Using Rare Decays to Probe the Standard Model at LHCb

Indrek Sepp

High Energy Physics

Blackett Laboratory

Imperial College London

A thesis submitted to Imperial College London

for the degree of Doctor of Philosophy

May 2014

Declaration

The work presented in this thesis was carried out between October 2010 and May 2014. It is the result of my own studies, with the support of members of the Imperial College HEP group and the broader LHCb collaboration. The work of others is explicitly referenced.

The entire thesis was written by myself. The entirety of the RICH performance analysis, presented in Chapter 4, was performed by myself, except for the PID efficiency measurements shown in Fig. 4.13 and Fig. 4.15. For the work presented in Chapter 5, I contributed to all aspects except for the selection development and items relating to the $\overline{B}^0 \to D^+\pi^-$ decay. I contributed to all aspects of the analysis described in Chapter 6, except for the selection development and the S-wave sub-analysis described in Appendix E.

This thesis has not been submitted for any other qualification.

Indrek Sepp, 22 May 2014

Abstract

The Large Hadron Collider beauty Experiment (LHCb) detector is one of the four main particle detectors on the Large Hadron Collider (LHC). It is dedicated to the study of physics processes involving b quarks. This thesis presents three analyses of data collected by LHCb.

The first measures the photoelectron yield of the Ring Imaging Cherenkov Detector (RICH) detector subsystem, which distinguishes between pions, kaons and protons. The yield is seen to be 15%(19%) less than that in the simulation for the C_4F_{10} (CF_4) radiator medium. The result is a Particle Identification (PID) performance which is sufficient for the physics goals of LHCb, albeit slightly less than expected from simulation. No evidence is found for the deterioration of the Hybrid Photon Detector (HPD) quantum efficiency, mirror reflectivity or RICH medium transparency over the course of 2011 and 2012 data collection.

The second analysis measures the dependence of the $b \to \Lambda_b^0$ to $b \to B^0$ hadronisation ratio on the p_{T} and η of the Λ_b^0 and B^0. An exponential function with a plateau provides the best fit for the p_{T} dependence. A linear dependence of the ratio to η is also observed. These observations are substantial improvements on previous measurements of the dependencies that can aid the development of QCD models and simulation frameworks that describe b quark hadronisation.

The third analysis presents the world's first search for the decays $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to \mu^+\mu^-\mu^+\mu^-$. Upper limits are set on the branching fractions of both decays that are ~ 2 orders of magnitude above Standard Model (SM) expectations. These limits begin to exclude the phase-space of supersymmetric models where the decays are mediated by S and P sgoldstinos.

Acknowledgements

I give many thanks to my supervisor Andrey Golutvin for all his advice, help and guidance throughout my PhD. I also thank all of my collaborators who worked with me on the research presented in this thesis.

I especially wish to acknowledge David Websdale for his input to my RICH studies, and for reviewing my entire thesis despite having no professional obligation to do so. Similarly, I thank everyone at Imperial College and CERN who took time out from their busy workload to help me.

Finally, many thanks to Isabel, Henry and the rest of my family for their patience.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Beginning with the discovery of the electron in 1897 by J.J. Thomson [20], the field of particle physics has yielded an impressive array of discoveries that have revolutionised our understanding of how nature operates at the fundamental level. This has culminated in the formulation of the Standard Model (SM) in the 1970's, a unified theoretical framework that describes all of the currently observed particles and forces of nature, except gravity. The SM is described in this thesis in Chapter 2. As of the time of writing the SM has withstood all direct experimental tests of its predictions[1]. Some of its predictions have been verified to extremely high degrees of precision, for example, the magnetic moment of the electron has been measured to a precision of 0.22 parts per billion [1]. All theories of physics provide only a partial description of nature. They are effective within certain scales and describe a limited set of phenomena, beyond which they lose their predictive power. The SM is not an exception, a hard limit on its effectiveness is set at the Planck energy scale, where gravity, which is not described by the SM, becomes non-negligible. In addition the SM does not provide a candidate for the particle(s) which constitute dark matter. Neither does it explain the dominance of matter over anti-matter in the universe. Sometimes a theory is successful to such a degree that it takes a long time for experiments to attain a sensitivity to phenomena which defy the theories predictions. Newtonian mechanics withstood direct experimental tests for many centuries, until falsification by the Michelson and Morley experiment in 1887 [21], the first experiment to be sensitive to relativistic phenomena. The longevity of the SM is already comparable to that of classical electrodynamics, which was formulated in 1861 [22] and falsified in 1905 [23], when the observation of the photoelectric effect ushered in the revolution of Quantum Mechanics. It may take many years, decades or more, but like classical mechanics or electrodynamics, the SM will be falsified eventually, hopefully bringing about a revolution in our understanding of nature on a par to those which began in 1887 and 1905.

The discovery of physics Beyond the Standard Model (BSM) can occur through the

[1] Barring the observation of neutrino oscillations, which although not predicted by the SM, can be incorporated into its framework

observation of new particles or forces when they are produced directly in particle colliders as they access ever-higher collision energies. This is one of the primary aims of the A Toroidal LHC ApparatuS (ATLAS) and Compact Muon Solenoid (CMS) detectors at the Large Hadron Collider (LHC). Alternatively, BSM physics can be observed indirectly through its effects on observables such as decay rates, which can be enchanced or suppressed relative to SM predictions. The Large Hadron Collider beauty Experiment (LHCb), described in Chapter 3, is designed to be sensitive to the indirect influence of new physics on processes involving b quarks. The work presented in this thesis is based on data collected at LHCb.

Chapter 2 provides an overview of the SM, its limitations and potential extensions. Chapter 3 describes the Large Hadron Collider beauty Experiment (LHCb) detector and its various subsystems. Chapter 4 describes the measurement of the photoelectron yield in the RICH subdetector systems, an important quantity that determines how well pions, kaons and protons are distinguished from each other at LHCb. Chapter 5 presents the measurement of the p_{T} and η dependence of the $b \to \Lambda_b^0$ to $b \to B^0$ hadronisation ratio. Chapter 6 details the search for the rare decays $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to \mu^+\mu^-\mu^+\mu^-$, processes which are suppressed in the SM but can be enchanced by BSM physics.

Chapter 2

The Standard Model and beyond

2.1 Overview

The Standard Model (SM) is a relativistic quantum field theory which describes how the Electromagnetic (EM), weak and strong forces interact with the elementary particles of matter in a unified framework. In the SM matter is composed of twelve fundamental spin-$\frac{1}{2}$ fermions. There are four types of fermion: up-type and down-type quarks, which participate in all the three SM forces; charged leptons, which interact via the weak and EM forces; and the uncharged neutrinos, which only participate in weak interactions. Each fermion type is made up of three generations of particles, for example, the charged leptons consist of the electron, the muon and the tau. The fermions interact with each other via the exchange of spin-1 gauge bosons. The gauge bosons include: the photon (γ), which mediates the EM force between charged particles; the charged $W^{\pm}$ and uncharged Z^0, which convey the weak nuclear force between the fermions; and the eight gluons (g), which transmit the strong nuclear force between the quarks. The SM also has one scalar spin-0 particle, the Higgs boson, which endows mass to the particles it interacts with. The particles described above are summarised in Fig. 2.1.

This chapter describes the basic features and mathematical construction of the SM. The limitations of the SM are also discussed as well as a brief introduction to the theory of supersymmetry and the phenomenology of sgoldstinos. This provides a theoretical background with which to motivate the search for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays, detailed in Chapter 6. The description of the SM provided in this chapter is derived from [24–26], unless referenced explicitly.

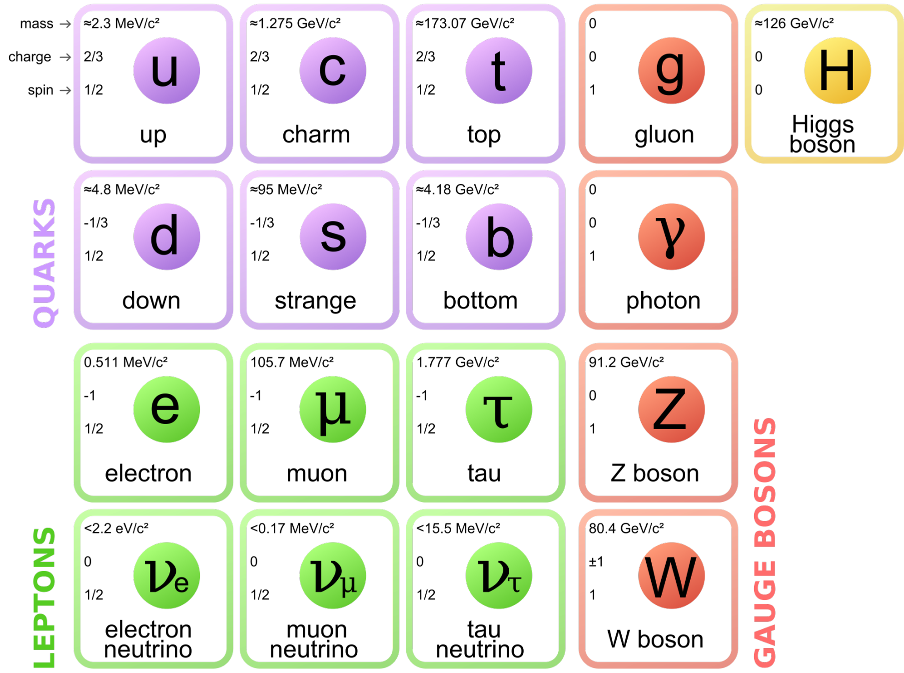


Figure 2.1: An overview of the fermions and gauge bosons of the SM, displaying the properties of charge, mass and spin. This figure is taken from `http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`.

2.2 Mathematical formalism of the SM

The SM is described by the matter particles described in Sect. 2.1 and the Lagrangian density[1] $\mathcal{L}$ which has kinetic, mass and interaction terms:

$$\mathcal{L} = \mathcal{L}_{kinetic} + \mathcal{L}_{mass} + \mathcal{L}_{interaction}. \tag{2.1}$$

The SM Lagrangian is invariant under the local transformations of the gauge group

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \tag{2.2}$$

The $SU(3)_C$ group describes the strong force through the framework of Quantum Chromodynamics (QCD). This force acts on the property of colour C. The $SU(2)_L \times U(1)_Y$ group (where L denotes that only fermions with left-handed chirality transform under the $SU(2)$ group) defines the unified EM and weak forces, refered to as the Electroweak (EW) force. This force couples to the hypercharge quantum number Y, defined as:

$$Y = 2(Q + T^3), \tag{2.3}$$

where Q is the EM charge and T^3 the weak isospin component.

[1]referred to as just the Lagrangian hereafter

2.2.1 Quantum Electrodynamics: An example of a gauge field theory

Local gauge invariance is a powerful feature of the SM, because its logical consequence is the introduction of force-mediating gauge bosons and interactions between them and the fermions. The construction of Quantum Electrodynamics (QED), which describes EM interactions, is a simple and elegant example of the principles of local gauge invariance. A fermion field $\psi(x)$ with mass m which does not interact with other particles is described by the free Dirac Lagrangian:

$$\mathcal{L}_{free} = \overbrace{i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi}^{\mathcal{L}_{kinetic}} - \overbrace{m\bar{\psi}\psi}^{\mathcal{L}_{mass}}, \tag{2.4}$$

where γ^{μ} are the Dirac matrices. The kinetic term of $\mathcal{L}_{free}$ is not invariant under local $U(1)$ gauge transformations, which for QED is written as:

$$\psi(x) \to \psi'(x) = e^{ie\theta(x)}\psi(x) \qquad \bar{\psi}(x) \to \psi'(x) = e^{-ie\theta(x)}\bar{\psi}(x), \tag{2.5}$$

where e is a constant, the unit of electric charge and $\theta(x)$ is an arbitrary real function of space-time x. This is because the derivative ∂_{μ} brings out an additional term $-e\bar{\psi}\gamma^{\mu}(\partial_{\mu}\theta(x))\psi$ in $\mathcal{L}_{free}$ after the gauge transformation. To make $\mathcal{L}_{free}$ invariant one must replace ∂_{μ} with a covariant derivative D_{μ}:

$$D_{\mu} = \partial_{\mu} - ieA_{\mu}, \tag{2.6}$$

where A_{μ} is a spin-1 vector field, a gauge boson, which is interpreted as the photon in QED. The vector field transforms as:

$$A_{\mu} \longrightarrow A'_{\mu} = A_{\mu} + \partial_{\mu}\theta(x). \tag{2.7}$$

Because the gauge boson field A_{μ} has been introduced, its kinetic term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, where $F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}$ needs to be added to the Lagrangian. The gauge boson field is required to be massless, because its corresponding mass term $\frac{1}{2}m^2A_{\mu}A^{\mu}$ would not be invariant under the gauge transformation shown in Equation (2.7). The resulting Lagrangian, after expanding D_{μ}, is the QED Lagrangian:

$$\mathcal{L}_{QED} = \overbrace{i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}}^{kinetic} - \overbrace{m\bar{\psi}\psi}^{mass} + \overbrace{e\bar{\psi}\gamma^{\mu}\psi A_{\mu}}^{interaction}. \tag{2.8}$$

The interaction term describes a coupling of the fermion current $j^{\mu} = \bar{\psi}\gamma^{\mu}\psi$ to the photon A_{μ}, with a strength defined by e. Demanding that a non-interacting fermion field obeys local $U(1)$ gauge invariance results in the fundamental equation describing electromagnetism.

2.3 Quantum Chromodynamics

QCD is the framework which describes the strong nuclear force, which acts on the quantum number of colour charge. Quarks can carry one of three types of colour charge: red, green and blue, the leptons carry no colour charge. In QCD the quarks are represented as fermion triplets:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}, \tag{2.9}$$

where the r, g, b subscripts denote the type of colour charge. QCD is constructed by requiring that the free Dirac Lagrangian, $\mathcal{L}_{free}$ in (2.4), is invariant under $SU(3)$ transformations:

$$\psi(x) \to \psi'(x) = e^{ig_s\lambda_i\theta_i(x)}\psi(x) \qquad \bar{\psi}(x) \to \psi'(x) = \bar{\psi}(x)e^{ig_s\lambda_i\theta_i(x)}, \tag{2.10}$$

where g_s is the strong coupling constant and λ_i are the set of eight 3×3 Gell-Mann matrices[2], the generators of the $SU(3)$ group. To make $\mathcal{L}_{free}$ invariant one has to replace the partial derivative with a covariant derivative:

$$D_\mu = \mathbb{I}\partial_\mu - ig_s\lambda^i G^i_\mu. \tag{2.11}$$

There are now eight gluon vector fields G^i_μ which transform as:

$$G^i_\mu \longrightarrow G^{i\prime}_\mu = G^i_\mu + \partial_\mu\theta^i(x) + f_{ijk}\theta^j(x)G^k_\mu, \tag{2.12}$$

where f_{ijk} are the structure constants of the $SU(3)$ group, as defined by the commutation relation $[\lambda_i, \lambda_j] = if_{ijk}\lambda_k$. The final term in Equation (2.12) is present because the $SU(3)$ group is 'non-Abelian', meaning that the Gell-mann matrices do not commute with each other. As in QED, the gauge bosons are required to be massless and their kinetic term, $-\frac{1}{4}F^i_{\mu\nu}F^{i\mu\nu}$, has to be added to $\mathcal{L}_{free}$, where:

$$F^{i\mu\nu} = \partial^\mu G^\nu - \partial^\nu G^\mu + g_s f_{ijk} G^j_\mu G^{k\nu}. \tag{2.13}$$

The final term in Equation (2.12) is also a result of the non-Abelian property of the $SU(3)$

[2]The summation of repeated indices such as i is implied throughout this text, i.e. $\lambda_i\theta_i(x) \equiv \sum_{i=1}^{8} \lambda_i\theta_i(x)$

group. The QCD Lagrangian with D_μ expanded is:

$$\mathcal{L}_{QCD} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - \frac{1}{4}F^i_{\mu\nu}F^{i\mu\nu} - m\bar{\psi}\psi + g_s\bar{\psi}\gamma^\mu\lambda^i\psi G^i_\mu. \quad (2.14)$$

The QCD Lagrangian contains interaction terms describing the coupling of the fermion current to each of the eight gluon fields. There is one instance of the QCD Lagrangian for each of the six quark flavours. After expanding out the gluon kinetic term in Equation (2.14) using Equation (2.13) one gets the following two interaction terms:

$$g_s f_{ijk}(\partial^\mu G^{i\nu} - \partial^\nu G^{i\mu})G^j_\mu G^k_\nu \text{ and } g_s^2 f_{ijk} f_{ilm} G^j_\mu G^k_\nu G^{l\mu} G^{m\nu}, \quad (2.15)$$

which describe the self coupling between three and four gluons. This gauge boson self-coupling is a feature of all non-Abelian gauge theories.

A unique feature of QCD is that the strength of the strong force between two quarks increases with separation. It would therefore take an infinite amount of energy to separate two quarks. The consequence of this is the property of 'confinement', which stipulates that the colour charge can never be observed directly. Therefore quarks can only exist in two different types of colourless bound states: mesons and baryons. Mesons are formed of quark-antiquark pairs, where the colour-charge of the quark is equal and opposite to that of the antiquark. An example is the B^0 meson, which is composed of a $\bar{b}$ and d quark. Baryons consist of three quarks, where each one of these quarks has a colour of blue, red or green. The combination of the three colours is a colourless state. An example is the proton, which consists of a uud quark combination.

2.4 Electroweak Theory

The combined EM and weak force, the electroweak force, arises from requiring the free Dirac Lagrangian, $\mathcal{L}_{free}$ in Equation (2.4), to be invariant under transformations of the $SU(2)_L \times U(1)_Y$ EW symmetry group. A fermion field can be split into two left/right handed chiral components: $\psi = (\psi_L + \psi_R)$, where $\psi_{L/R} = (1 \pm \gamma^5)\psi$. The $SU(2)_L$ component of the EW symmetry group acts only on left-handed fermion field ψ_L, the $U(1)_Y$ group acts on both ψ_R and the individual elements of ψ_L. The SM fermions are grouped into left-handed doublets $\vec{l^i}_L$, $\vec{Q}^i_L$ and right-handed singlets l^i_R, u^i_R and d^i_R, where i denotes the generation. $\vec{l^i}_L$ and $\vec{Q}^i_L$ are defined as:

$$\vec{l^i}_L = \begin{pmatrix} \nu^i \\ l^i_L \end{pmatrix} = \left(\begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix} \right) \qquad \vec{Q}^i_L = \begin{pmatrix} u^i_L \\ d^i_L \end{pmatrix} = \left(\begin{pmatrix} u_L \\ d_L \end{pmatrix} \begin{pmatrix} c_L \\ s_L \end{pmatrix} \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right), \quad (2.16)$$

they can transform under the whole EW symmetry group. The right-handed singlets are defined as:

$$l^i_R = \begin{pmatrix} e_R & \mu_R & \tau_R \end{pmatrix} \quad u^i_R = \begin{pmatrix} u_R & c_R & t_R \end{pmatrix} \quad d^i_R = \begin{pmatrix} d_R & s_R & b_R \end{pmatrix}, \tag{2.17}$$

they can only partake in $U(1)$ transformations. Note that the neutrino field does not have a right-handed component. As in QCD and QED, $\mathcal{L}_{free}$ is made gauge invariant by introducing a covariant derivative, which for the $SU(2)_L \times U(1)_Y$ group is:

$$D_\mu = \partial_\mu - \frac{ig}{2}\sigma^i W^i_\mu - \frac{ig'}{2} Y B_\mu, \tag{2.18}$$

where g and g' are the coupling constants of the $SU(2)_L$ and $U(1)_Y$ groups, respectively, Y is the weak hypercharge quantum number, σ^i are the three 2×2 Pauli matrices, the generators of the $SU(2)$ group. W^i_μ and B_μ are vector fields. By construction, only the left handed doublets can engage in the $SU(2)_L$ component of the gauge transformation and couple to the W^i_μ fields. The physical $W^\pm_\mu$,Z_μ and A_μ (photon) gauge bosons are defined as a mixture of the W^i_μ and B_μ fields:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp i W^2_\mu) \qquad \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} A^3_\mu \\ B_\mu \end{pmatrix}, \tag{2.19}$$

where θ_W is the Weinberg angle, defined as $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$ and $\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$. Writing the covariant derivative in Equation (2.18) in terms of the physical gauge boson fields yields:

$$D_\mu = \partial_\mu - i\frac{e}{\sqrt{2}\sin\theta_W}(T^+W^+_\mu + T^-W^-_\mu) - i\frac{e}{\cos\theta_W \sin\theta_W}(T^3 - \sin^2\theta_W Q)Z_\mu - ieQA_\mu, \tag{2.20}$$

where $T^\pm = \frac{1}{2}(\sigma^1 \pm i\sigma^2)$, $T^3 = \sigma^3/2$,$e = \frac{gg'}{\sqrt{g^2+g'^2}}$ and $Q = (T^3 + \frac{Y}{2})$. In equation (2.20) e is the coupling constant of the EM force and Q is the generator of electric charge. The EW theory has to yield the massive $W^\pm$ and Z gauge bosons and a massless photon. This is done by breaking the EW symmetry via the Higgs mechanism. To break the symmetry a complex scalar doublet field ϕ is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}. \tag{2.21}$$

In the SM the Lagrangian of the ϕ field is written as:

$$\mathcal{L}_{Higgs} = \overbrace{(D_\mu\phi)^\dagger(D^\mu\phi)}^{kinetic} - \overbrace{(\mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2)}^{potential} \quad (\mu^2, \lambda)\epsilon\mathbb{R}_{>0}. \tag{2.22}$$

The minimum of the potential term in Equation (2.22) is at $\phi^\dagger\phi = \mu^2/\lambda = v^2$, where v is the Vacuum Expectation Value (VEV). Therefore, to make the ground state of the ϕ field consistent with the minimum of the potential in $\mathcal{L}_{Higgs}$, it has to be redefined in terms of a scalar Higgs field H, which has a ground state at the VEV:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{2.23}$$

Expanding out the kinetic term in $\mathcal{L}_{Higgs}$ in terms of the physical bosons defined in (2.19) results in mass terms for the $W^{\pm}$ and Z bosons, but not for the photon:

$$(D_\mu\phi)^\dagger(D^\mu\phi) = \frac{v^2e^2}{4\sin^2\theta_W} W_\mu^+ W^{-\mu} + \frac{v^2e^2}{8\sin^2\theta_W\cos^2\theta_W} Z_\mu Z^\mu + (\text{kinetic and interaction terms}). \tag{2.24}$$

From Equation (2.24) the $W^{\pm}$ and Z boson mass terms can be extracted:

$$m_W = \frac{ve}{\sqrt{2}\sin\theta_W} \qquad m_Z = \frac{ve}{\sqrt{2}\sin\theta_W\cos\theta_W}. \tag{2.25}$$

Fully expanding the kinetic terms in the EW Lagrangian yields results in three types of fermion current. The EM current couples to the photon and is proportional to the EM charge Q of the fermion:

$$j^\mu_{EM} = Q\bar{\psi}\gamma^\mu\psi. \tag{2.26}$$

The charged current couples to the $W^{\pm}$ bosons:

$$j^{\mu+}_W = \frac{1}{\sqrt{2}}(\bar{\nu}^i_L\gamma^\mu l^i_L + \bar{u}^i_L\gamma^\mu d^i_L) \qquad j^{\mu-}_W = \frac{1}{\sqrt{2}}(\bar{l}^i_L\gamma^\mu \nu^i_L + \bar{d}^i_L\gamma^\mu u^i_L). \tag{2.27}$$

The neutral current couples to the Z boson and is proportional to the weak isospin T^3 component and EM charge of the fermion:

$$j^\mu_Z = \frac{1}{\cos\theta_W}(\bar{\psi}\gamma^\mu(T^3 - \sin^2\theta_W Q)\psi). \tag{2.28}$$

The strength of the coupling to j^μ_Z is different for the ψ_L and ψ_R, as the former can have a T^3 eigenvalue of $\pm\frac{1}{2}$ while for the latter it is always zero.

