridgeway, and A. Morral, *Propensity score estimation with boosted regression for evaluating causal effects in observational studies*, Psychological methods **9** (2004) 403.

[191] Student, *The probable error of a mean*, Biometrika **6** (1908) 1.

[192] J. F. Kenney, *Mathematics of Statistics*, no. pt. 1 in Mathematics of Statistics, D. Van Nostrand Company, 1939.

[193] The LHCb Collaboration, *Measurement of resonant and $\overline{B}{}^0_s \to J/\psi\pi^+\pi^-$ decays*, Physical Review D **89** (2014) .

[194] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **764** (2014) 150–155.

[195] The LHCb Collaboration, *Measurement of the $\bar{B}^0 - B^0$ and $\bar{B}^0_s - B^0_s$ production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Physics Letters B **739** (2014) 218–228.

[196] Y. Amhis *et* al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021*, Phys. Rev. D **107** (2023) 052008, arXiv:2206.07501.