2.4.1 Fermion masses and quark flavour mixing

In the SM the fermion masses have to be set to zero before the $SU(2)_L \times U(1)_Y$ symmetry is broken. This is because the mass term $m\bar{\psi}\psi = m(\bar{\psi_L}\psi_R + \bar{\psi_R}\psi_L)$ is not invariant under the EW gauge transformation, as the left and right-handed components of the fermion fields transform differently under the $SU(2)_L \times U(1)_Y$ group. The Higgs mechanism is used to generate the fermion masses by introducing a Yukawa coupling [27] between the fermions and the ϕ field. An example of the Yukawa coupling for the electron is written as:

$$\mathcal{L}_{Y,e} = -Y_e(l_L^{1\dagger} \cdot \phi)e_R + h.c. \quad , \tag{2.29}$$

where Y_e is the Yukawa coupling constant of the electron and $l_L^{1\dagger}$ is the first generation lepton doublet, as defined in (2.16). By substituting the expression in (2.23) for ϕ after symmetry breaking, Equation (2.29) yields a term for the electron mass and its coupling to the Higgs boson:

$$\mathcal{L}_{Y,e} = -\overbrace{\frac{Y_e v}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)}^{\text{mass}} + \overbrace{\frac{Y_e H}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)}^{\text{interaction}}, \tag{2.30}$$

where the electron mass is $m_e = \frac{Y_e v}{\sqrt{2}}$. This mass generation mechanism can be applied to all SM fermions except the neutrinos, as they have no right-handed singlet expression and are therefore defined as massless in the SM. The complete Yukawa Lagrangian with all three generations considered is:

$$\mathcal{L}_Y = -Y_d^{ij}(\vec{Q}_L^{i\dagger} \cdot \phi)d_R^j - Y_u^{ij}(\vec{Q}_L^{i\dagger} \cdot \vec{\phi}_c)u_R^i - Y_l^{ij}(l_L^{i\dagger} \cdot \vec{\phi}_c)l_R^i + h.c. \quad , \tag{2.31}$$

where $\vec{\phi}_c = i\sigma^2\phi^*$, i and j denote the quark generation and $Y_{d,u,l}^{ij}$ are complex matrices. The Y^{ij} matrices can be diagonalised by applying a unitary transformation to the fermions:

$$\begin{aligned} u_L^i &\to (U_L^u)^{ij}u_L^j & \quad u_R^i &\to (U_R^u)^{ij}u_R^j \\ d_L^i &\to (U_L^d)^{ij}d_L^j & \quad d_R^i &\to (U_R^d)^{ij}d_R^j \\ l_L^i &\to (U_L^l)^{ij}l_L^j & \quad l_R^i &\to (U_R^l)^{ij}l_R^j, \end{aligned} \tag{2.32}$$

where U are unitary matrices defined such that $U_L^\dagger Y U_R$ is a diagonal matrix. The EM and neutral currents, shown in Equations (2.26) and (2.28) are invariant under the transformations defined in (2.32). For example, the transformation of the left-handed component of the lepton EM current transforms as:

$$\bar{l}_L^i\gamma^\mu l_L^i \longrightarrow \bar{l}_L^i(U_L^{l\dagger})^{ij}\gamma^\mu(U_L^l)^{jk}l_L^k = \bar{l}_L^i\gamma^\mu(\overbrace{U_L^{l\dagger}U_L^l}^{=\mathbb{I}})^{ik}l_L^k = \bar{l}_L^i\gamma^\mu l_L^i. \tag{2.33}$$

As a result Flavour Changing Neutral Current (FCNC) processes are forbidden at the tree level in the SM. These can only occur via loop diagrams and as such are strongly suppressed. The quark charged current interaction, $J_W^{\mu\pm}$ in (2.27), does allow flavour to change at the tree level, as it is not invariant under the U transformations:

$$\bar{u}_L^i \gamma^\mu d_L^i \longrightarrow \bar{u}_L^i \gamma^\mu (U_L^{u\dagger} U_L^d)^{ik} d_L^k = \bar{u}_L^i \gamma^\mu (V_{\text{CKM}})^{ik} d_L^k, \tag{2.34}$$

where V_{CKM} is a complex 3×3 unitary matrix, called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The individual components of the CKM matrix define the relative strength by which the different quark generations couple to each other through the charged current:

$$j_W^{\mu+} = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma^\mu \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \qquad j_W^{\mu-} = (j_W^{\mu+})^\dagger. \tag{2.35}$$

The magnitudes of the individual components of V_{CKM} can be measured by observing processes which partake in the currents detailed in Equation (2.35), for example, $|V_{bc}|$ can be obtained by measuring the rate of the decay $B^- \to \overline{D} e^- \overline{\nu}_e$. The current average values of the CKM amplitudes are measured to be [1]:

$$\begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}. \tag{2.36}$$

The CKM matrix can be expressed in terms of four independent parameters. A commonly used scheme is the Wolfenstein parameterisation [28], where the matrix is written in terms of the parameters A,λ,ρ and η:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.37}$$

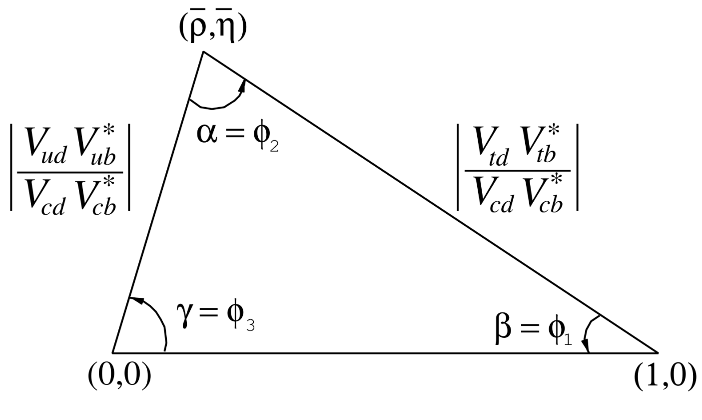


Figure 2.2: A visual representation of the CKM unitarity triangle on a complex plane, taken from [1]. The parameters shown are defined in (2.38)

Experimentally confirming whether V_{CKM} is a unitary matrix is an important test of whether the SM offers a complete description of flavour mixing. One requirement of unitarity is:

$$V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0. \tag{2.38}$$

This relation can be represented visually on a complex plane, as shown in Fig. 2.2, where the angles and the ρ,η parameters are defined as:

$$\begin{aligned} \alpha &= \arg\left(\frac{V_{td}V_{tb}^{*}}{V_{ud}V_{ub}^{*}}\right) \quad \beta = \arg\left(\frac{V_{cd}V_{cb}^{*}}{V_{td}V_{tb}^{*}}\right) \quad \gamma = \arg\left(\frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}}\right) \\ \bar{\rho} &= \rho(1-\frac{\lambda^2}{2}) + \mathcal{O}(\lambda^4) \qquad \bar{\eta} = \eta(1-\frac{\lambda^2}{2}) + \mathcal{O}(\lambda^4). \end{aligned} \tag{2.39}$$

The current experimental constraints on the unitarity parameters in (2.39) are set at [1]:

$$\bar{\rho} = 0.131^{+0.026}_{-0.013} \quad \bar{\eta} = 0.345^{+0.013}_{-0.014} \quad \gamma = (68^{+10}_{-11})^{\circ} \quad \beta = (21.4\pm0.8)^{\circ} \quad \alpha = (89.0^{+4.4}_{-4.2})^{\circ}, \tag{2.40}$$

these constraints are shown in Fig. 2.3.

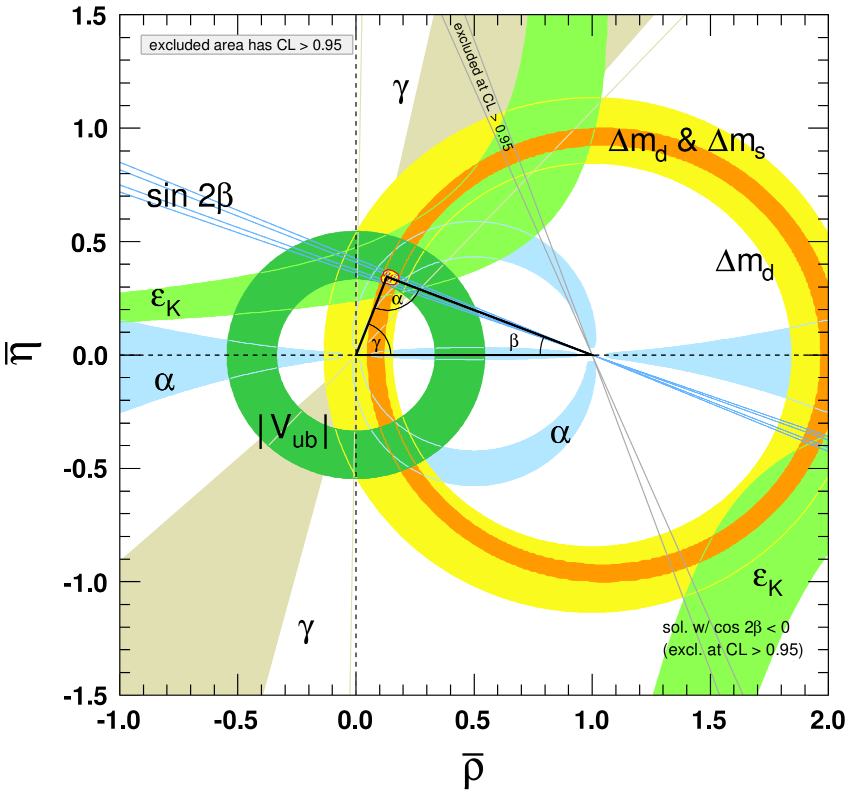


Figure 2.3: Current experimental constraints on the CKM unitarity triangle, as of May 2014, taken from [1].

$\mathcal{CP}$ violation

The CKM matrix contains a complex phase. This leads to the violation of the combined $\mathcal{CP}$ symmetry, where the $\mathcal{CP}$ operator exchanges particles with their antiparticles and inverts the spatial co-ordinates of a physical system, such that $\psi(\vec{x},t) \to \bar{\psi}(-\vec{x},t)$. The violation of $\mathcal{CP}$ results in asymmetries between certain processes and their $\mathcal{CP}$ conjugates. The combined symmetry of $\mathcal{CP}$ with $\mathcal{T}$ (time reversal, where $t \to -t$) is required to be invariant in a locally invariant Quantum Field Theory (QED), such as the SM. No $\mathcal{CPT}$ violation has been experimentally observed at the time of writing. In the SM there are three different mechanisms for $\mathcal{CP}$ violation.

$\mathcal{CP}$ violation in decay, also called 'direct' $\mathcal{CP}$ violation, results in the amplitude of the decay of a meson to a final state, A_f, being different from the amplitude for the $\mathcal{CP}$ conjugate process, such that $|A_f| \neq |\bar{A}_{\bar{f}}|$. Direct $\mathcal{CP}$ violation occurs in decays where A_f is the sum of at least two individual amplitudes which have different $\mathcal{CP}$-invariant 'strong' phases originating from QCD processes and also different $\mathcal{CP}$-odd 'weak' phases resulting from charged-current components of the diagrams. The interference between these amplitudes causes the $\mathcal{CP}$ violation. An example of direct $\mathcal{CP}$ violation is the difference between the decay rates of $B^0_{(s)} \to K^+\pi^-$ and $\overline{B}_{(s)} \to K^-\pi^+$ as defined by $A_{CP}(B^0_{(s)} \to K^+\pi^-)$:

$$A_{CP}(B^0_{(s)} \to K^+\pi^-) = \frac{\Gamma(\overline{B}_{(s)} \to K^-\pi^+) - \Gamma(B^0_{(s)} \to K^+\pi^-)}{\Gamma(\overline{B}_{(s)} \to K^-\pi^+) + \Gamma(B^0_{(s)} \to K^+\pi^-)}, \quad (2.41)$$

which has been measured by LHCb to be [29]:

$$A_{CP}(B^0 \to K^+\pi^-) = 0.080 \pm 0.007(\text{stat}) \pm 0.003(\text{syst}) \quad (2.42)$$
$$A_{CP}(B^0_s \to K^+\pi^-) = 0.27 \pm 0.04(\text{stat}) \pm 0.01(\text{syst}). \quad (2.43)$$

$\mathcal{CP}$ violation in mixing, or 'indirect' $\mathcal{CP}$ violation occurs in neutral flavoured mesons such as kaons and $B^0_{(s)}$ mesons. The flavour eigenstates of these mesons ($|M^0>$, $|\bar{M}^0>$) are not the same as their mass eigenstates ($|M^0_{H,L}>$) due to quark mixing via the CKM matrix. The latter are expressed as a superposition of the former:

$$|M^0_H>= p|M^0> +q|\bar{M}^0> \qquad |M^0_L>= p|M^0> -q|\bar{M}^0>, \quad (2.44)$$

where p and q are complex coefficients that satisfy $|p^2| + |q^2| = 1$. The probability of measuring one of the two flavour eigenstates varies with time as the mass eigenvector propagates through space. When $|q/p| \neq 1$ the rates of oscillation are different for the two flavour states, such that $\Gamma(M^0 \to \bar{M}^0) \neq \Gamma(\bar{M}^0 \to M^0)$. As a result the meson will exhibit a time-integrated preference for being detected as one flavour eigenstate rather than its $\mathcal{CP}$-conjugate. An example of the consequences of indirect $\mathcal{CP}$ violation is the asymmetry between the branching fractions of the $K^0_L \to \pi^-\mu^+\nu_\mu$ and $K^0_L \to \pi^+\mu^-\bar{\nu}_\mu$ decays, measured to be [1]:

$$A_L(\mu) = \frac{\Gamma(\pi^-\mu^+\nu_\mu) - \Gamma(\pi^+\mu^-\bar{\nu}_\mu)}{\Gamma(\pi^-\mu^+\nu_\mu) + \Gamma(\pi^+\mu^-\bar{\nu}_\mu)} = (30.4 \pm 2.5) \times 10^{-4}. \quad (2.45)$$

$\mathcal{CP}$ violation can also occur through the interference between mixing and decay in processes where M^0 and $\bar{M}^0$ can decay into the same final state f directly or by oscillation and subsequent decay, such that

$$\Im\left(\frac{q}{p} \times \frac{A_f}{\bar{A}_f}\right) \neq 0. \quad (2.46)$$

This form of $\mathcal{CP}$ violation has been observed in the $B^0 \to J/\psi\, K^0_\mathrm{s}$ decays [30].

2.5 Limitations of the SM

The SM has proved to be exceptionally resilient to direct experimental tests of its predictions. However, it only offers a description of a limited set of observed phenomena and has fundamental theoretical shortcomings that prevent it from offering a complete description of nature. Some of the main limitations of the SM are described below.

The neutrinos are massless in the most basic formulation of the SM, as described in section 2.4.1, which is in direct contradiction to the experimental observation of neutrino oscillation [31]. When a neutrino of a given flavour is produced, there is a probability that it will be detected as a different flavour of neutrino after propagation. This oscillation means that the neutrinos have different flavour and (non-zero) mass eigenstates, which mix in an analogous way to the CKM mechanism in the SM. The SM has to be amended to describe massive neutrinos.

The SM particles account for only 4.9% of the observed mass-energy of the universe, 26.8% of which is made of dark matter [32]. Dark matter is composed of weakly interacting non-SM particles, the presence of which is inferred from their gravitational interaction with matter. The remaining 68.5 % is the energy-density associated with the acceleration of the rate of expansion of the universe, again the SM provides no insight into the origin or cause of this expansion.

The gravitational force is described by Einstein's theory of General Relativity, which is incompatible with the SM as it is not a quantum theory. A quantum theory of gravity (Quantum Gravity) is not required to describe phenomena occurring at the $\mathcal{O}(\text{ TeV})$ energy scales being probed by current collider experiments, as its strength is negligible compared to the other three forces. At the Planck energy scale $E_p = 2.4 \times 10^{15}\,\text{TeV}$ the strength of gravity becomes non-negligible, therefore a theory of Quantum Gravity is needed to understand phenomena occurring inside the event horizon of black holes or the universe within $\sim 10^{-43}$ seconds after the big bang.

The inferred ratio of antimatter to matter in the universe is $\mathcal{O}(10^{-9})$, the SM mechanism of CP violation via the CKM matrix yields a ratio of just $\mathcal{O}(10^{-20})$. This suggests that the dominant source for the asymmetry is through non-SM mechanisms.

Calculating the beyond tree-level contributions to the Higgs mass term results in divergent terms. To cancel out these divergences the parameters of the Higgs sector have to be tuned by hand to a very high degree, this approach is referred to as 'fine-tuning'. Although the predictive power of the SM is not directly compromised by fine-tuning, it is often regarded as an un-natural way to address the divergences in the theory. This problem is referred to as the 'hierarchy' problem.

2.6 Supersymmetry

Supersymmetry (SUSY) introduces a symmetry between fermions and bosons via a fermionic operator Q:

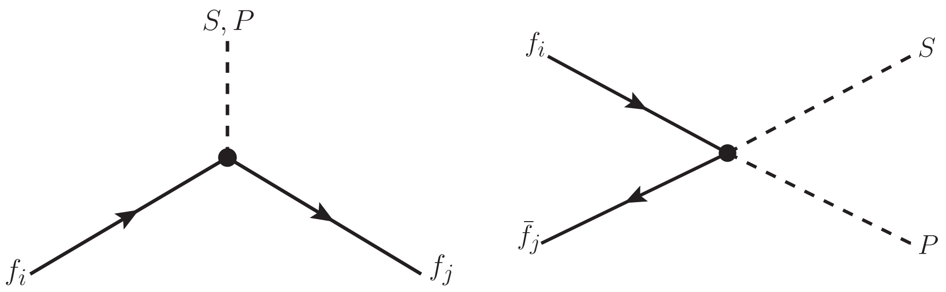
$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle \qquad Q|\text{Fermion}\rangle = |\text{Boson}\rangle, \tag{2.47}$$

where Q obeys the following commutation and anti-commutation relations:

$$\{Q, Q^\dagger\} = P^\mu \qquad \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0 \qquad [P^\mu, Q] = [P^\mu, Q^\dagger] = 0, \tag{2.48}$$

where P^μ is the generator of space-time translations. Extending this symmetry to the SM particles requires the introduction of a set of 'superpartner' particles to the SM particles. The simplest supersymmetric extension of the SM is called the Minimal Supersymmetric Standard Model (MSSM). In this framework each chiral fermion has its own unique 'superpartner' particle; for example the left and right handed top quarks each have their own spin-0 'stop' superpartner particle, while spin-1 bosons have fermionic superpartners. If SUSY were to be an unbroken symmetry, the superpartners would be of the same mass as their SM counterparts, but as no superpartner has yet been observed the symmetry has to be broken. In the MSSM the SUSY symmetry is broken explicitly. One can also construct scenarios where the symmetry is broken spontaneously, in a fashion analogous to EW symmetry breaking.

SUSY could potentially address some of the limitations of the SM, as stated in Section 2.5. The lightest superpartner particle fulfills the basic properties required of a dark matter particle. The superpartner contributions to the Higgs mass naturally cancel out the divergent SM contributions, removing the need for fine-tuning. However, since the superpartners are not of the same mass as their SM counterparts the cancellation is not perfect so some fine-tuning is still required, though to a much lesser extent than in the SM. The SUSY framework also allows for the natural inclusion of General Relativity, thereby providing a path towards building a theory of Quantum Gravity.



(a) (b)

Figure 2.4: Feynman diagrams illustrating type I (a) and type II (b) couplings between S and P sgoldstinos and massive SM fermions f (up-type quarks, down-type quarks, charged leptons), where i, j denotes the fermion flavour.

2.6.1 Sgoldstinos

Goldstone's theorem states that new massless particles called 'Goldstone bosons' are produced for each spontaneously broken continuous symmetry [33]. In the SM case of EW symmetry breaking three spin-0 Goldstone bosons are produced. These provide the longitudinal polarization components of the $W^{\pm}$ and Z bosons to give them mass. Spontaneously broken SUSY yields a massless spin-$\frac{1}{2}$ goldstino $\tilde{G}$, the superpartner of which is a complex scalar field $\phi_{\tilde{G}} = \frac{1}{\sqrt{2}}(S + iP)$. The real components of $\phi_{\tilde{G}}$ are the massive scalar S and pseudoscalar P sgoldstinos. The sgoldstinos can couple to all massive SM fermions to produce neutral currents, including ones that mix between different fermion flavours, resulting in tree-level FCNC processes via two different types of coupling. For type I couplings, either the S or P couples to a fermion pair in a three-pronged vertex. In type II couplings, both the S and P couple to a fermion pair in a four-pronged vertex. Both coupling types are also shown in Fig. 2.4.

The large variety of possible sgoldstino processes allows for a rich phenomenology. The FCNC processes in particular are ideal to probe experimentally as these allow for the enhancement of decays which would otherwise be strongly suppressed in the SM. Stringent experimental limits have already been set on the type I couplings by searches for such decays that are rare in the SM. The HyperCP collaboration searched for the decay $\Sigma^+ \to p\mu^+\mu^-$ [34], which has an expected SM branching fraction of $(1.6 - 9.0) \times 10^{-8}$ [35]. Three signal events of the decay were observed, with a branching fraction of $\mathcal{B}(\Sigma^+ \to p\mu^+\mu^-) = (8.6^{+6.6}_{-5.4}(\text{stat}) \pm 5.5(\text{syst})) \times 10^{-8}$, which is consistent with the SM expectation. However, a peaking structure was found in the dimuon invariant mass spectrum of the three events at $214.3 \pm 0.5\,\text{MeV}/c^2$, HyperCP calculated a 0.8% probability that the SM decay of $\Sigma^+ \to p\mu^+\mu^-$ would yield such a structure. The helicity structure of the decay allows for the interpretation that it was mediated by either an S or P sgoldstino with mass $214.3 \pm 0.5\,\text{MeV}/c^2$: $\Sigma^+ \to pX(\to \mu^+\mu^-)$ $(X = P, S)$, via a type I $s \to dX$ quark transition, as shown in Fig. 2.4(a). The HyperCP result is not statistically significant enough to strongly favour either the SM or sgoldstino hypothesis. The E865 collaboration observed 430 events of the decay $K^+ \to \pi^+\mu^+\mu^-$ [36] with a branching fraction of $9.22 \pm 0.60(\text{stat}) \pm 0.49(\text{syst})) \times 10^{-8}$, no peaking structure in the low dimuon mass spectrum was found, which sets an upper limit for the branching fraction at 95% confidence level (CL): $\mathcal{B}(K^+ \to \pi^+ S) \lesssim 8.7 \times 10^{-9}$ [37], where S has a mass of $214.3\,\text{MeV}/c^2$. This excludes the scalar S sgoldstino interpretation of the HyperCP result, but not the pseudoscalar P mode as the helicity structure of $K^+ \to \pi^+\mu^+\mu^-$ allows only for a scalar resonance to mediate the decay. Another complementary study was performed by the Belle collaboration which searched for the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to \rho^0\mu^+\mu^-$ in a dimuon invariant mass range of $212\,\text{MeV}/c^2 < m_{\mu^+\mu^-} < 300\,\text{MeV}/c^2$ [38], these decays would be sensitive to the type I $b \to dX$ quark transition. No resonant structure was found and 90% CL limits were set on the branching fractions: $\mathcal{B}(B^0 \to K^{*0}X(\to \mu^+\mu^-)) < 2.26 \times 10^{-8}$ and $\mathcal{B}(B^0 \to \rho^0 X(\to \mu^+\mu^-)) < 1.73 \times 10^{-8}$ where X has a mass of $214.3\ \text{MeV}/c^2$.

Type II couplings have recently been explored by the LHCb collaboration through the search for the decays $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to \mu^+\mu^-\mu^+\mu^-$ [39]. These decays can be propagated via $B^0_{(s)} \to SP$, where both S and P decay into muon pairs, making them sensitive to both type I and type II couplings. This search is detailed in Sect. 6.

Chapter 3

The LHCb detector

3.1 The Large Hadron Collider

The LHCb detector forms part of a broad experimental particle physics programme based around the Large Hadron Collider (LHC), a proton-proton ($p-p$) collider[1] located at the European Organization for Nuclear Research (CERN). The other main detectors around the LHC are ATLAS and CMS, which are general purpose detectors, and ALICE, which specialises in lead ion collisions. The LHC is housed in the circular tunnel with a circumference of 27 km and mean depth of 100m. The tunnel previously contained the Large Electron-Positron Collider (LEP) which was dismantled in 2000. The protons are accelerated through a series of smaller accelerators, shown in Fig. 3.1, before they are fed into the LHC. The first stage of the accelerator chain is the Linac 2 linear accelerator. Hydrogen gas is fed in at one end of the accelerator, where the gas is ionised to extract the protons, which are then accelerated by Radio Frequency (RF) cavities to an energy of 50 MeV. The Linac 2 protons are fed into the Proton Synchrotron Booster (PSB), which is composed of four superimposed synchrotron rings that accelerate the protons to 1.2 GeV and inject them into the Proton Synchrotron (PS). The PS increases the proton energy to 25 GeV. It has been operating in CERN since 1959, famously supplying the neutrino beam to the Gargamelle detector where the electroweak neutral current was discovered in 1973 [40]. The protons from the PS are fed into the Super Proton Synchrotron (SPS), which was the highest energy accelerator at CERN from 1976 to 1989. It functioned as a $p-\bar{p}$ collider for the UA1 and UA2 experiments, leading to the discovery of the $W^{\pm}$ and Z bosons [41,42]. The SPS accelerates the protons from the PS to 450 GeV before injecting them into the LHC, where they are accelerated to the final beam energy, which was 3.5 TeV in 2011 and 4 TeV in 2012. A 'consolidation upgrade' of the collider is taking place in 2013-2014, after which the beam energy will be increased to 7 TeV. A brief summary of the LHC running conditions employed to date is shown in Table 3.1.

[1] The LHC can also be configured for lead-lead and lead-proton collisions

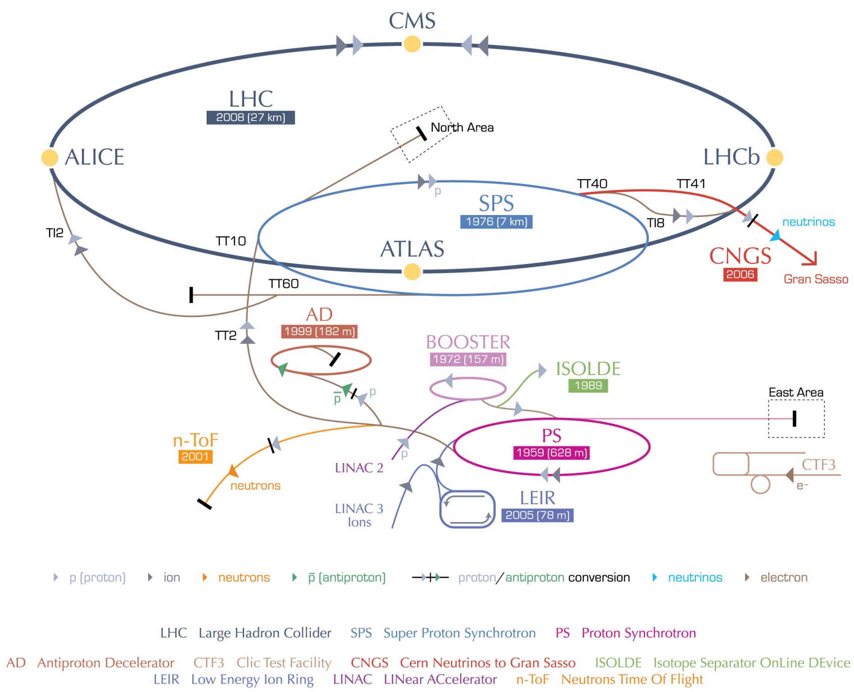


Figure 3.1: An overview of the accelerators based at CERN as of December 2008, taken from [2]

Table 3.1: Summary of the LHC running conditions over the dominant data collection periods in 2011 and 2012.

	$\sqrt{s}$ (TeV)	$\mathcal{L}$ $(10^{33}cm^{-2}s^{-1})$	No. of bunches	bunch spacing (ns)
2011	7	~ 2	1380	50
2012	8	~ 7	1380	50
design	14	~ 10	2808	25

3.2 LHCb design and layout

The principal physics aim of the LHCb detector is to study processes involving b quarks, these are produced predominantly via the gluon fusion process $gg \to b\bar{b}$. The resulting $b\bar{b}$ pairs have trajectories which are close to the axis of the LHC beam, as shown in Fig. 3.2. This removes the need for the detector to have a full angular coverage. LHCb has a vertical(horizontal) angular coverage of $\sim 10-300(250)$ mrad with respect to the beam direction. The detector only covers one direction along the beamline, so as to utilise the available cavern space to build larger subdetectors to improve performance at the expense of not detecting the particles travelling in the other beam direction. The pp collisions produce a large shower of particles in the LHCb acceptance due to soft QCD processes. At LHCb a maximum of ~ 600 charged tracks can be reliably reconstructed in a single event. To ensure that the charged track occupancy does not exceed this number, the number of pp interactions per bunch crossing is kept between $1-2$. This is done by reducing the amount of focussing applied to the LHC proton beams prior to collision at

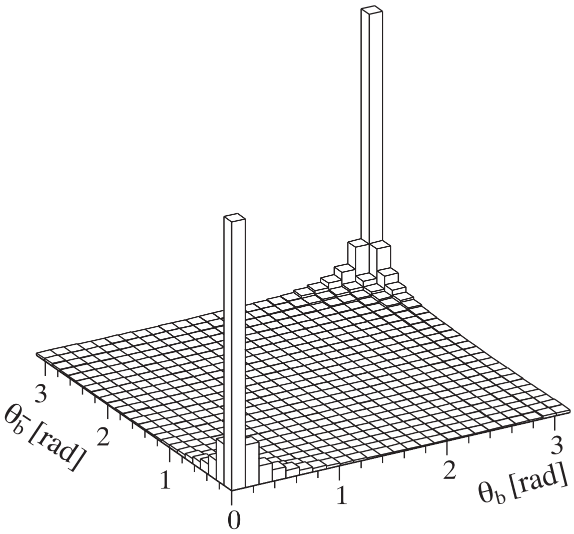


Figure 3.2: The angular distribution of $b-\bar{b}$ quark pairs with respect to the beam axis. Generated from simulated $p-p$ collisions at 14 TeV [3].

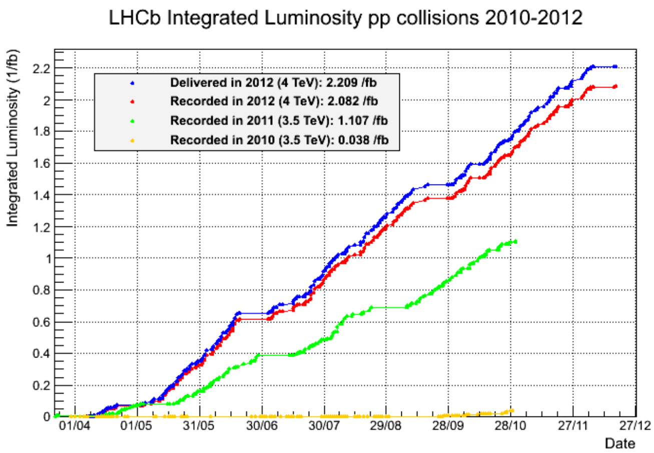


Figure 3.3: A diagram of the integrated luminosity collected by LHCb in 2010-2012, taken from [4].

LHCb, reducing the instantaneous luminosity to $(3-4)\times 10^{32}\,\text{cm}^2\text{s}^{-1}$. This is much less than the luminosity at CMS or ATLAS, as shown in Fig. 3.4. The reduction of luminosity is more than compensated by the large $pp \to b\bar{b}X$ cross section, which is measured to be $75\,\mu\text{b}$ within the LHCb acceptance at $\sqrt{s} = 7\,\text{TeV}$ [43]. The low number of pp collisions allows for very accurate association between particles and the pp interaction where they were produced. LHCb was collecting pp data throughout 2010-2013, with short technical stops used for quick maintainance of the detector and the LHC, the integrated luminosity collected during this period is shown in Fig. 3.3. The layout of the LHCb detector, with subdetector systems labelled, is shown in Fig. 3.5. The functions of main detector subsystems can be grouped into three categories: tracking, Particle Identification (PID) and calorimetery. There is some overlap between the functions of the subsystems, i.e. the calorimeter system also provides some PID information.

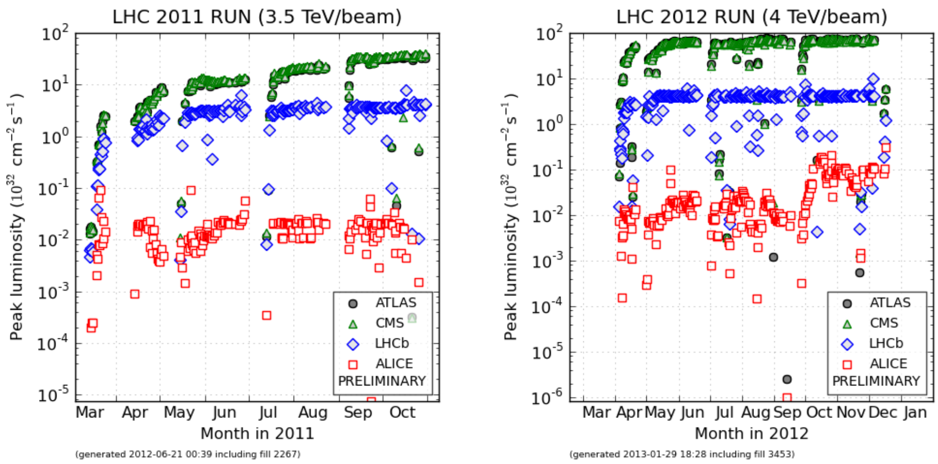


Figure 3.4: Plots showing the peak instantaneous luminosity at the four main LHC detectors in 2011 (left) and 2012 (right), taken from [5].

The tracking system

The purpose of the tracking system is to measure the momentum and trajectory of charged particles in LHCb, enabling the reconstruction and discrimination between primary and secondary decay vertices. The first component of the tracking system is the Vertex Locator (VELO), a high-precision tracking system made of silicon strips surrounding the pp interaction point which is described in more detail in section 3.3. The VELO's primary functions are to identify charged particle tracks and vertices from the Primary Vertex (PV), the pp interaction vertex, and, along with the Tracker Turicensis (TT) (a silicon strip tracker), to provide detailed tracking information upstream of the magnet. The magnet is a warm dipole magnet with a bending power of ~ 4 Tm that bends the trajectories of charged particles, allowing for measurement of their momentum. The magnetic field polarity can be flipped to 'up' or 'down'. Data are collected in roughly equal amounts for each polarity. Downstream of the magnet are three tracking stations, labelled as T1,T2 and T3 in Fig. 3.5. These are composed of an inner silicon strip tracker and an outer tracker which uses straw drift tubes. The momentum resolution is $\delta p/p \sim 0.4-0.6\%$ for tracks with momenta of 5-100 GeV/c^2

The PID system

The PID system provides information on the flavour of reconstructed particles. Kaons, pions and protons are identified in a momentum range of $2-100$ GeV/c by the two Ring Imaging Cherenkov Detector (RICH) detectors. Muons are identified with the five muon stations, labelled M1-M5 in Fig. 3.5. The RICH and muon systems are described in more detail in sections 3.4 and 3.5. Additional PID information is provided by the calorimeter system, to identify electrons and photons.

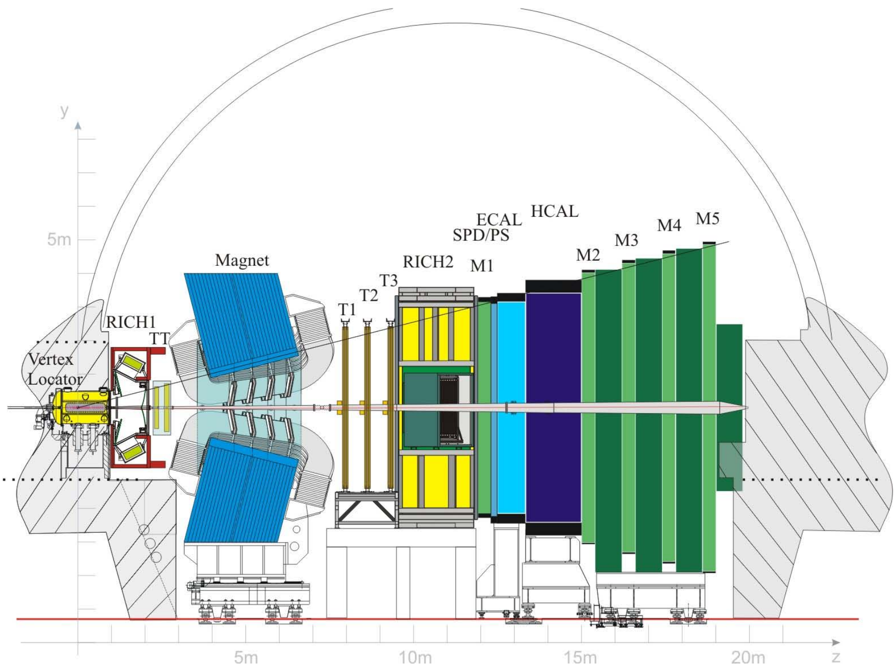


Figure 3.5: A cross-section of the LHCb detector in the y-z plane [6].

The calorimeter system

The calorimeter system measures the magnitude and location of energy deposited by charged and neutral particles. It is composed of the Scintillating Pad Detector (SPD), Pre-Shower Detector (PSD), Electromagnetic Calorimeter (ECAL) and Hadron Calorimeter (HCAL). The SPD is a segmented scintillator that is sensitive to the passage of single charged particles. It is followed by the PSD, composed of a 15mm thick lead plate, designed to induce electron and photon showers, followed by another scintillator. The main purpose of the SPD is to distinguish between photons and electrons. The ECAL is composed of alternating layers of lead plate and scintillating tiles. Charged tracks deposit energy in the ECAL, which is measured to a precision of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$ (E in GeV). The HCAL measures the energy of hadrons with a precision of $\sigma_E/E = 80\%/\sqrt{E} \oplus 10\%$ (E in GeV). The HCAL is composed of alternating layers of iron and scintillator tiles, which induce the hadrons to shower, measuring their energy.

3.3 The VELO

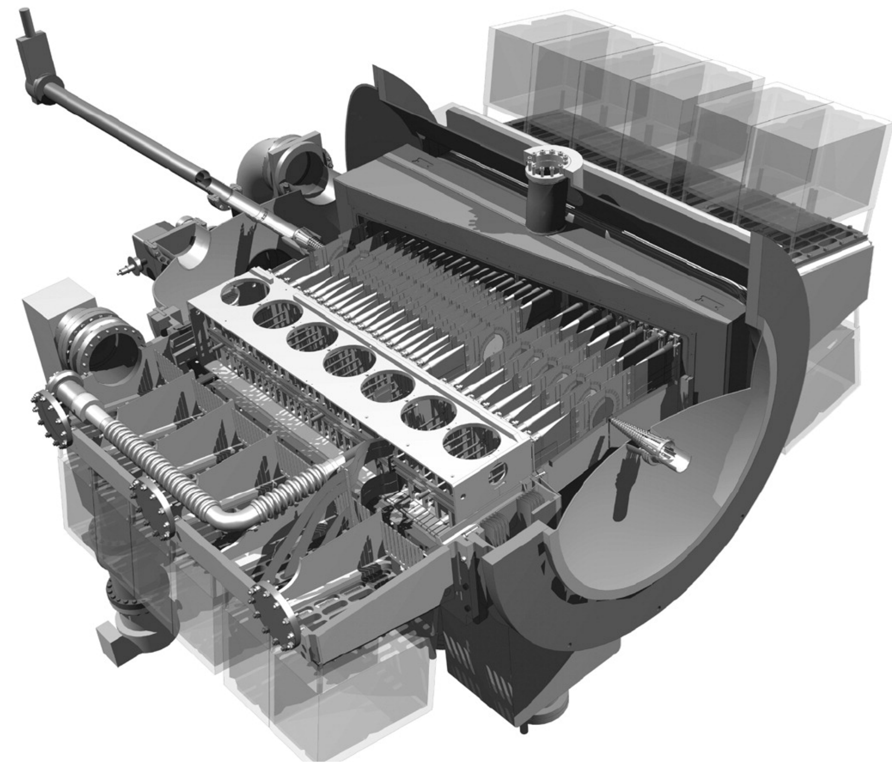


Figure 3.6: A CAD image of the VELO detector, including the vacuum vessel that houses the 21 modules and the beampipe. Taken from [7].

The VELO consists of 21 disc-shaped silicon tracker modules housed in a vacuum vessel. The layout of the VELO is shown in Fig. 3.6 and Fig. 3.7. The modules are composed of two partially overlapping semicircular discs which are separated from each other and the LHC beam by a 200 µm-thick aluminium foil, which preserves the LHC vacuum. When the proton beam is injected and accelerated in the LHC, the beam radius is greater than the 8 mm radius of the inner edge of the VELO disk. To prevent radiation damage the semicircular discs are retracted from each other horizontally by 3 cm. The VELO was exposed to a 1 MeV neutron-equivalent fluence of $(2.5-6.5)\times10^{13}$ per cm^2 per fb^{-1} in the innermost disc regions [44]. This is the harshest radiation condition experienced by any silicon detector in the LHC. To minimise radiation damage the modules are are maintained at a temperature of -10 to $0\,^\circ$ C. Each VELO module has 2048 silicon strips with a pitch ranging from $40-100\,\mu$m. The arrangement of the strips on each disc side is shown in Fig. 3.7, where the strips are aligned either to a constant radial angle ϕ or radial distance R with respect to the beam axis.

The precision in measurement of the Impact Parameter (IP), the perpendicular distance between a track and the PV, is $\sigma(IP)\sim 25\,\mu$m for tracks with a p_{T} of around 2 GeV. The $\sigma(IP)$ resolution improves with increasing track p_{T}, as shown in Fig. 3.8(a). Fig. 3.8(b) shows that the z-axis position of a PV reconstructed with 35-40 tracks (typ-

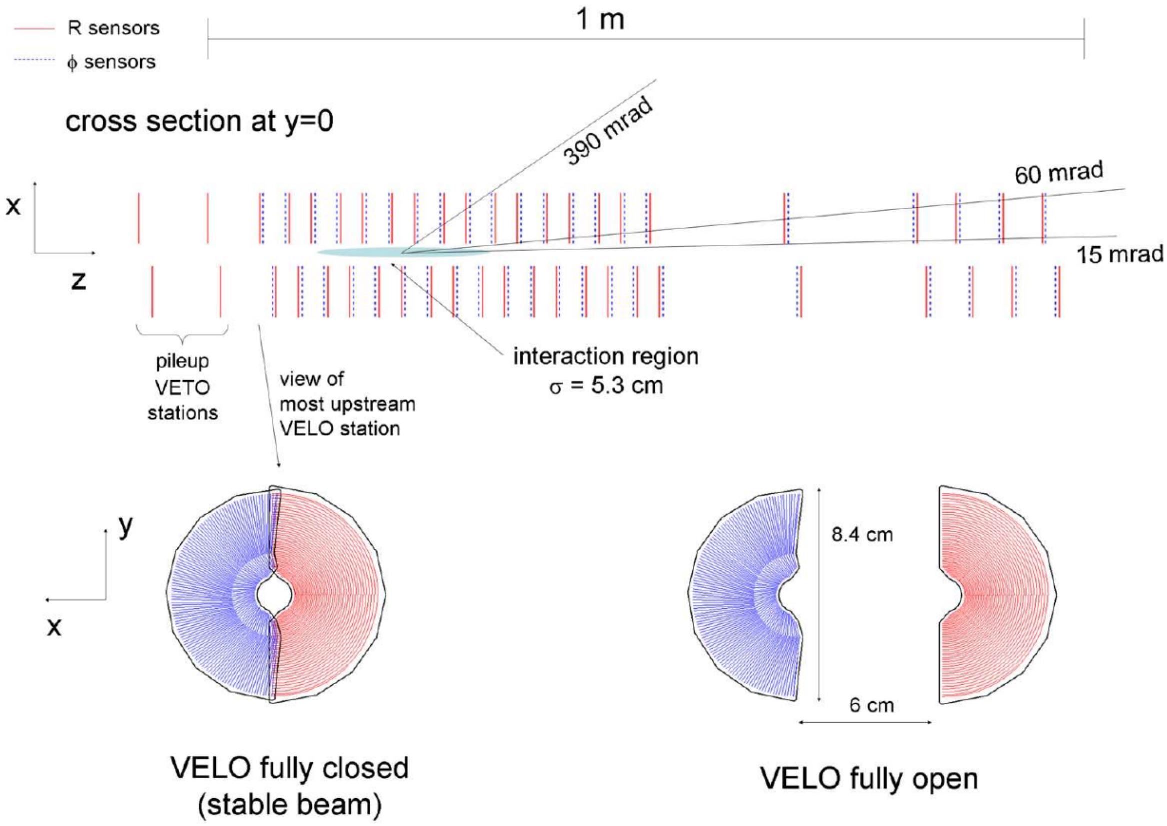
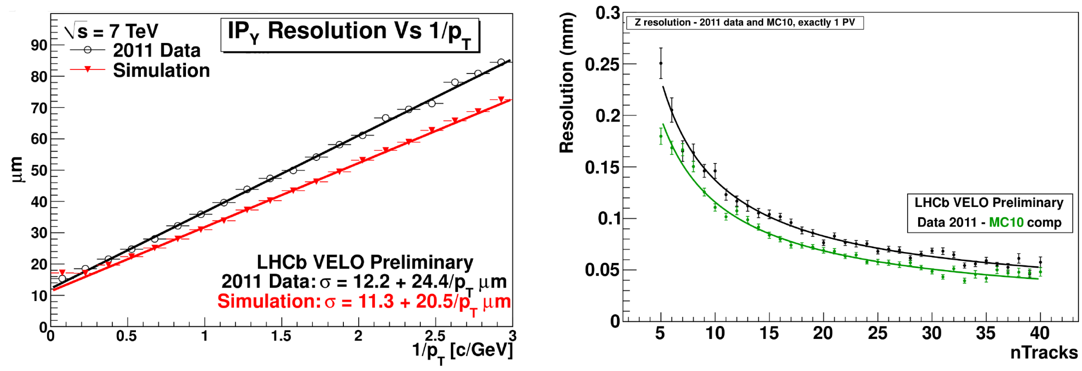


Figure 3.7: Cross-sectional diagrams of the VELO subdetector in the $x - z$ and $y - x$ planes [6].

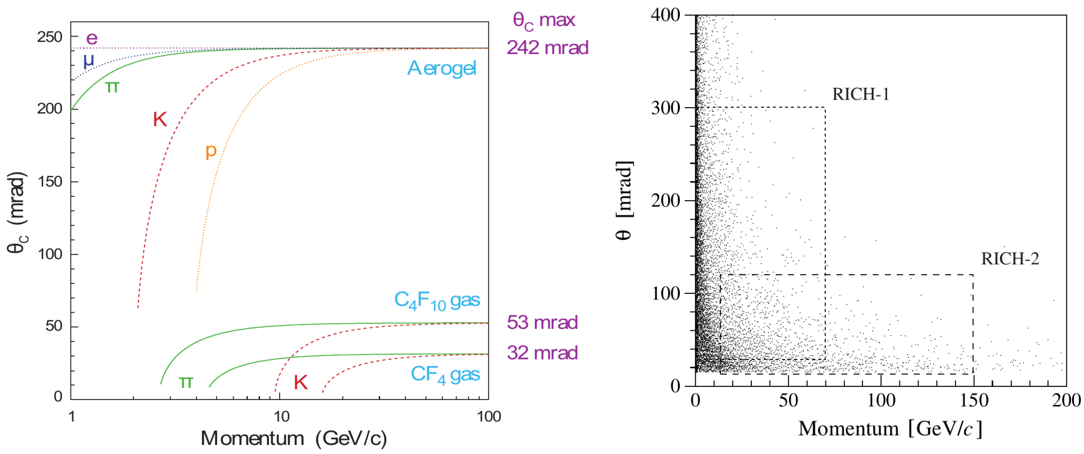
ical for a PV in 2011 data) can be measured to a precision of roughly $50 - 60\,\mu\text{m}$ for a Primary Vertex (PV). The VELO can measure the separation between the PV and a displaced Secondary Vertex (SV) very accurately. This allows for the measurement of important properties such as particle lifetime. For instance, the decay time resolution for $B_s^0 \to D_s^- \pi^+$ decays is 44 fs [45]. Accurate IP and vertex displacement measurements allow LHCb to distinguish between B meson decays and background processes very effectively, as a B meson typically travels $\sim 1\,\text{cm}$ in LHCb before decaying into lighter particles, which tend to have high IP as the B decay imparts transverse momentum to them.



(a) (b)

Figure 3.8: The performance of the VELO in 2011 compared with simulation. The dependence of the IP resolution on the inverse p_{T} of a track (a) and the z-position resolution of the PV as a function of the number of tracks (nTracks) used to reconstruct it (b). Both plots used events containing only one reconstructed PV [7].

3.4 The RICH system



(a) (b)

Figure 3.9: (a): A graph showing how the Cherenkov angle θ_C varies with particle momentum for different species of particle for the three RICH radiators, also shown are the saturation momenta for each radiator, taken from [6]. (b): A plot showing the track angle vs momentum coverage of the two RICH detectors, the particles in the distribution are pions from simulated $B^0 \to \pi^+ \pi^-$ decays at $\sqrt{s} = 14\,\mathrm{TeV}$, taken from [8].

When a charged particle travels through a dielectric medium at a speed faster than the phase velocity of light in the medium, a cone of light is emitted at an angle θ_C with respect to the particle's trajectory [46]. The Cherenkov angle θ_C is related to the velocity of the particle and n, the refractive index of the medium:

$$\cos\theta_C = \frac{1}{n\beta} \quad , \text{ where } \beta = \frac{v}{c} = \frac{pc}{\sqrt{m^2c^4 + p^2c^2}}. \tag{3.1}$$

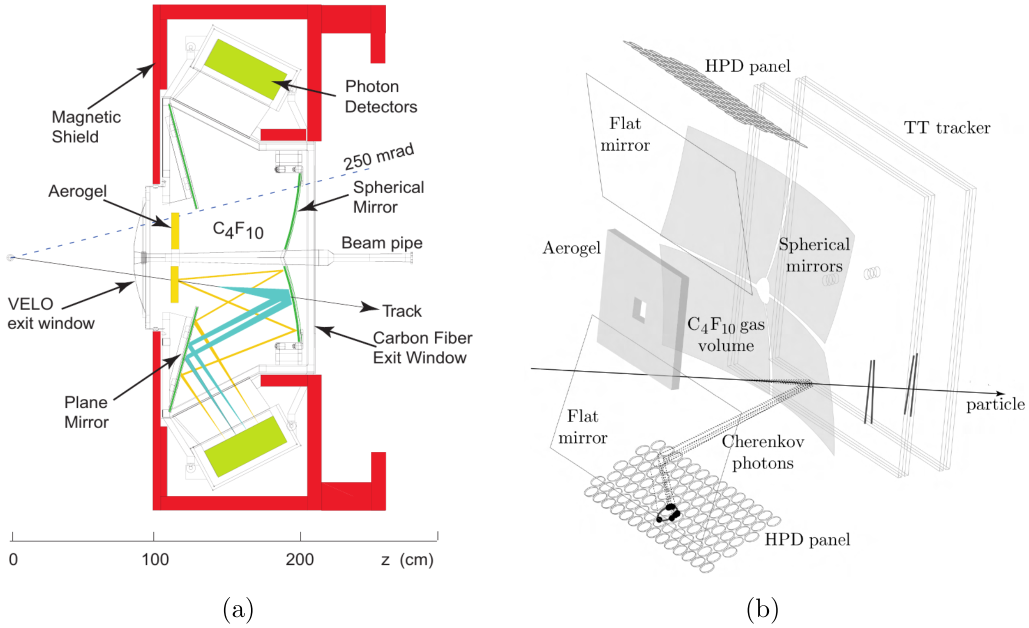


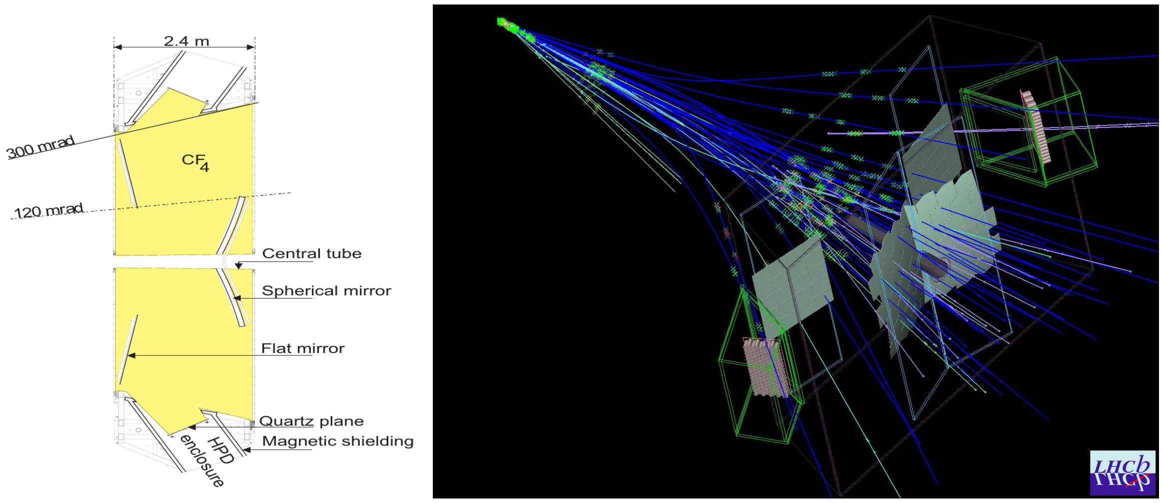
Figure 3.10: (a): A cross-section of the RICH1 detector in the $y-z$ plane, taken from [6]). (b): RICH1 with the TT tracker, taken from [9].

The mass of a given particle can therefore be calculated by measuring n, θ_C and p. The PID of a particle can be determined within a defined momentum range, the lower limit of which is the 'threshold' momentum, where $n\beta = 1$. Below this momentum no Cherenkov light is emitted. The upper limit is called the 'saturation' momentum, where $p >> mc$ and θ_C approaches the saturation angle

$$\theta_{sat} = \cos^{-1}(1/n) \tag{3.2}$$

Heavier particles and media with lower refractive indices increase both the threshold and saturation momenta. The purpose of the RICH at LHCb is to distinguish between kaons, pions and protons in a momentum range of 2-100 GeV/c, which is provided by three media called 'radiators', made of silica aerogel, C_4F_{10} gas and CF_4 gas. Fig. 3.9(a) shows a graph of θ_C vs. track momentum for these radiators. The radiators are housed in two separate subdetectors, RICH1 and RICH2, which provide a particle angular coverage shown in Fig. 3.9(b).

RICH1, shown in Fig. 3.10, contains the aerogel and C_4F_{10} radiators. It is located upstream of the TT and magnet, covering an angular acceptance of ±25 mrad to ±250 mrad (vertical) and ±300 mrad (horizontal). RICH1 provides PID for particles with momenta of $2-60$ GeV/c. The aerogel is composed of 16 tiles which are 50 mm thick, positioned at the front of RICH1. The refractive index of aerogel is 1.03, providing PID for particles with momenta less than 10 GeV/c. The C_4F_{10} gas is located behind the aerogel in RICH1, with an effective particle track path-length of 95 cm. It has a refractive index of 1.0014 at temperature $T = 0\ ^\circ$ C, pressure $P = 101.325$ kPa and radiation with a



(a) (b)

Figure 3.11: (a): A cross-section of the RICH2 detector in the $y-z$ plane, taken from [6]. (b): An event display showing the mirrors and photo-detectors of RICH2 with tracks originating from a simulated pp collision at 14 TeV, taken from [10].

wavelength of $\lambda = 400\,\text{nm}$.

RICH2, shown in Fig. 3.11, houses the CF_4 gas radiator, providing PID for particles in a momentum range of $15-100$ GeV/c. RICH2 is positioned downstream of the magnet, between the T3 tracking station and the SPD, as shown in Fig. 3.10. The angular coverage of RICH2 is ±15 mrad to ±100 mrad (vertical) and ±120 mrad (horizontal). The CF_4 has an effective particle track path-length of 180 cm and $n = 1.0005$ at $T = 0°\text{C}, P = 101.325\text{kPa}$ and $\lambda = 400\,\text{nm}$,.

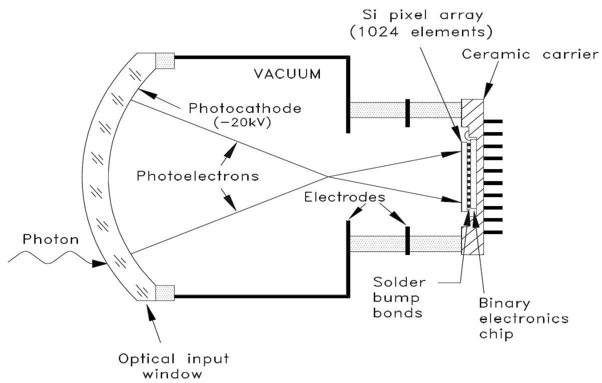


Figure 3.12: A schematic of an RICH HPD, taken from [6].

In both of the RICH detectors, charged particles traverse the radiator media and produce Cherenkov light, which is focused and reflected outside of the LHCb acceptance by a combination of spherical and flat mirrors onto planes of hexagonally arranged Hybrid Photon Detector (HPD) elements. Each RICH has two HPD planes, RICH1 and RICH2 employ 196 and 288 HPDs respectively. A schematic of an HPD is shown in Fig. 3.12. It consists of a vacuum tube with a quartz window, coated with a multialkali photocathode layer that converts incident photons into electrons (referred to as ‘photoelectrons‘). These are then accelerated across an 18 kV electric potential onto an array of 1024 silicon pixel detectors. Each pixel is 500 µm × 500 µm in size and has a quantum efficiency of $\sim 30\%$ at $\lambda = 270\,\text{nm}$. The pixels have a binary readout, which registers a hit if the charge deposited by one or more photoelectrons exceeds a preset threshold. The fringe fields of the LHCb dipole magnet can influence the trajectory of a photoelectron inside a HPD. To minimise this effect the HPD planes are surrounded by magnetic shielding which reduces the fringe field strength from 60 mT in RICH1 to < 1 mT and 15 mT in RICH2 to $0.2-0.6$ mT.

3.4.1 RICH reconstruction and performance

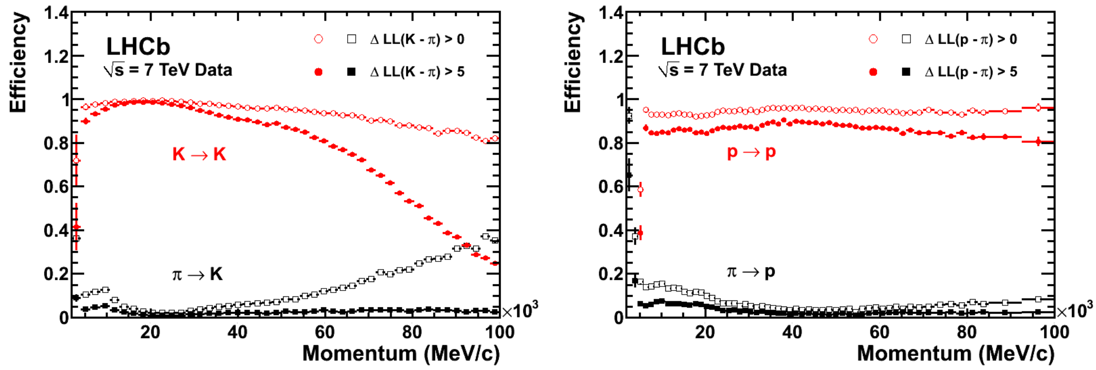
The RICH assigns a PID to a track using a ‘global pattern recognition’ algorithm [47]. For every recorded *pp* collision event the RICH software calculates the global likelihood for the observed distribution of HPD hits being consistent with an expected distribution with every charged track being a pion (the most abundant particle produced in hadronic collisions). The algorithm iterates through each track in turn and recalculates the global likelihood when the track PID hypothesis is changed to that of an electron, muon, kaon or proton (for electrons and muons additional information from the calorimeter and muon systems is also used). The hypothesis which maximises the likelihood is assigned to the track and the iteration process is continued until all tracks have been assigned hypotheses. The ‘strength’ of a given PID hypothesis is determined by the difference between the natural logarithm of its global likelihood and the pion likelihood:

$$DLL_{\alpha} = \ln(\mathcal{L}_{\alpha}) - \ln(\mathcal{L}_{\pi}) \qquad \alpha = e, \mu, K, p. \tag{3.3}$$

The greater the value of DLL_{α}, the more likely the α PID hypothesis is. Two non-pion hypotheses can be compared by combining different DLL variables:

$$DLL_{\alpha\beta} = \ln(\mathcal{L}_{\alpha}) - \ln(\mathcal{L}_{\beta}) \qquad \alpha, \beta = e, \mu, K, p. \tag{3.4}$$

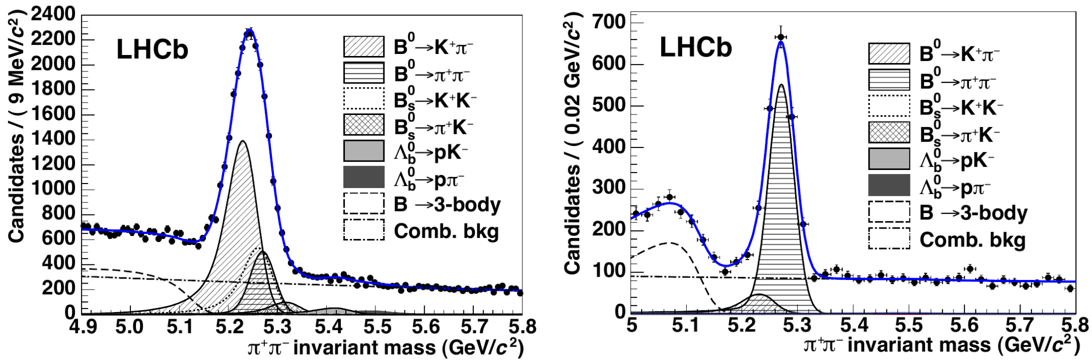
Figure 3.13(a) compares the probability for a kaon or a pion in 2011 data to pass a DLL_K cut. The kaon probability (correct hypothesis) is referred to as the ID efficiency and the pion probability (incorrect hypothesis) is referred to as the MisID efficiency. The mean kaon-pion ID and misID efficiencies are $\sim$ 95% and $\sim$ 10% respectively for a $DLL_K > 0$ cut, the equivalent rates for a tighter $DLL_K > 5$ cut are $\sim$ 85% and $\sim$ 3% respectively. The best performance is achieved in the momentum range of 10-50 GeV/c, where both C_4F_{10} and CF_4 radiators can distinguish between the two hypotheses. The performance degrades as momenta approach 100 GeV/c as the kaon approaches the



(a) (b)

Figure 3.13: Plots comparing the probability for a kaon (a) or a proton (b) being assigned the correct PID hypothesis (circular points) or the pion hypothesis (square points) in 2011 data for different momenta. The hollow and solid points compare the rates for different DLL cuts. The plots are produced using pions, kaons and protons from the PID calibration samples described in Appendix B. Taken from [11].

saturation momentum. The performance drops at momenta less than 10 GeV/c, where the kaon is below threshold for the gas radiators so it can only be identified via the absence of a pion ring. The aerogel performance is also not as good as for the gas radiators for reasons described in Sect. 4.4.3. The proton-pion PID separation is shown in figure 3.13(b), the performance is markedly better at high momentum than for kaon-pion PID because the proton saturation momentum is much greater than the kaon saturation momentum.

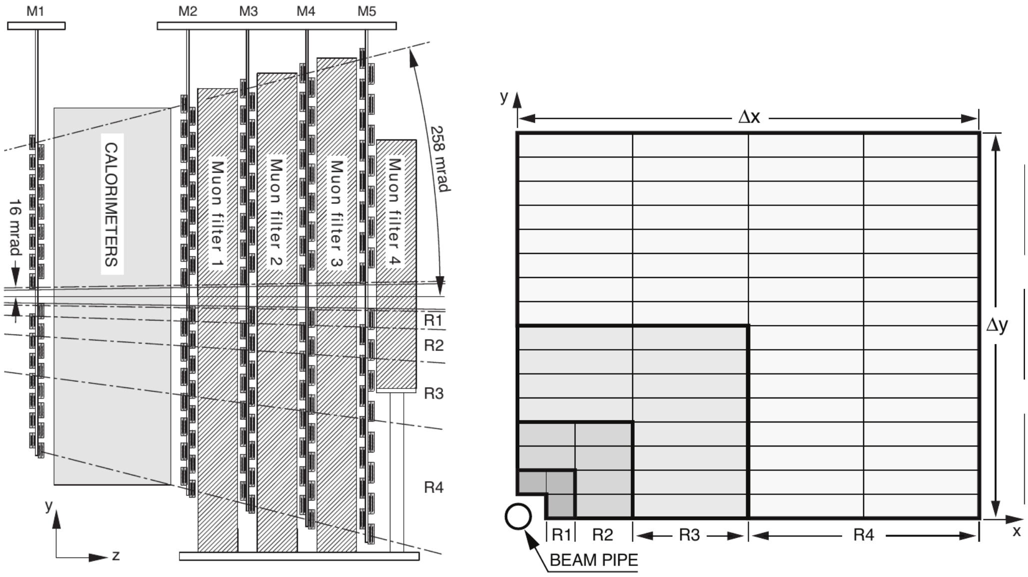


(a) (b)

Figure 3.14: Plots showing the $\pi^+ \pi^-$ invariant mass distributions obtained after applying a kinematic selection only (a) and applying kinematic and PID selections (b) to isolate a $B^0 \to \pi^+\pi^-$ decay [12].

The DLL cuts enable LHCb physics analyses to distinguish between kinematically similar decay modes with different hadronic final states, such as B^0 and B^0_s mesons decaying into h^+h^-, where $h = \pi, K$. Figure 3.14 shows how the application of DLL cuts can be used to isolate the $B^0 \to \pi^+\pi^-$ decay from the other two-body B decays, in particlular the much more numerous $B^0 \to K^+ \pi^-$ decays.

3.5 Muon System



(a) (b)

Figure 3.15: (a): Cross-sectional schematics of the muon system in the $y - z$ plane. (b) A quadrant of a muon station in the $x - y$ plane with the individual chambers shown. Taken from [12].

The muon system is composed of five stations covering an angular acceptance of 20-306 (16-258) mrad in the bending (non-bending) plane, denoted as M1-M5. Their layout relative to the calorimeter system is shown in Fig. 3.15(a). The station M1 is positioned in front of the calorimeter, with M2-M5 behind. The M2-M5 stations are separated by 80 cm-thick iron walls which absorb all charged particles except muons, which can penetrate all layers if their momentum exceeds 5 GeV/c. The stations are composed of multiple Multi-Wire Proportional Chambers (MWPCs) filled with $Ar - CO_2 - CF_4$ gas, except for the inner region of M1 which uses Gas Electron Multiplier (GEM) detectors, which are more resistant to radiation. Each muon station is split into 276 rectangular chambers. The granularity of the chambers increases at closer radial distance to the beampipe, as shown in Fig. 3.15(b). This enables the $p_{\rm T}$ of a muon to be measured by the muon system alone to a precision of $\sim 20\%$ across the entire angular acceptance, as is done at the hardware stage of the LHCb trigger.

Muon PID information is compiled into three different variables: IsMuon, muDLL and DLL_{μ}.

IsMuon is a boolean variable which is true for muon candidates. This is determined by defining a 'field of interest' around an extrapolated track trajectory through the muon

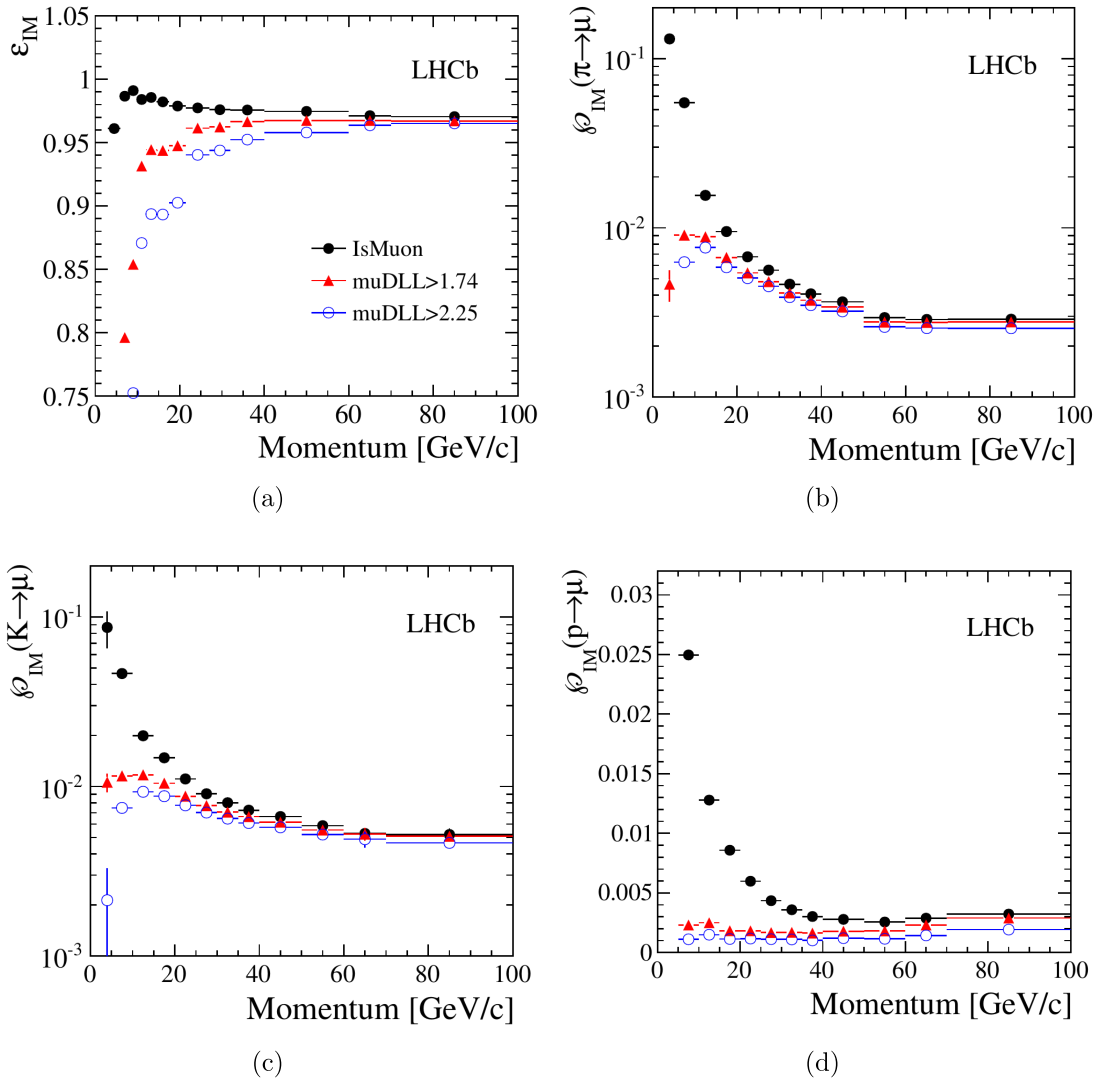


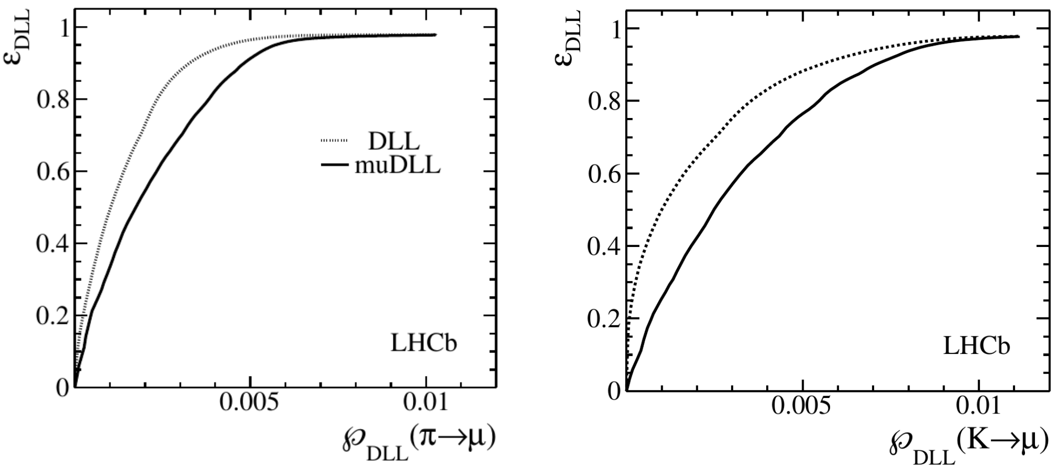
Figure 3.16: Plots showing (a) the muon ID efficiency and (b) the $\pi \to \mu$, (c) $K \to \mu$ and (d) $p \to \mu$ misID rates for different track momenta in 2011 data, after application of either an isMuon = true cut (black points) or a combination of isMuon = true and muDLL cuts (red and blue points, grey in B&W) [13].

chambers. The variable is set to true if muon chamber hits in multiple muon stations are found in the field of interest. Tracks with higher momenta require hits in more muon stations.

muDLL is the difference in the logarithm of the likelihood that the pattern of hits in the muon system is consistent with the extrapolated track being either a muon or a non-muon particle. The larger the value for muDLL the more 'muon-like' the track is.

DLL_{μ} is the difference in the log-likelihoods of the muon and pion PID hypotheses, where the likelihoods are calculated using information from the RICH, calorimeter and muon systems. The muon hypothesis can be compared to a non-pion hypotheses by combining with other DLL variables as shown in Eq. (3.4).

The PID performance is measured in data using the μ, π, K and p PID calibration



(a) (b)

Figure 3.17: Plots showing the muon ID efficiency ϵ_{DLL} against $\pi \to \mu$ (a) and $K \to \mu$ (b) misID rates when applying either a $muDLL$ or a DLL cut for 2011 data [13].

samples defined in Appendix B. The effectiveness with which the IsMuon and muDLL variables discriminate between these particles is shown in Fig. 3.16. The average isMuon ID efficiency for muons with $p > 3\,\text{GeV}/c$ and $p_{\text{T}} > 0.8\,\text{GeV}/c$ is $98.13 \pm 0.04\%$, the corresponding π, K and p misID rates are $1.025{\pm}0.003\%$,$1.111{\pm}0.003\%$ and $1.033{\pm}0.003\%$ respectively. Kaons and pions have greater misID rates than protons due to 'decays in flight', where π and K decay into muons inside the detector, which are subsequently identified by the muon system. The IsMuon performance drops at lower momentum due to an increased contribution from multiple scattering and because fewer muon stations are required to have matching hits to satisfy the IsMuon=true for tracks with $p < 10\,\text{GeV}/c$. Requiring additional muDLL criteria improves the performance further; even more so when DLL_{μ} is used (Fig. 3.17), as more information is used to discriminate between muon and non-muon hypotheses.

3.6 Trigger

In LHCb a sequence of hardware and software triggers select events containing decays of charm and beauty hadrons to be stored offline for use in physics analyses. The trigger system consists of a hardware stage, Level 0 (L0), followed by two software 'High Level Trigger' (HLT) stages, HLT1 and HLT2. Each level of the trigger is composed of multiple trigger 'lines' which select specific categories of final state particles and decays. Physics analyses can require events to be selected by a certain combination of trigger lines, i.e. an analysis searching for decays into muons would require the events to pass the muon specific trigger lines. The performance of the main trigger lines which select B meson decays with either muon or charged hadron final state particles are detailed in sections 3.6.1 to 3.6.3.

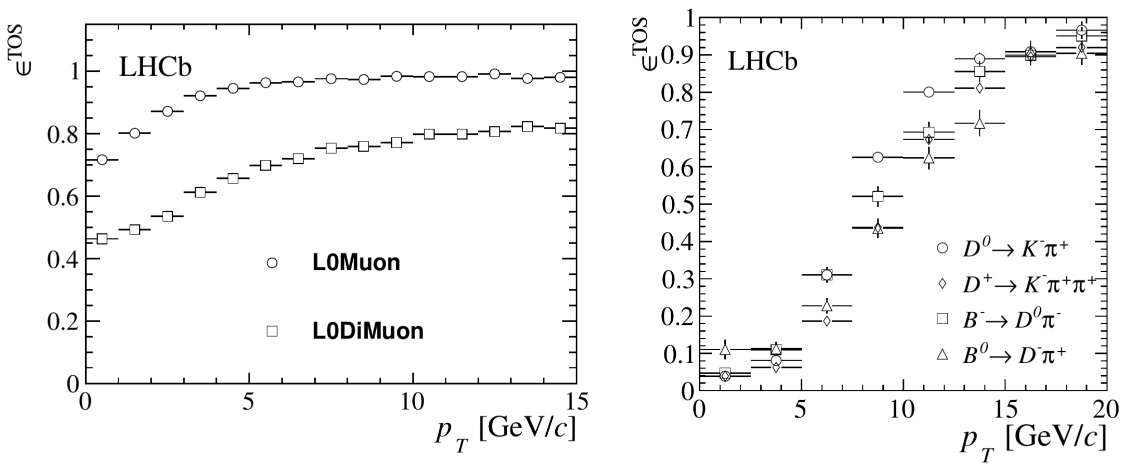
Given a decay of interest, there are three categories of trigger line decision: Trigger on Signal (TOS), Trigger Independent of Signal (TIS) and TIS&TOS. In a TOS decision, the trigger line conditions are satisfied exclusively by the decay and its final state particles, i.e. the event would be triggered if all other tracks in the event were removed except for the signal decay products. For a TIS decision, the event is triggered independently of the decay, i.e. if the decay products are removed from the event, the event would still be triggered. The event would not be triggered if the decay products are kept and the other tracks removed. The TIS&TOS decision defines an event that is triggered by both the signal and background particles, i.e. the event would not be triggered if either the signal decay products or the non-signal particles are removed from the event.

The performance of specific trigger lines is measured in data by calculating its TOS efficiency:

$$\epsilon^{TOS} = \frac{N^{TIS\&TOS}}{N^{TIS}}, \tag{3.5}$$

where N^{TIS} and $N^{TIS\&TOS}$ are the number of events that satisfy the TIS and TIS&TOS conditions, respectively.

3.6.1 L0



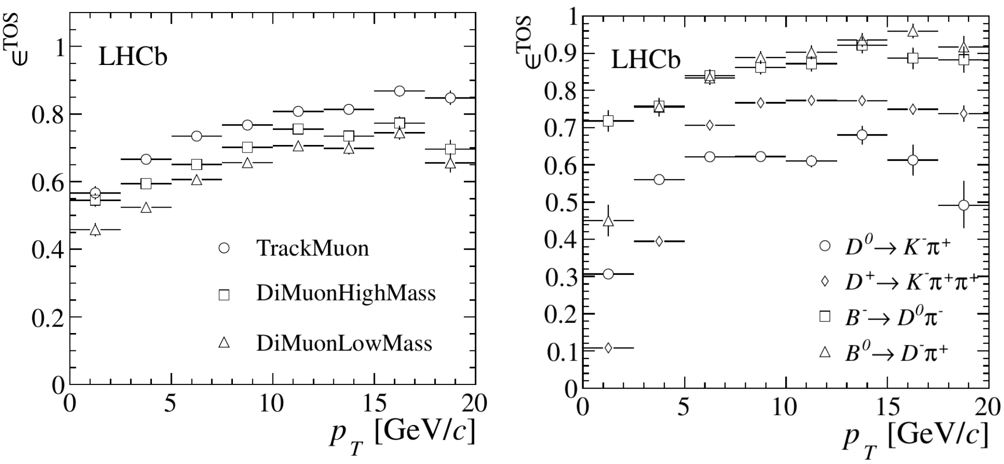
(a) (b)

Figure 3.18: Diagrams showing the TOS efficiency of the L0 muon lines (a) for $B^+ \to J/\psi$ $(\to \mu^+ \mu^-)$ K^+ decays and the L0 hadron line (b) for various B and D meson decays [14]. Both plots use 2011 data and show the dependency of the efficiency on the p_T of the J/ψ, B and D mesons.

The L0 is a hardware level trigger stage which reduces the LHC bunch crossing rate of 40 MHz to 1 MHz, at which the output of all the detector subsystems can be fully read out. In 2011 data collection conditions the L0 was configured to reduce the observed

pp interaction rate of $\sim 11\,\text{MHz}$, to $870\,\text{kHz}$. The L0 trigger uses information from the muon system to calculate the muon p_{T} and information from the calorimeters to extract the transverse energy E_T which is deposited in a cluster of the calorimeter. The L0 muon trigger requirements are satisfied if the event contains a muon with $p_{\text{T}} > 1.5\,\text{GeV}/c$ (L0Muon line) or the product of the two highest p_{T} muons is $\sqrt{p_{\text{T}}^{\text{largest}} \times p_{\text{T}}^{\text{2nd largest}}} > 1.3\,\text{GeV}/c$ (L0DiMuon line). Hadrons are selected by the L0Hadron line, which requires that the event contains an HCAL cluster with $E_T > 3.5\,\text{GeV}/c$. The performance of the L0 muon and hadron lines is shown in Fig. 3.18, the TOS efficiency of both increases with the p_{T} of the signal decay particles as more events exceed the trigger p_{T} thresholds.

3.6.2 HLT1

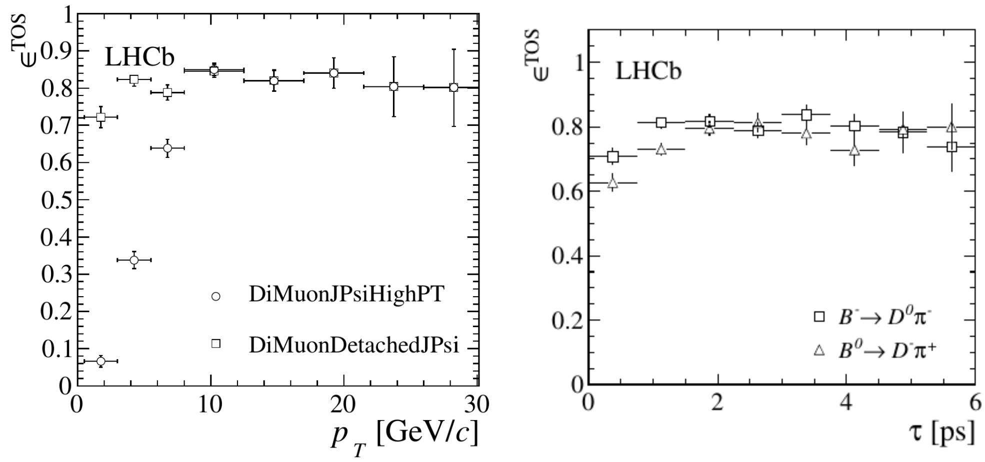


(a) (b)

Figure 3.19: Diagrams showing the TOS efficiency of the HLT1 muon lines (left) for $B^+ \to J/\psi\ (\to \mu^+\ \mu^-)\ K^+$ decays and the HLT1 charged track line (right) for various B and D meson decays [14]. Both plots use 2011 data and show the dependency of the efficiency on the p_{T} of the J/ψ, B and D mesons.

The HLT1 trigger reduces the readout rate to 43 kHz by adding information from the tracking system and performing a partial event reconstruction, which enables the reconstruction of primitive decay vertices and the calculation of the IP of a charged track. The HLT1 muon lines select either single muons with high IP and momentum (HLT1TrackMuon line) or muon pairs that make well defined vertices with an invariant dimuon mass of $m_{\mu^+\mu^-} > 1\,\text{GeV}/c^2$ (HLT1DiMuonLowMass line) or $m_{\mu^+\mu^-} > 2.7\,\text{GeV}/c^2$ (HLT1DiMuonHighMass line). Events containing charged particles with high IP, p_{T} and momentum, typical of B and D decays, are selected by the HLT1TrackAllL0 line. The 2011 data performance of the muon and single track HLT1 lines is shown in Fig. 3.19, using the decays described in section 3.6.1. The efficiency is seen to increase with the p_{T} of the meson. The single track lines are less efficient for D than B decays, as the charmed mesons have a lower mass than the beauty mesons, resulting in decay products with lower IP and p_{T}.

3.6.3 HLT2



(a) (b)

Figure 3.20: Diagrams showing the TOS efficiency of the HLT2 detached dimuon lines (left) for $B^+ \to J/\psi\ (\to \mu^+\ \mu^-)\ K^+$ decays and any of the HLT2Topo$n$Body topological trigger lines (right) for $B^- \to$ D0 $(\to K^+\ \pi^-)\ \pi^-\ (n = 2, 3)$ and $B^0 \to D^-\ (\to K^+\ \pi^-\ \pi^-)\ \pi^+\ (n = 2, 3, 4)$ decays [14]. Both plots use 2011 data and show the dependency of the efficiency on the p_{T} of the J/ψ and the lifetime τ of the B mesons.

At the HLT2 stage a full event reconstruction using information from all the subdetectors is performed to reduce the output rate to 3 kHz in 2011. When an event passes this trigger stage the raw digitised output from all the detector subsystems is stored in a compressed 'RAW' format for use in subsequent physics analyses. In principle an HLT2 line can use all the selection techniques deployed by full offline physics analyses, provided that the CPU time and output rate is within acceptable limits and detector calibration data are available. This allows for the construction of highly flexible and efficient trigger lines that can be tailor-made to select either specific 'exclusive' decays or general 'inclusive' groups of decays. An example of the latter are the topological 'HLT2Topo$n$Body' lines, which select generic decays of beauty hadrons into $n = 2, 3, 4$ bodies using multivariate selection techniques, as detailed in [14]. Fig. 3.20 shows the performance in 2011 of two lines which select $J/\psi \to \mu^+\ \mu^-$ decays and of the topological trigger lines. The line 'DiMuonJPsiHighPT' selects J/ψ decays with high p_{T} (as shown by the increase of efficiency with p_{T}), while 'DiMuonDetachedJpsi' selects J/ψ decays which are displaced from the PV, the efficiency of this line stays at $\sim 80\%$ for all p_{T}. The topological line efficiencies for purely hadronic 3 and 4 body B meson decays with lifetimes $\tau > 1\,\mathrm{ps}$ are $70 - 80\%$.

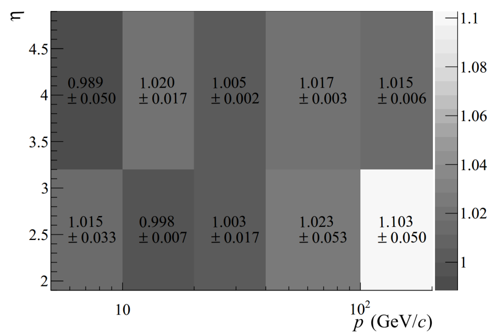
3.7 Offline analysis

The stored RAW LHCb data is analysed by reconstructing the data into low-level physical quantities, such as calorimeter cluster energies, tracking station hits and PID information, from which the momentum four-vectors corresponding to particles can be built and combined to form decay vertices. The event size of reconstructed data is roughly twice that of the RAW data. To save storage space only events which pass a pre-selection process called 'stripping' are stored in the fully reconstructed format. The stripping selection consists of a set of custom-built 'lines' designed to select specific candidate particles and decays of interest. The selected candidates are stored alongside the reconstructed data allowing for easy access by physics analysis.

3.7.1 Simulation

Simulated events containing decays of interest are used extensively in LHCb analyses, mainly to calculate and understand the efficiencies and systematic uncertainties related to the selection and reconstruction of decays of interest. The underlying *pp* collision events containing the signal decay are generated using Pythia 6.4 [48] configured with the parameters detailed in Ref. [49]. The decay and evolution of the resulting particles (including the signal decay) is modelled with the EvtGen [50] package. Final state QED radiative corrections are included using the Photos package [51]. The interaction of the particles with the LHCb detector is simulated with the Geant 4 [52, 53] package. All the simulation samples used in the analyses in Chapter 4, Chapter 5 and Chapter 6 are generated with detector performance parameters tuned to closely resemble the running conditions in 2011.

Corrections to the simulation



(a)

Figure 3.21: A 2D histogram showing the tracking efficiency ratio between 2011 data and the simulation, in bins of track momentum and pseudorapidity. Taken from [15]

The LHCb simulation provides a broadly accurate representation of the data. However, there are differences observed for three important quantities that are important for physics analysis: the impact parameter resolution; track reconstruction efficiencies and the PID DLL variable distributions. These contribute to systematic uncertainties in physics analyses and have to be corrected for *a-posteriori*.

The IP resolution is better in the simulation than the data, as shown in Fig. 3.8. This is likely caused by the distribution of material in the VELO and the scattering of particles by the VELO not being modelled precisely in the GEANT 4 stage of the simulation. The correct IP resolution is achieved by smearing the (x, y) position of the tracks with a Gaussian distribution, the width of which defines a 'smearing scale', calibrated such that a scale of 1.0 reproduces the data IP distribution and 0 corresponds to no smearing being applied.

The overall track reconstruction efficiency is replicated by the simulation to within 1%, where the efficiency is measured in data using the 'tag-and-probe' technique with $J/\psi \to \mu^+\mu^-$ decays [15]. There are slightly larger data-simulation deviations in some bins of track p and η, as shown in Fig. 3.21, caused by small inaccuracies in the material description of the tracking system, the relative alignment between its different components and the modelling of scattering interactions. When calculating the reconstruction efficiency of a decay each final state particle is assigned a weight which corrects the efficiency to that measured in data, using the values shown in Fig. 3.21.

The DLL variables are not well replicated in the simulation. This is primary because the photoelectron background is underestimated in the simulation. This in turn is caused by the multiplicity of low momentum tracks originating from secondary interactions in the detector and beampipe being lower than in data. To correct for this the likelihood for a particle to pass a given DLL cut is extracted from the data using the PID calibration samples detailed in Appendix B.

Chapter 4

Studies of the RICH performance

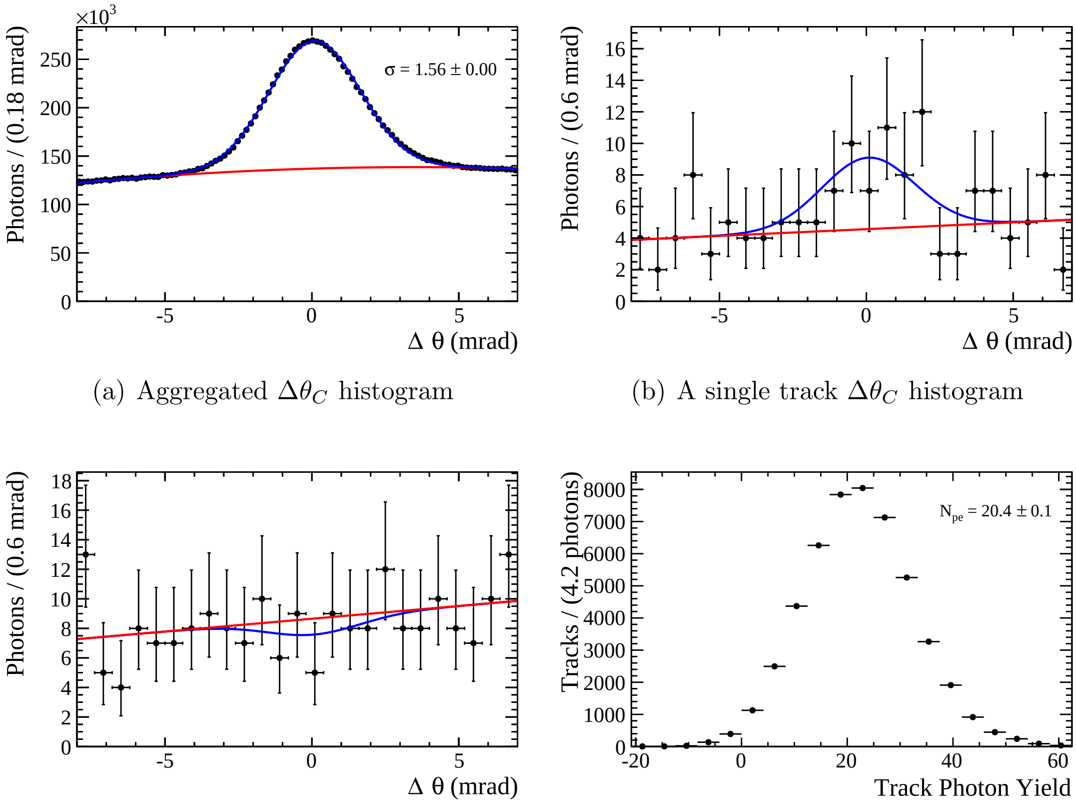
This chapter describes the measurement of the photoelectron yield of the RICH detector subsystem. The yield is a key input into the performance of the RICH. The yield is seen to be lower than expected from simulation. This result is published in European Physical Journal C [11]. A subsequent unpublished study is performed to analyse the evolution of the yield through 2011 and 2012. The author performed all of the studies detailed in this chapter, except the PID efficiency measurements shown in Fig. 4.13 and Fig. 4.15.

4.1 Introduction

The performance of the RICH PID algorithm, detailed in Sect. 3.4.1, is dependent on the precision of the reconstruction of the Cherenkov angle θ_C of a photon and the photoelectron yield, N_{pe}. N_{pe} is defined as the mean number of Cherenkov photons emitted by a charged track that are detected by the HPD's. Larger values of N_{pe} enable the RICH global pattern recognition software to discriminate more strongly between the pattern of HPD hits which arise from the true PID hypothesis of a track and other (false) hypotheses. The analytic expression for the expected N_{pe} of a saturated track ($\beta \approx 1$) with unit electric charge is [54]:

$$N_{pe} = \frac{\alpha}{\hbar c} L \epsilon_A \eta \int_{1.5\,\mathrm{eV}}^{6.2\,\mathrm{eV}} \mathcal{QRT} sin^2\theta_C dE_\gamma, \tag{4.1}$$

where $\alpha = 1/137$ is the fine structure constant, L is the path length in the radiator, ϵ_A is the fraction of the Cherenkov ring area that projects onto active HPD pixels, η is the efficiency with which a single photoelectron is detected by an HPD after conversion by the photocathode, $\mathcal{Q}$ is the HPD quantum efficiency, $\mathcal{R}$ is the mirror reflectivity, $\mathcal{T}$ is the combined transparency of the media and detector elements that the photon travels through before conversion and E_γ is the photon energy. During the design phase of the RICH system the expected values for N_{pe} for the three radiators were calculated by Eq. (4.1) to be ~ 6.5, ~ 30 and ~ 22 for unit-charge tracks with $\beta \approx 1$.



(a) Aggregated $\Delta\theta_C$ histogram

(b) A single track $\Delta\theta_C$ histogram

(c) A single track $\Delta\theta_C$ histogram

(d) Track N_{pe} distribution

Figure 4.1: Histograms illustrating the N_{pe} calculation method described in the text, produced using the full 2011 kaon and pion PID calibration sample dataset, with the 'normal' event selection applied.

The following sections describe the method used to analyse the photoelectron yield in LHCb. N_{pe} is measured for data collected in 2011 and compared with simulation. A subsequent analysis of the variation of N_{pe} through 2011 - 2012 and its impact on the RICH PID performance is detailed in Sect. 4.5.

4.2 Method

It is not possible to determine whether an HPD hit was produced by a Cherenkov photon from a specific track, or from a background source. Therefore a statistical method is used to extract N_{pe} by using the variable $\Delta\theta_C$, which defines the difference between the measured and expected Cherenkov angle for a track and HPD pixel hit combination:

$$\Delta\theta_C = \theta_C - \theta_{exp}, \tag{4.2}$$

where θ_{exp} is the expected Cherenkov angle for the particle, given its momentum and mass, as defined by Eq. (3.1) and θ_C is the measured angle between a track and a photon candidate. θ_C is reconstructed using optical ray tracing from an HPD hit that is projected

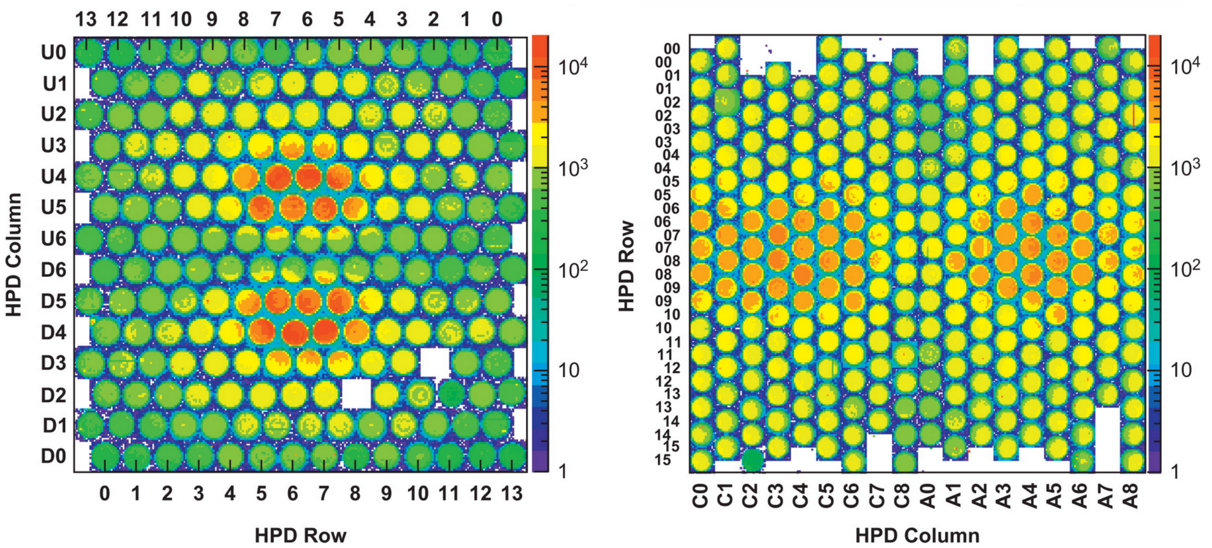
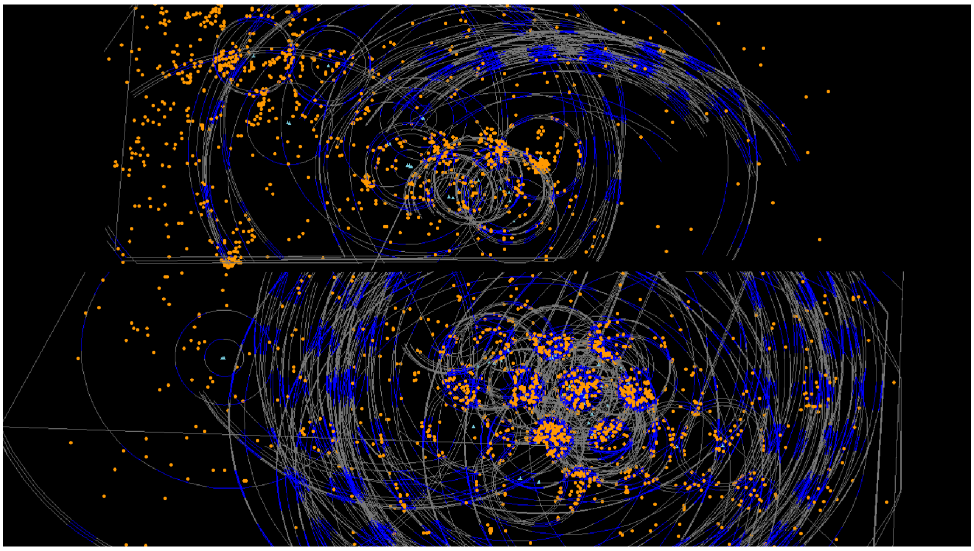
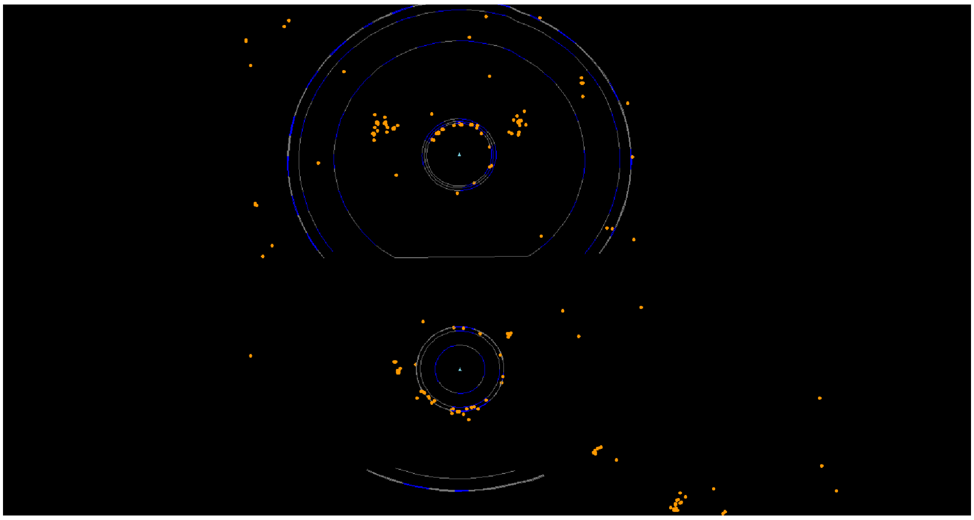


Figure 4.2: HPD hit distribution of the RICH1 (left) and RICH2 (right) HPD planes for a typical 2011 data collection run, taken from [16].

through the RICH optical system onto the halfway point of the track's trajectory through the radiator. The $\Delta\theta_C$ values for all track-photon combinations are collated into a histogram, as shown in Fig. 4.1(a). Cherenkov photons that are correctly associated to their parent track produce a Gaussian $\Delta\theta_C$ distribution centred around zero (within the precision with which the refractive index of the radiator is measured), whereas the background hits and incorrect track-photon associations produce a smooth, non-peaking background. The width of the Gaussian signal defines the resolution with which θ_C is measured. The N_{pe} is calculated by fitting the aggregated $\Delta\theta_C$ distribution with a combined Gaussian signal and a second order polynomial background Probability Density Function (PDF), as shown in Fig. 4.1(a). The non-linear component of the background arises from the variation in HPD hit density across the whole HPD plane, as shown in Fig. 4.2. The $\Delta\theta_C$ distribution of each individual track is then fitted with a Gaussian PDF with a mean set to zero and a width fixed to that obtained from the aggregated $\Delta\theta_C$ fit. The background is fitted with a linear PDF. The N_{pe} of the individual track is calculated as the area under the signal shape. Examples of individual track fits are shown in Fig. 4.1(b) and Fig. 4.1(c). The signal fit is allowed to have a negative yield, as this can arise when background fluctuations in the upper and lower $\Delta\theta_C$ sidebands are greater than those in the signal region, as is the case in Fig. 4.1(c). This is compensated for by tracks where the fluctuations enhance the signal yield. The overall N_{pe} is calculated as the mean of the individual track N_{pe} distribution, as shown in Fig. 4.1. The error is assigned as the standard error on the mean.



(a) A $pp \to D^{*+}(\to D^0(\to K^-\pi^+)\pi^+)X$ event



(b) A $pp \to pp\mu^+\mu^-$ event

Figure 4.3: Event displays showing the RICH1 HPD planes, recorded during 2011 data collection. The dots show the HPD hits, the triangles show the position of charged tracks which are propagated through the mirror system onto the HPD plane. The circular lines show the expected Cherenkov rings produced by the aerogel and C_4F_{10} radiators for the pion, kaon and proton PID hypotheses, the blue (dark grey in B&W) sections of the rings show the ring overlap with active HPD regions.

Table 4.1: Momentum cuts applied to the different species of tracks used in the N_{pe} study.

Particle type	Minimum momentum (GeV/c)		
	Aerogel	C_4F_{10}	CF_4
π, μ	5	25	40.4
K	9.8	37	74.8

4.3 Datasets and selection

The N_{pe} is measured in data using tracks belonging to two different categories of RICH events: 'normal' and 'ideal' events.

Normal events are representative for typical data-collection conditions at LHCb in 2011. The tracks used for this category are kaons and pions from the PID calibration sample, detailed in Appendix B.

Ideal events have optimal conditions for RICH operation, with very low HPD backgrounds and tracks with unobstructed Cherenkov rings. Muons from $pp \to pp\mu^+\mu^-$ events are chosen, which arise from a diffractive scattering process [55]. These events are selected by requiring that the event does not contain a $p-p$ collision vertex (referred to as the 'primary vertex' from hereon) and that the two muons make a vertex which is located within a ± 1.5 mm radius from the centre of the detector in the $x - y$ plane, both muons have $p_{\rm T} > 400\,{\rm MeV}/c$ and satisfy 'IsMuon = true'. The muons are required to have $\epsilon_A > 0.5$ (from Eq. (4.1)), reducing the number of Cherenkov photons which are lost due to the Cherenkov cone being obstructed by the beampipe, or projecting to regions outside the HPD plane or to the gaps between the HPDs.

Examples of the RICH1 HPD hit patterns produced by the two event types are shown in Fig. 4.3, for the $pp \to pp\mu^+\mu^-$ event the Cherenkov rings produced by both muons can be seen very clearly. Momentum cuts as detailed in Table 4.1 are applied to all tracks from both categories, such that $\theta_{exp} \sim \theta_{sat}$ defined in Eq. (3.2), in order to minimise the uncertainty on θ_{exp} resulting from the measurement of the track momentum.

The validity of the N_{pe} calculation method is checked using simulated Monte Carlo (MC) samples of the kaon and pion PID calibration events described in Appendix B. These events contain MC truth information on the origin of HPD hits which can be used to identify correct track-photon combinations to access the 'true' N_{pe}. The simulation track sample has an $\epsilon_A > 0.5$ cut applied and the charged-track multiplicity of the events is required to be less than 50. This enables a like-for-like comparison of N_{pe} to be made with the data $pp \to pp\mu^+\mu^-$ events. The N_{pe} of the simulated events is calculated for three different configurations: the 'true N_{pe}', the 'normal N_{pe}' and the 'digital readout N_{pe}'. The true N_{pe} is calculated by counting the number of photons which have the correct track-pixel association for each track and then taking the mean of the resulting individual track N_{pe} distribution. The normal N_{pe} is obtained in the same way as for the data. The digital readout N_{pe} is a special simulation configuration where the RICH HPD pixels have digital readout. In this configuration pixels hit by more than one photoelectron register the total number of hits, as opposed to registering just one hit for the binary readout configuration, present in the physical detector. The N_{pe} values are calculated in the same way as for data.

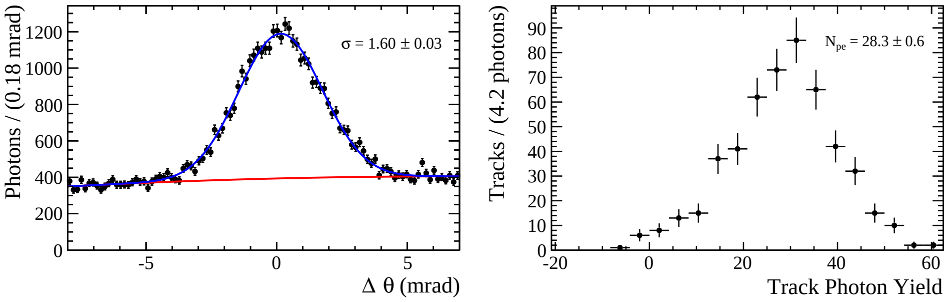
The photon yield calculated using the method detailed in Sect. 4.2 can be less for the binary readout setting than the digital readout setting. This occurs when HPD pixels are hit by a signal photoelectron and also one or more background photoelectrons. In the binary readout setting only one photon would be registered, in the delta-theta fit these photons would, on average, be assigned as background hits, resulting in a reduction in the calculated photon yield. This should not be interpreted as an inefficiency, as the HPD pixel still registers a hit by a signal photoelectron. This effect becomes more pronounced in events with greater HPD occupancies, where the increase in photon background makes it more likely for signal photon hits to coincide with background hits. This artificial yield suppression would not occur for the digital readout setting, where all photons are counted when a pixel is hit by multiple photons, or for simulated events when the yield is calculated using the truth information of the HPD pixels.

4.4 Results

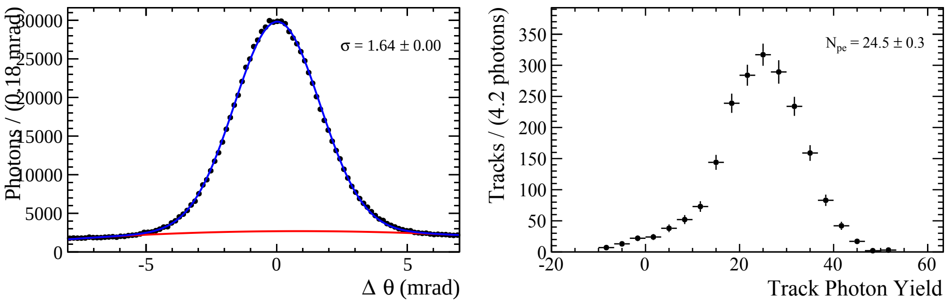
The photoelectron yield is compared for the different simulated and data events detailed in Sect. 4.3 for all three radiators. The results of the ideal event category are obtained using data collected in May and June 2011, when the RICH performance was found to be most stable. The time dependence of the photon yield and the RICH performance is discussed further in Sect. 4.5. The dependence of N_{pe} on the charged track multiplicity is investigated as this was found to affect the RICH PID performance. The track N_{pe} and $\Delta\theta_C$ distributions are also compared.

4.4.1 C_4F_{10}

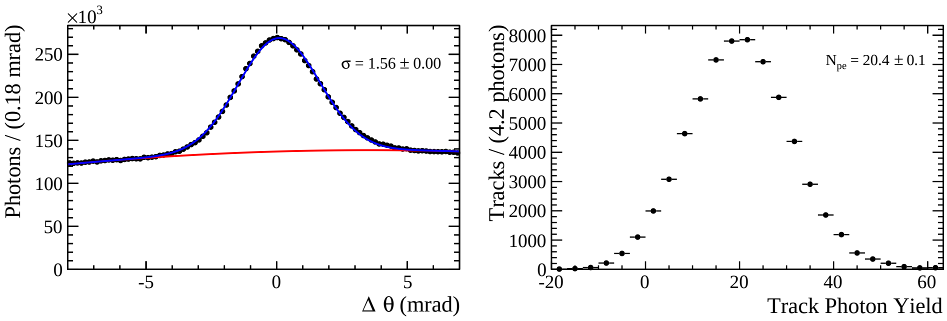
A comparison of the C_4F_{10} $\Delta\theta_C$ and track N_{pe} distributions for data and simulation is shown in Fig. 4.4, all three $\Delta\theta_C$ distributions are fitted very well by the signal and background PDFs. The widths of both data $\Delta\theta_C$ signal shapes are consistent with that of the simulated events. All three shapes are consistent with $\sigma = 1.6$ mrad, as predicted from the simulation. This indicates that the alignment of the RICH1 optical system, the correction of the fringe magnetic field distortion of the HPD images and the precision of the tracking system, are all accurate to within the design specifications of the LHCb experiment. The $pp \to pp\mu^+\mu^-$ data N_{pe} of 24.5 ± 0.3 is 13% less than the N_{pe} value of 28.3 ± 0.6 for the equivalent simulated events. The $pp \to pp\mu^+\mu^-$ N_{pe} cannot be suppressed



(a) C_4F_{10}, simulation, 'ideal' event selection, charged track multiplicity < 50



(b) C_4F_{10}, May - June 2011 $pp \to pp\mu^+\mu^-$ data, 'ideal' event selection



(c) C_4F_{10}, full 2011 D^* tagged $D^0 \to K^-\pi^+$ dataset, 'normal' event selection

Figure 4.4: Aggregated C_4F_{10} $\Delta\theta_C$ distributions (left) and the corresponding track N_{pe} distributions (right).

by either high track multiplicities (only two tracks in each $pp \to pp\mu^+\mu^-$ event), or the high L0 trigger rates which suppress the yield after June 2011 (see Sect. 4.5). This implies that the lower data yield is may be due to the photoelectron efficiency terms in Eq. (4.1) being slightly lower than measured during the design phase of the LHCb experiment. The N_{pe} corresponding to the full 2011 D^* tagged $D^0 \to K^-\pi^+$ dataset is 20.4 ± 0.1 photons, which is less than the $pp \to pp\mu^+\mu^-$ yield due to no $\epsilon_A > 0.5$ cut being applied, the higher charged track multiplicities and high L0 rate N_{pe} suppression.

The dependence of the RICH1 N_{pe} on charged track multiplicity is shown in Fig. 4.5. A clear dependence of the RICH1 N_{pe} on charged track multiplicity is seen in both data and the 'normal' binary readout MC, the gradients of the fitted linear functions for these plots are $(-7.0 \pm 0.4) \times 10^{-3}$ N_{pe}/track and $(-10.0 \pm 1.3) \times 10^{-3}$ N_{pe}/track respectively.

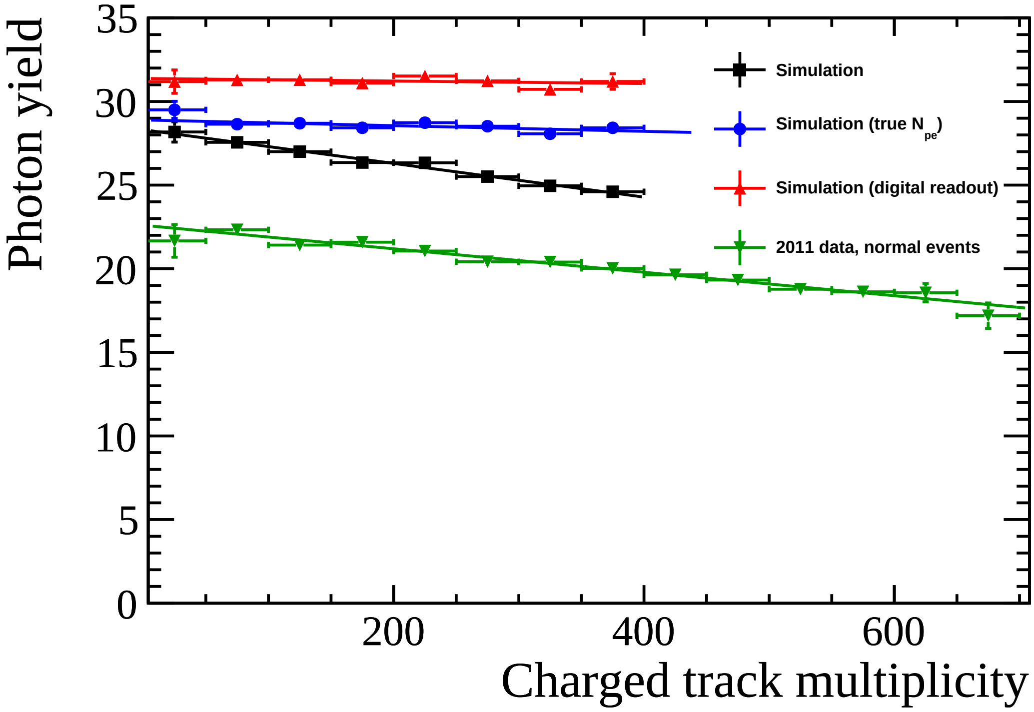
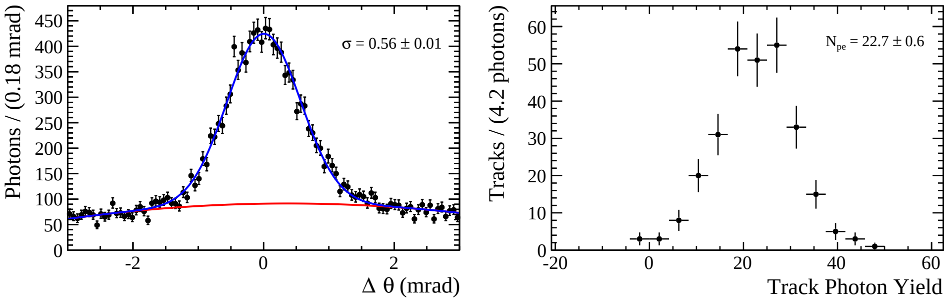


Figure 4.5: Plots showing the dependence of the N_{pe} on the event charged track multiplicity for the C_4F_{10} radiator, illustrated by linear fit lines to the data points. Simulation and 2011 data are compared.

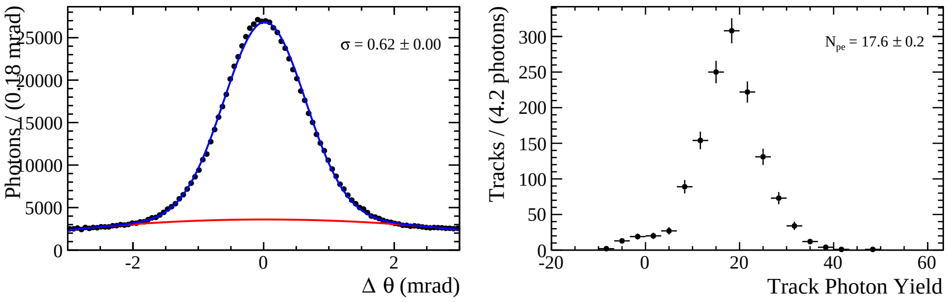
The 'true' and digital readout MC yields are consistent with having no dependency on track multiplicity, their linear fit gradients are $(-1.7 \pm 0.8) \times 10^{-3}$pe/track and $(-0.7 \pm 1.2) \times 10^{-3}$pe/track respectively. This indicates that at higher track multiplicities the C_4F_{10} N_{pe} gets increasingly suppressed as more photon information is lost due to the binary HPD readout via increased HPD occupancies. For low track multiplicities of < 50 the true and calculated MC N_{pe} values are 29.5 ± 0.5 and 28.3 ± 0.6 respectively, which are consistent with each other to within ~ 2 standard deviations. It can therefore be concluded that at low track multiplicities the N_{pe} suppression from excess photoelectron background is minimal. The digital readout N_{pe} is ~ 2 photoelectrons greater than the true MC yield for all bins of track multiplicity. This can be the result of binary readout suppression from the signal Cherenkov photons themselves.

4.4.2 CF_4

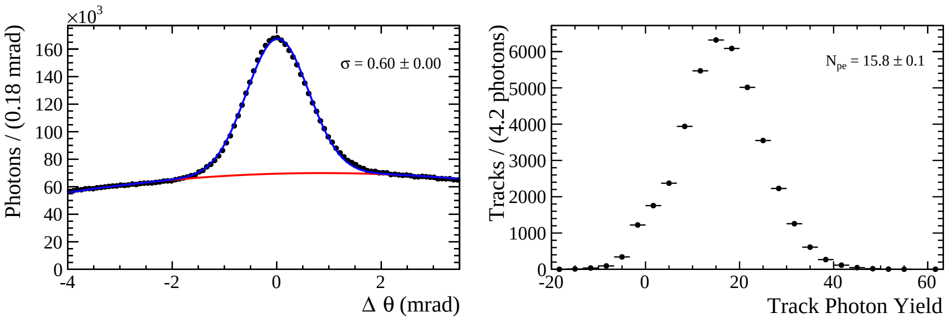
The $\Delta\theta_C$ and track N_{pe} distributions obtained from the RICH2 gas, CF_4, for data and simulation are compared in Fig. 4.6. All $\Delta\theta_C$ distributions are fitted well with the Gaussian signal and polynomial background PDF shapes. The signal widths are consistent with 0.6 mrad, as expected from the simulation. The simulated events have an N_{pe} value of 22.7 ± 0.6, larger than $N_{pe} = 17.6 \pm 0.2$ for the ideal data $pp \to pp\mu^+\mu^-$ events. The 29% relative difference between the two yields is larger than the equivalent 13% difference



(a) CF_4, simulation, 'ideal' event selection, charged track multiplicity < 50



(b) CF_4, May - June 2011 $pp \to pp\mu^+\mu^-$ data, 'ideal' event selection



(c) CF_4, full 2011 D^* tagged $D^0 \to K^-\pi^+$ dataset, 'normal' event selection

Figure 4.6: Aggregated CF_4 $\Delta\theta_C$ distributions (left) and the corresponding track N_{pe} distributions (right).

observed in C_4F_{10}. The cause of the discrepancy may be due to overestimation of the HPD photon detection efficiencies used in the simulation. The N_{pe} obtained from $D^0 \to K^-\pi^+$ data is less than the $pp \to pp\mu^+\mu^-$ yield due to the absence of a track $\epsilon_A > 0.5$ cut and the inclusion of the post-June 2011 dataset which has a reduced N_{pe} due to the increased L0 rate.

The track multiplicity dependence of N_{pe} for the CF_4 is shown in Fig. 4.7. No dependency is seen in the simulation, and a very slight reduction in N_{pe} is seen at high multiplicities in $D^0 \to K^-\pi^+$ data. In CF_4 the N_{pe} suppression caused by the binary readout being saturated by background photoelectrons, as seen in C_4F_{10}, is minimal as the occupancies of the pixels in the busiest HPD region are ~ 5 times less in CF_4 than for C_4F_{10} (see Fig. 4.2).

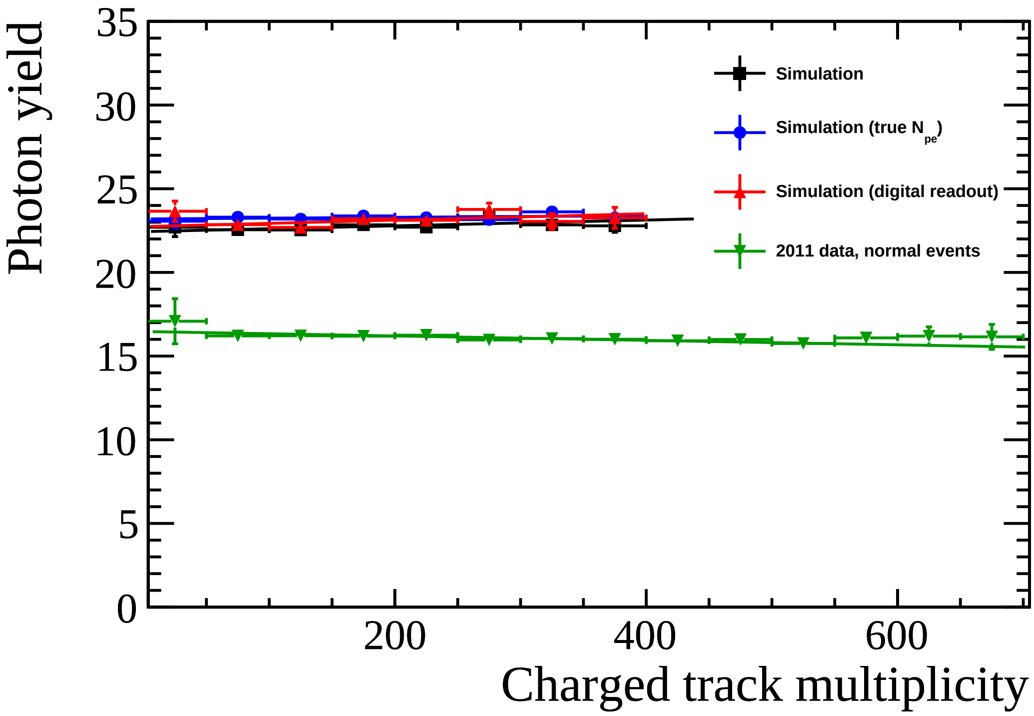
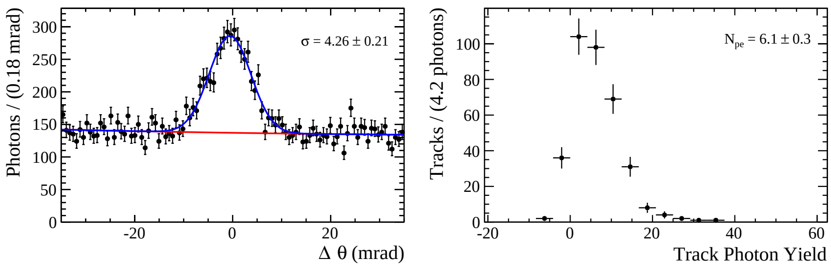


Figure 4.7: Plots showing the dependence of the N_{pe} on the event charged-track multiplicity for the CF_4 radiator. The lines are linear fits to the data points. Simulation and 2011 data are compared.

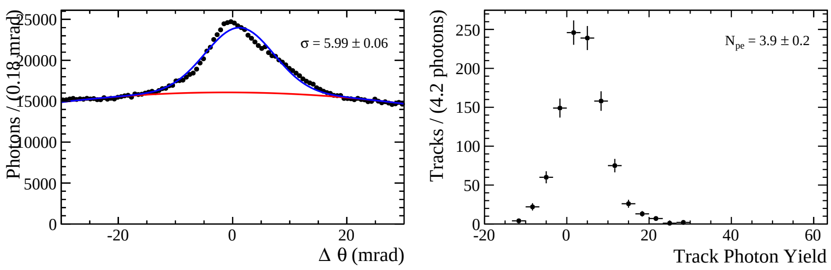
The CF_4 HPD Cherenkov ring HPD image circumference is ~ 4 times greater than for C_4F_{10}, due to the longer focal length of the RICH2 optical system. So the CF_4 has a lower single-track Cherenkov photon density on the HPD plane. As a result the probability of two Cherenkov photons striking the same pixel is negligible in CF_4 compared to C_4F_{10}.

4.4.3 Aerogel

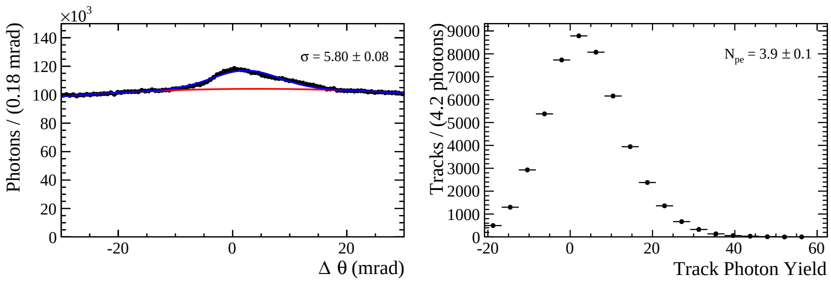
The $\Delta\theta_C$ and track N_{pe} distributions extracted from the aerogel simulation and data are shown in Fig. 4.8. The aerogel signal shape in data (Fig. 4.8(b,c)) is composed of a superposition of multiple Gaussian shapes with different values of θ_C. As such it is not well fitted by a single Gaussian PDF. This also causes the width of the fitted Gaussian shape to be much larger in data ($\sigma \sim 5.8$) than for the simulation ($\sigma \sim 4.3$ mrad). The difference in the values of θ_C is a consequence of the variation of the refractive index across the aerogel plane, which is at least partially caused by the absorption of C_4F_{10} gas by the porous aerogel surface. Calibration software partially corrects for the variation, by producing distributions of θ_C for tracks passing through different 40 mm $\times$40 mm 'subtile' regions of the aerogel plane and extracting θ_{exp} for each individual region. The aerogel calibration software cannot correct for variations in n within the subtiles themselves and as such has a limited effectiveness in producing a uniform $\Delta\theta_C$ signal shape. The signal-to-background ratio in the aerogel $\Delta\theta_C$ plot is very low compared to the gas radiators,



(a) Aerogel, simulation, 'ideal' event selection, charged track multiplicity < 50



(b) Aerogel, May - June 2011 $pp \to pp\mu^+\mu^-$ data, run No. < 95k, 'ideal' event selection



(c) Aerogel, full 2011 D^* tagged $D^0 \to K^-\pi^+$ dataset, 'normal' event selection

Figure 4.8: Aggregated aerogel $\Delta\theta_C$ distributions (left) and the corresponding track N_{pe} distributions (right). The aerogel refractive index calibration has been applied to the data distributions (b,c).

especially for the $D^0 \to K^-\pi^+$ events in data shown in Fig. 4.8(c). There are three primary causes of this. First, the aerogel emits fewer Cherenkov photons than the gas radiators, as most photons are Rayleigh scattered by the aerogel itself. These scattered photons contribute to the photoelectron background in RICH1. Second, the Cherenkov angle of aerogel is much larger than that of the gas radiators (see Fig. 3.9(a)). As a result the aerogel Cherenkov photons are distributed in a Cherenkov ring with a diameter far wider than that of C_4F_{10}, as seen in Fig. 4.3, over which background HPD hits are collected into the $\Delta\theta_C$ distribution. Third, the large width of the aerogel signal shape also increases the area of the HPD plane over which the signal photons are distributed, further reducing the Cherenkov photon HPD hit density of the aerogel.

The N_{pe} values of 3.9 ± 0.1 and 3.9 ± 0.2 for the respective $D^0 \to K^-\pi^+$ and

$pp \to pp\mu^+\mu^-$ data tracks are in good agreement with each other, these are significantly less than the simulated event yield of 6.1 ± 0.3.

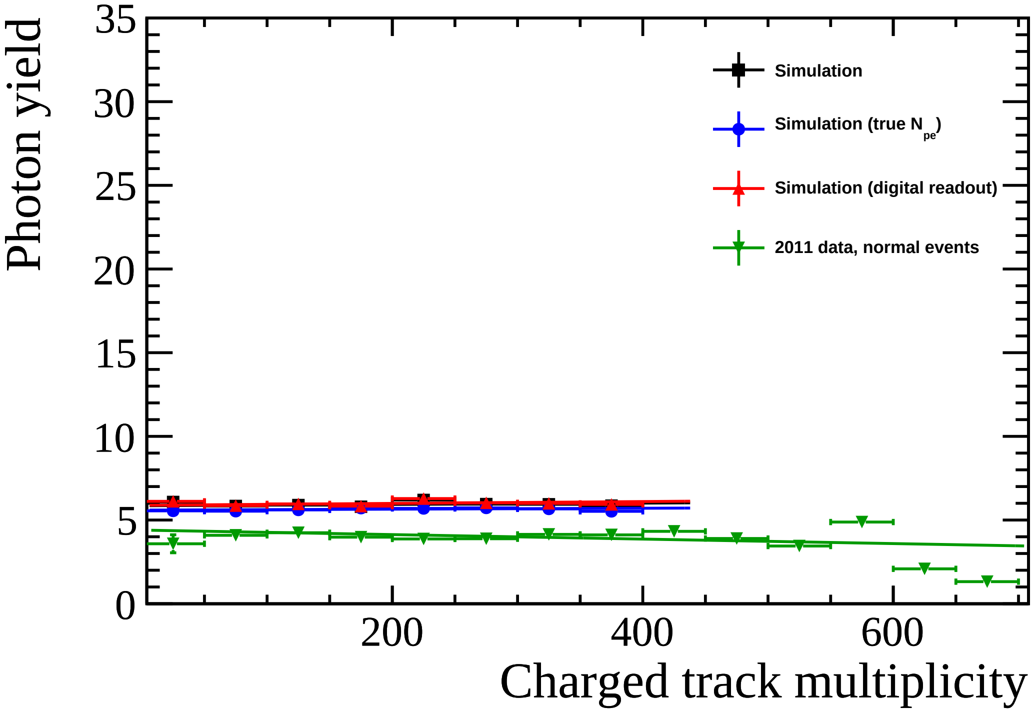


Figure 4.9: Plots showing the dependence of N_{pe} on the event charged-track multiplicity for the aerogel radiator. The lines are linear fits to the data points. Simulation and 2011 data are compared.

The dependence of the aerogel N_{pe} on the charged-track multiplicity of the event is shown in Fig. 4.9. The yields of the three simulation configurations are in good agreement, indicating that the N_{pe} value returned by the photon yield calculation method is accurate, at least for the case where the signal shape is a pure Gaussian. No dependence on track multiplicity is seen for either the data or the simulation, as is to be expected because the aerogel rings fall on the outer RICH1 HPD regions which have a photoelectron hit density of ~ 100 times less than the inner HPD regions. As such there are few instances where photons are lost due to pixels being hit by multiple photoelectrons. This is confirmed by the digital readout configuration returning the same N_{pe} values as the binary readout configuration. The N_{pe} values from $D^0 \to K^-\pi^+$ data deviate from the trend at high charged-track multiplicities. This is due to the poor $\Delta\theta_C$ signal-to-background ratio at high track multiplicities, as shown in the $\Delta\theta_C$ and track N_{pe} distributions for the 650-700 multiplicity bin in Fig. 4.10. For this set of tracks the signal is virtually indistinguishable from the background and no valid N_{pe} value can be extracted.

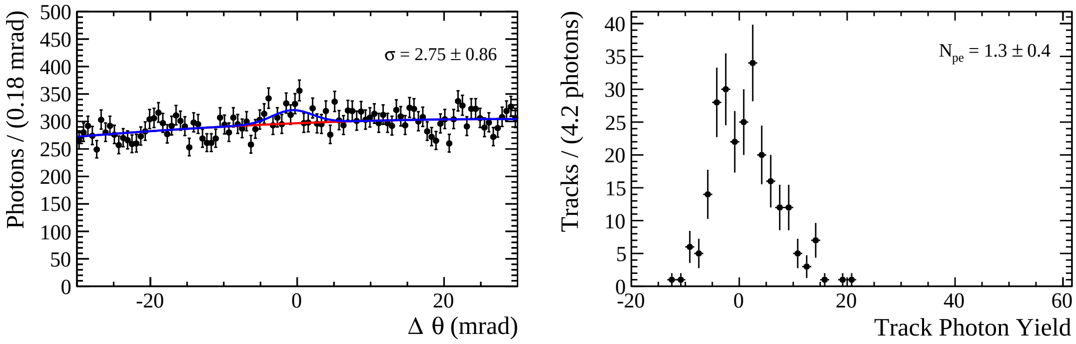


Figure 4.10: The aggregated aerogel $\Delta\theta_C$ distribution (left) and the corresponding track N_{pe} distribution (right) for aerogel events with a charged track multiplicity of 650 to 700, for the 2011 D^* tagged $D^0 \to K^-\pi^+$ dataset with the 'normal' event selection applied. The aerogel refractive index calibration is applied.

Table 4.2: Comparison of N_{pe}'s for 2011 data and simulation. Ideal event cuts: $\epsilon_A > 0.5$ and May-June data.

radiator	N_{pe} from 2011 Data			N_{pe} from simulation	
	Tagged $D^0 \to K^-\pi^+$		$pp \to pp\mu^+\mu^-$	Calculated N_{pe}	True N_{pe}
	No extra cuts	'ideal' event cuts			
Aerogel	3.9 ± 0.1	4.8 ± 0.3	3.9 ± 0.2	6.1 ± 0.3	5.5 ±0.1
C_4F_{10}	20.4 ± 0.1	24.2 ± 0.3	24.5 ± 0.3	28.3 ± 0.6	29.5 ±0.5
CF_4	15.8 ± 0.1	18.4 ± 0.3	17.6 ± 0.2	22.7 ± 0.6	23.3 ±0.5

4.4.4 Summary

The N_{pe} yields for all three radiators are compared and summarised in Table 4.2. For all radiators the simulation values of N_{pe} calculated using the method described in Sect. 4.2 agree with those calculated using the MC truth information to within ~ 2 standard deviations, indicating that the method reproduces the correct N_{pe}, given a low event track multiplicity and a $\Delta\theta_C$ distribution with well fitted signal and background components. The yields in $pp \to pp\mu^+\mu^-$ data are lower than in simulation for all radiators, suggesting that the photon detection efficiency terms shown in Eq. (4.1) are overestimated in the simulation, as these 'ideal' RICH events are selected to minimise other effects that can suppress N_{pe}, such as high track multiplicity or high L0 rates. The $D^0 \to K^-\pi^+$ yields are representative for the bulk data-taking conditions in 2011. These are lower than the $pp \to pp\mu^+\mu^-$ yield, as no special cuts, as used for the 'ideal' events, are applied. This is verified by applying the 'ideal' event selection criteria to the $D^0 \to K^-\pi^+$, which results in N_{pe} values that are consistent with $pp \to pp\mu^+\mu^-$ events for the gas radiators. The aerogel data N_{pe} yields have an additional uncertainty arising from the non-Gaussian structure of the $\Delta\theta_C$ signal shape. The effect on the yield produced by this is difficult to quantify as it is not possible to determine the precise $\Delta\theta_C$ signal shape.

The consequence of a lower than expected N_{pe} in data is a slight reduction in RICH PID performance. In the simulation a $\sim 95\%$ Kaon ID efficiency is associated with a $\pi \to K$ misID rate of $\sim 5\%$ for tracks with $2 < p < 100$ GeV/c. In 2011 data the equivalent $\pi \to K$ misID rate is $\sim 10\%$ for a 95 % kaon ID efficiency. This performance is still sufficient to achieve the physics goals of LHCb.

4.5 Time dependence of the photoelectron yield

The evolution of the photoelectron yield through 2011 and 2012 is measured by calculating the N_{pe} of $pp \to pp\mu^+\mu^-$ events in bins of run number, which is a chronologically designated label assigned to a period of data-collection under stable detector running conditions. The N_{pe} can be affected by variations in the mean number of pp interactions per bunch crossing, as the events recorded in runs with a higher number of interactions have larger charged track multiplicities which suppress N_{pe} via the binary readout effect discussed in the previous section. The increase in collision energy from $\sqrt{s} = 7\,\text{TeV}$ in 2011 to 8 TeV in 2012 also resulted in an increase in track multiplicity. The $pp \to pp\mu^+\mu^-$ events are not sensitive to these effects as they contain only two charged tracks. So the N_{pe} of these events is therefore exclusively sensitive to the RICH photon detection efficiency parameters in Eq. (4.1), enabling a like-for-like comparison between N_{pe} values in 2011 and 2012. The aerogel N_{pe} is not considered here owing to the uncertainties in the aerogel $\Delta\theta_C$ signal shape, as detailed in Sect. 4.4.3.

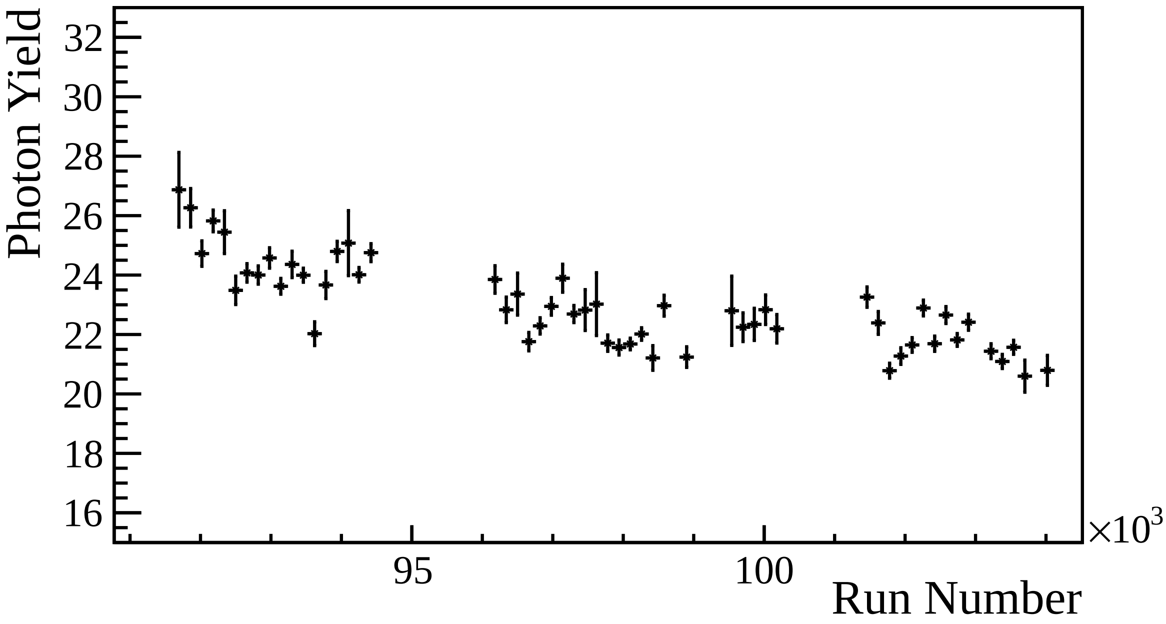
4.5.1 2011

The variation of N_{pe} in 2011 for the C_4F_{10} and CF_4 radiators is shown in Fig. 4.11. In C_4F_{10} a gradual fall in N_{pe} is seen, from ~ 24.5 photoelectrons for run numbers $< 95k$[1] to ~ 22 photoelectrons in run numbers $> 101k$[2], a decrease of $\sim 10\%$. Trends of increasing and decreasing N_{pe} are also seen across smaller timescales, for example, N_{pe} is seen to increase incrementally for each of the four run groups between runs 96740 and 97380, before decreasing again in the subsequent run groups. The CF_4 radiator exhibits the same behaviour, with a $\sim 10\%$ reduction in N_{pe} from ~ 17.5 to ~ 15.5 over the course of the 2011 data-collection period.

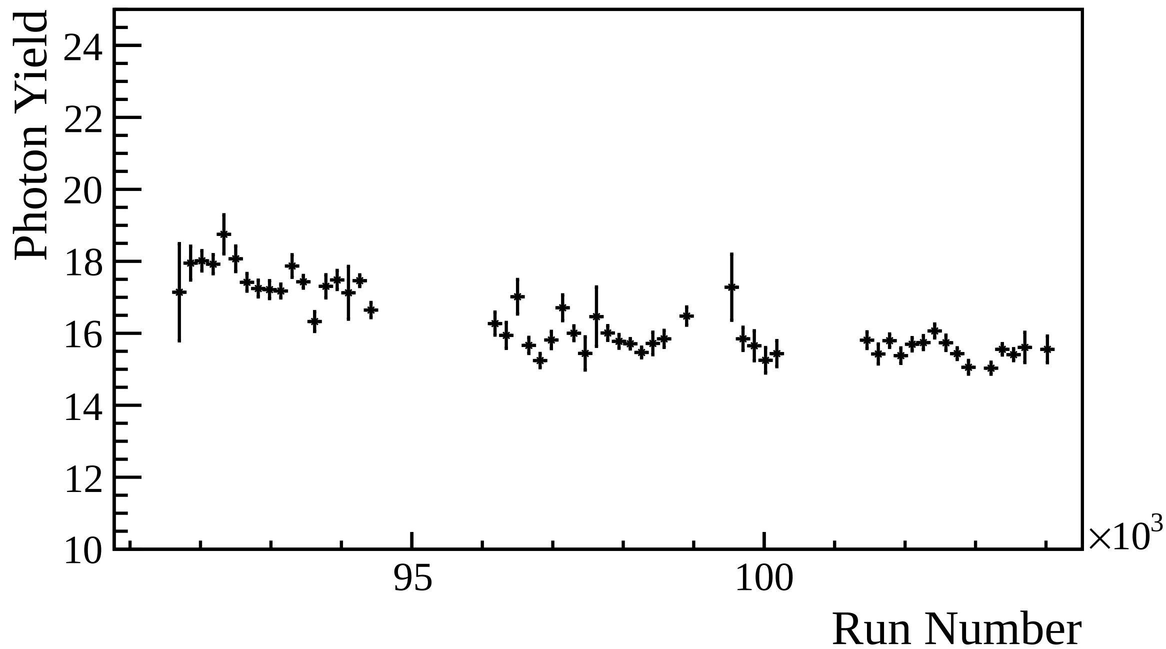
The reduction in N_{pe} across 2011 is reflected by a slight degradation in the PID performance, shown in Fig. 4.13. A 95 % kaon ID efficiency is associated with a mean $\pi \to K$ misID rate of $\sim 6\%$ for data collected in runs 89333 to 90207, which increases to $\sim 10\%$ for runs 103954-104414.

The underlying cause of the degradation of the photoelectron yield and PID performance is unclear. It was found that both could be partially recovered by adjusting a reference voltage in the HPD pixel readout chips. The re-adjustment was performed in June 2012 (run number 119.5k in Fig. 4.14), after which the N_{pe} in RICH1 was restored to the level observed in early 2011. In RICH2 the recovery in N_{pe} was less substantial than that in RICH1. The designers of the readout chips are currently investigating the causes of the loss in photon detection efficiency.

[1] Corresponding to data collected in May and June 2011
[2] Data collected in September and October 2011



(a) C_4F_{10}



(b) CF_4

Figure 4.11: The time variation of the N_{pe} yield of $pp \rightarrow pp\mu^+\mu^-$ events in 2011 for the C_4F_{10} and CF_4 radiators. Run number bins with less than 10 track N_{pe} values are not shown.

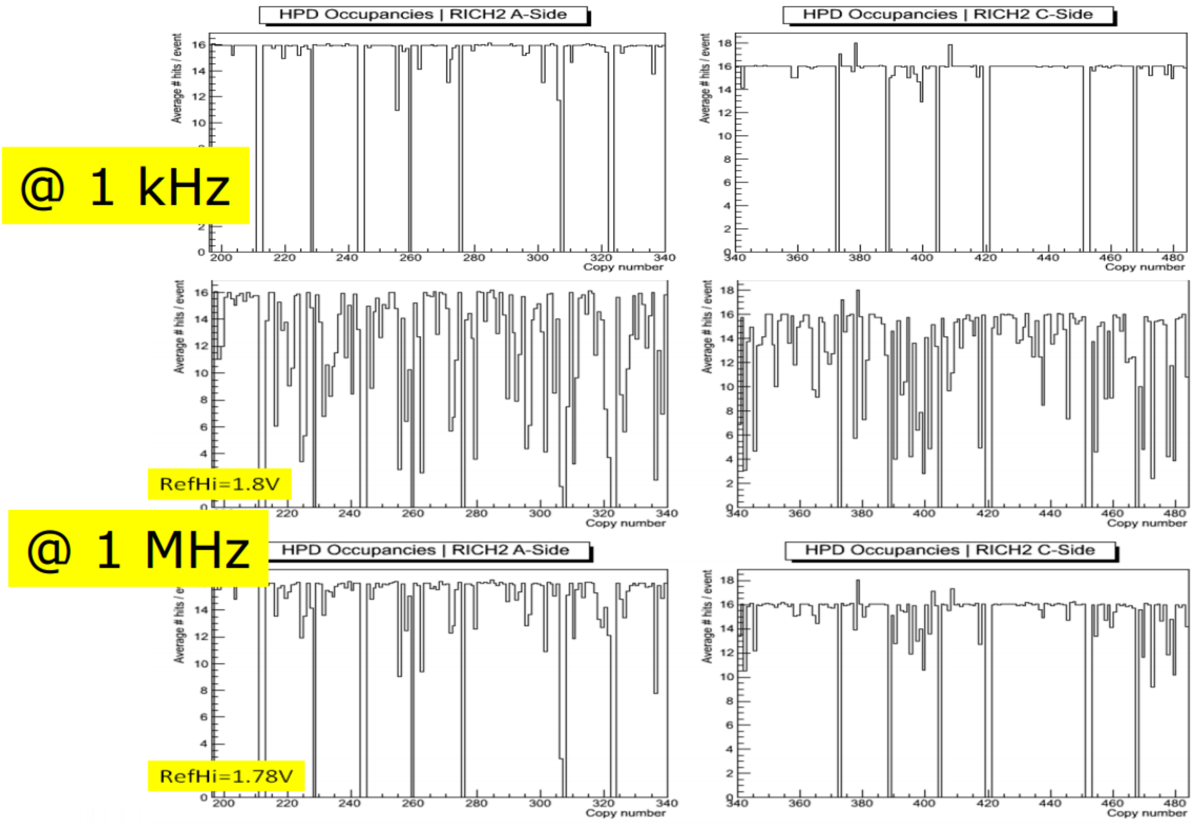


Figure 4.12: The average number of photons registered by each HPD (denoted by the 'copy number' ID on the x-axis) in RICH2 in events with no $p-p$ collisions, a pulsed laser in the RICH is fired which illuminates the HPD plane with $\approx$ 16 photons per HPD. For the top histograms the laser is fired and the are HPD's read out at 1 kHz, for the middle and bottom histograms the rate is increased to 1 MHz. For the bottom histogram a reference voltage of the HPD pixel readout chips is changed from the default value of 1.8V to 1.78V. Taken from https://indico.cern.ch/event/226173/session/1/contribution/7/material/slides/1.pdf .

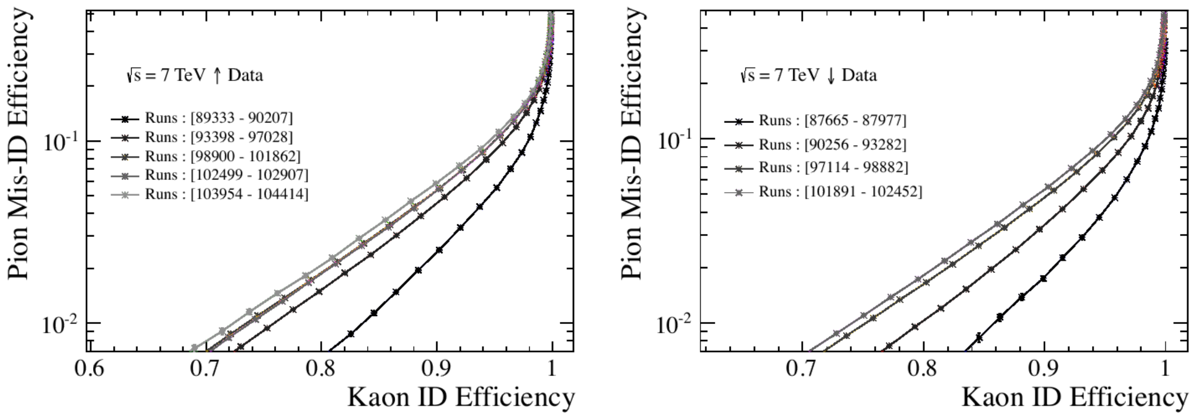


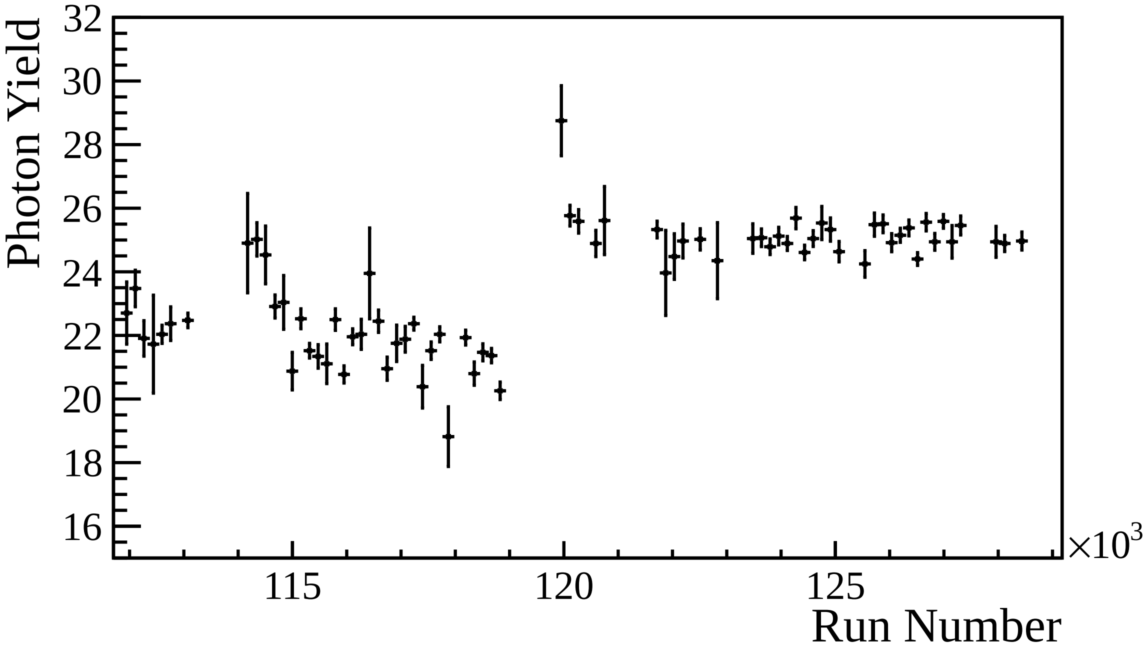
Figure 4.13: Plots showing the kaon ID vs pion to kaon misID rates for 2011 data collected in the 'magnet up' (left) and 'magnet down' (right) polarity configurations. This analysis used the pion and kaon calibration samples described in Appendix B. The performance is shown for the different groups of run number denoted in the legend. The performance is seen to be unaffected by the magnet polarity.

4.5.2 2012

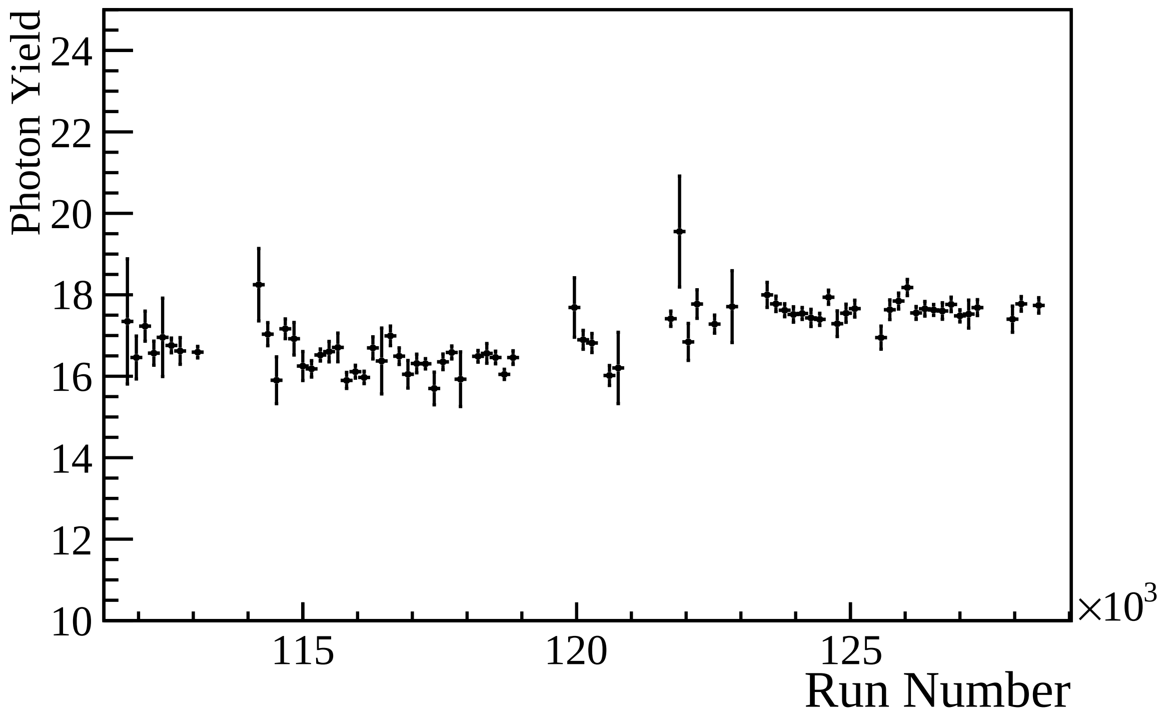
The N_{pe} variation in 2012 is shown in Fig. 4.14. The readout pixel adjustment described in the previous section was performed in a technical stop[3] which took place between runs 119k and 120k. Before the intervention the N_{pe} value for C_4F_{10} remains suppressed at $\sim$ 22 photoelectrons, the same level seen towards the end of 2011. The L0 rate was maintained at $\sim$ 1 MHz throughout 2012, except for runs 114k-114.5k where the rate was kept low ($\sim$ 400 kHz) following a technical stop. Non-statistical fluctuations in N_{pe} over smaller timescales can also be seen, similar to that observed in 2011. After the intervention the C_4F_{10} N_{pe} is seen to increase to $\sim$ 25 photoelectrons, recovering the performance lost due to the increase in L0 rate. The N_{pe} fluctuations are reduced and more consistent with statistical fluctuations post-intervention, indicating a stabilisation in the RICH1 performance. A similar improvement is seen in CF_4, where the readout pixel adjustment increases N_{pe} from $\sim$ 15.5 to $\sim$ 17 photoelectrons, although the early 2011 N_{pe} is not completely recovered as a RICH2 readout pixel adjustment that recovered $\sim$ 100% HPD detection efficiency by the test pulse laser was not achieved.

The effect of the N_{pe} recovery after the HPD pixel reconfiguration in 2012 can be seen in the RICH PID performance, shown in Fig. 4.15. The pre-intervention run groups of 111761-113146 and 114205-114287 (shown in blue in Fig. 4.15), which both have a low L0 rate, have a $\pi \to K$ misID rate of $\sim$ 8% at 95% kaon ID efficiency. This increases to $\sim$ 12% for the pre-intervention runs which have an L0 rate of $\sim$ 1 MHz. After the intervention the PID performance returns to the level seen for low L0 rate events.

[3] A period of intervention at the LHC when no proton beam is present in the accelerator.



(a) C_4F_{10}



(b) CF_4

Figure 4.14: The time variation of the N_{pe} yield of $pp \to pp\mu^+\mu^-$ events in 2012 for the C_4F_{10} and CF_4 radiators. Run number bins with less than 10 track N_{pe} values are not shown.

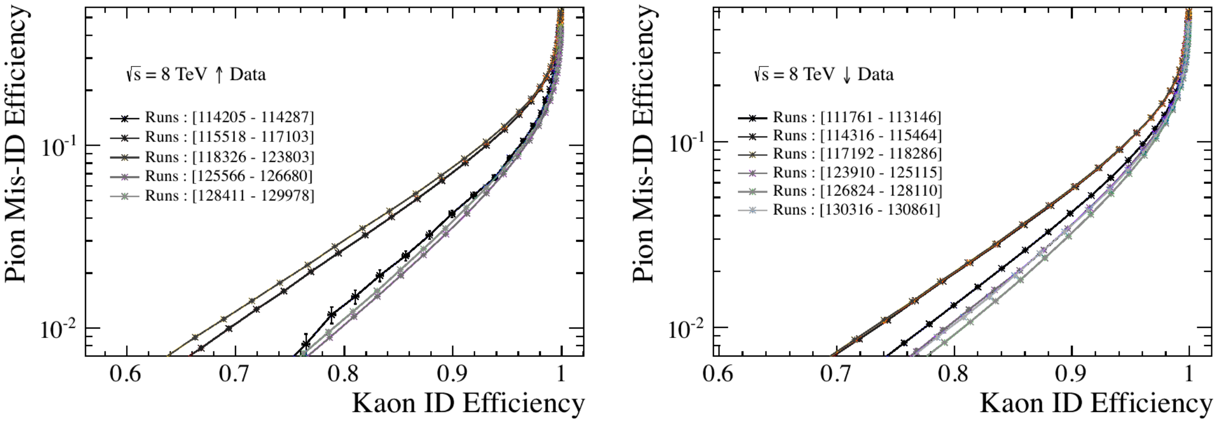


Figure 4.15: Plots showing the kaon ID vs pion to kaon misID rates for 2012 data collected in the ‘magnet up’ (left) and ‘magnet down’ (right) polarity configurations. This analysis used the pion and kaon calibration samples described in Appendix B. The performance is shown for the different groups of run number denoted in the legend. The performance is seen to be unaffected by the magnet polarity.

4.6 Conclusions

The photoelectron yield in 2011 is found to be $\sim 15\%$ less than that expected from simulated events, implying that one or more of the photon detection efficiency terms in (4.1) is smaller than in the simulation. The result is a slight reduction in the PID performance. Despite this the performance is still sufficient to achieve the physics goals of LHCb. The stability of the photon yield throughout 2011 and 2012 indicates that there is no deterioration of the HPD quantum efficiency, mirror reflectivity or RICH medium transparency over the course of 2011 and 2012 data collection. The configuration of the HPD pixel readout chips was seen to have an effect on N_{pe}. The precise cause of this dependence is being investigated by the designers of the readout chips. Continuous monitoring of N_{pe} will be maintained for the data collection period following the 2014 consolidation upgrade of LHCb.

Chapter 5

Measurement of the p_T and η dependence of $f_{\Lambda_b^0}/f_d$.

This chapter describes the measurement of the p_T and η dependence of the $b \to \Lambda_b^0$ to $b \to B^0$ hadronisation ratio, using LHCb data collected in 2011 with $\sqrt{s} = 7\,\mathrm{TeV}$ corresponding to an integrated luminosity of 1.0 fb^{-1}. The results of this analysis will be submitted to the Journal of High Energy Physics. The author contributed to all parts of the analysis, except for the selection development and all items referring to the $\overline{B}^0 \to D^+\pi^-$ decay. These are detailed for reference and completeness only.

5.1 Introduction

The probabilities for a b quark hadronising into a B meson of a given flavour (f_u,f_d,f_s,f_c) or a Λ_b^0 baryon ($f_{\Lambda_b^0}$) have to be extracted experimentally. They are difficult to calculate analytically because, at the energy scale at which fragmentation occurs, α_{QCD} becomes large enough such that QCD cannot be solved perturbatively. Knowledge of the b quark hadronisation rates is important for many B-physics analyses, for example, the search for rare $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays described in Sect. 6. The ratio $f_{\Lambda_b^0}/(f_d + f_u)$ has been measured at CDF [56] with 360 pb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\mathrm{TeV}$, and at LHCb [57] with 3 pb^{-1} of pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$, both using semileptonic $\Lambda_b^0 \to \Lambda_c^+ l\nu$ and $B^0 \to Dl\nu$ decays. $f_{\Lambda_b^0}/(f_d + f_u)$ was also measured at the ALEPH detector at LEP using $\Lambda_b^0 \to pX$ and $B^0 \to pX$ decays produced via $Z^0 \to \bar{b}b$ [58]. These results have been collated by the Heavy Flavour Averaging Group (HFAG) [17]. The $f_{\Lambda_b^0}/(f_d + f_u)$ values obtained from the three experiments are shown in Fig. 5.1. A dependence is seen with respect to the p_T of the charmed hadron + lepton system where linear and exponential functions both provide valid fits to the data points in Fig. 5.1.

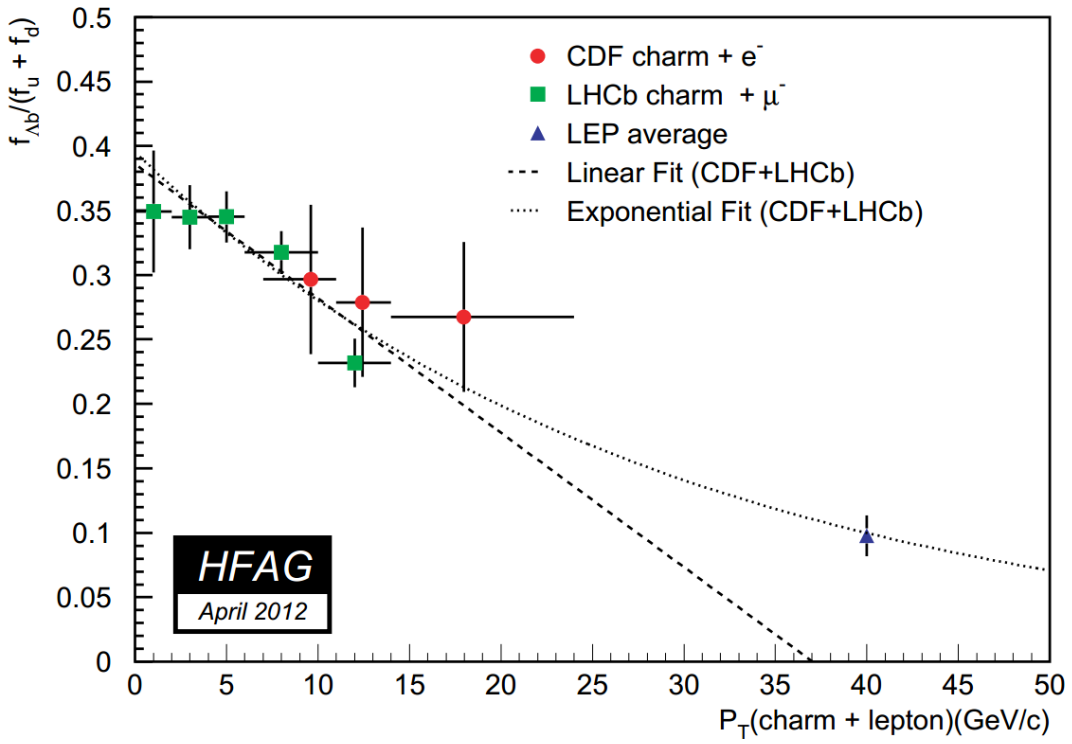


Figure 5.1: The Λ_b^0 production fraction as measured by LEP, LHCb with 2010 data and CDF in bins of p_{T} (charm+lepton). Linear (long-dashed) and exponential (short-dashed) fits to the LHCb and CDF data points are also shown. Taken from [17].

5.1.1 Analysis strategy

The analysis presented in this chapter extracts the dependence on p_{T} and η of $f_{\Lambda_b^0}/f_d$, using 1.0 fb^{-1} of LHCb data collected in 2011 at $\sqrt{s} = 7\,\mathrm{TeV}$. This is used to check the validity of the $f_{\Lambda_b^0}/(f_d + f_u)$ fit functions shown in Fig. 5.1, as $f_{\Lambda_b^0}/f_d$ is equivalent to $2 \times f_{\Lambda_b^0}/(f_d + f_u)$, given that $f_u = f_d$. This is due to the is due to the isospin symmetry between the B^0 and the B^+ mesons. The dominant mechanism for $b \to B^0$ and $b \to B^+$ hadronisation is by combination with a u or d quark originating from a quark-antiquark pair produced by a gluon. As u and d quarks have the same colour charge, the rate of a gluon splitting to $u\bar{u}$ is effectively the same as to $d\bar{d}$, resulting in $f_u = f_d$. The measurement is made by extracting the ratio $\mathcal{R}$ of the yields of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ decays, corrected for reconstruction and selection efficiencies. $\mathcal{R}$ is measured in 20 bins of p_{T} (H_b^0) and 10 bins of $\eta(H_b^0)$, where H_b^0 refers to both the Λ_b^0 and B^0 from hereon:

$$\mathcal{R}(p_{\mathrm{T}}, \eta) = \left(\frac{N_{\Lambda_b^0 \to \Lambda_c^+\pi^-}(p_{\mathrm{T}}, \eta) \times \epsilon_{\overline{B}^0 \to D^+\pi^-}(p_{\mathrm{T}}, \eta)}{N_{\overline{B}^0 \to D^+\pi^-}(p_{\mathrm{T}}, \eta) \times \epsilon_{\Lambda_b^0 \to \Lambda_c^+\pi^-}(p_{\mathrm{T}}, \eta)} \right), \tag{5.1}$$

where N_X is the total number of selected X decay candidates, and ϵ_X is the total detection efficiency for measuring X decays. The binned values of $N_{\overline{B}^0 \to D^+\pi^-}$ and $\epsilon_{\overline{B}^0 \to D^+\pi^-}$ have already been measured by the LHCb analysis of f_s/f_d, using 1 fb^{-1} of $\sqrt{s} = 7\,\mathrm{TeV}$ data [18]. These results are used in the analysis reported here. The only new measurements are the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ yields and efficiencies. The selection and measurement of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays are designed to be as similar as possible to those of the $\overline{B}^0 \to D^+\pi^-$ decays. This ensures that the efficiencies of the two modes are very similar and that potential systematic uncertainties are minimised. $\mathcal{R}$ is related to $f_{\Lambda_b^0}/f_d$ by a ratio of branching fractions (denoted by $\mathcal{B}$):

$$\frac{f_{\Lambda_b^0}}{f_d}(p_{\mathrm{T}},\eta) = \left(\frac{\mathcal{B}(\overline{B}^0 \to D^+\pi^-)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\pi^-)}\frac{\mathcal{B}(D^+ \to K^-\pi^+\pi^+)}{\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)}\right) \times \mathcal{R}(p_{\mathrm{T}},\eta). \tag{5.2}$$

The dominant sources of uncertainty in (5.2) are those of $\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\pi^-)$=$(5.7^{+4.0}_{-2.6}) \times 10^{-3}$ and $\mathcal{B}(\Lambda_c^+ \to pK^-\pi^+)$=$(5.0 \pm 1.3) \times 10^{-3}$, resulting in a combined uncertainty of $^{+74.8}_{-52.5}$%. The equivalent branching-fraction-related uncertainty of the semileptonic LHCb measurement of $f_{\Lambda_b^0}/(f_d + f_u)$ is ± 27.2%. As such the hadronic analysis reported here cannot measure the scale of $f_{\Lambda_b^0}/f_d$ more precisely than the semileptonic LHCb analysis. Only the scale-independent component of the $f_{\Lambda_b^0}/f_d$ dependence, which is encapsulated in $\mathcal{R}$, is measured in the hadronic analysis.

5.2 Dataset

The analysis reported in this chapter is performed on data collected by the LHCb detector at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ between March and October 2011, corresponding to an integrated luminosity of 1.0 fb^{-1}. The dominant part of the data was delivered by the LHC with a 50 ns bunch spacing scheme and 1380 circulating bunches, with 1296 colliding bunches in LHCb. The peak luminosity in LHCb was continuously leveled in order not to exceed $3 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for the first part of data taking (370 pb^{-1}), and $3.5 - 4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for the remaining part of the data. The average number of visible pp interactions per bunch crossing was ~ 1.5.

5.3 Simulation samples

Simulated samples of both signal decays and their associated backgrounds are generated. These are listed in Table 5.1. They are used to calculate efficiencies and to determine the invariant mass shapes for background decays. Each simulated event consists of a pp interaction at $\sqrt{s} = 7\,\mathrm{TeV}$, which is forced to produce a single instance of the signal decay. The resonant structures of the intermediate D^+, D_s^+ and Λ_c^+ decays are modeled in the simulation samples, based on studies of these decays performed at the E791 [59], CLEO [60] and Babar [61] experiments.

Channel	Events
signal channels :	
$\Lambda_b^0 \to \Lambda_c^+ (\to pK^-\pi^+)\ \pi^-$	2M
$B^0 \to D^+ (\to K^-\ \pi^+\ \pi^+)\ \pi^-$	1M
background channels:	
$\Lambda_b^0 \to \Lambda_c^+ (\to pK^-\pi^+)\ K^-$	2M
$B_s^0 \to D_s^+ (\to K^-\ K^+\ \pi^+)\ \pi^-$	1M
$B^0 \to D^{*+} (\to D^+ (\to K^-\ \pi^+\ \pi^+)\ \pi^0)\ \pi^-$	1M
$B^0 \to D^+ (\to K^-\ \pi^+\ \pi^+)\ \rho^-\ (\to \pi^-\ \pi^0)$	1M
$B^0 \to D^+ (\to K^-\ \pi^+\ \pi^+)\ K^-$	1M

Table 5.1: Monte Carlo simulation samples used in the reported analysis. The B^0 and B_s^0 decay mode simulation samples are the same ones used for the LHCb f_s/f_d analysis [18].

5.4 Event Selection

The selection of $\Lambda_b^0 \to \Lambda_c^+\pi^-$ candidates is designed to be as similar as possible to that of the $\overline{B}^0 \to D^+\pi^-$ channel. This ensures that potential systematic effects common to both channels cancel out, such as those related to differences between the simulation and data. The $\overline{B}^0 \to D^+\pi^-$ selection is described here for completeness. It was already optimised for the f_s/f_d analysis [18].

5.4.1 Trigger and pre-selection

An initial pre-selection is used to reconstruct and select $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ candidates in the LHCb dataset before the full offline selection. The pre-selection criteria are listed in Table 5.2. These are identical for both channels, except for the invariant mass requirements. For both channels, the decays of the D^+ and Λ_c^+ (referred to as 'H_c^+' from hereon) are reconstructed from three tracks which make a high quality vertex that is displaced from the PV. The invariant mass of the H_c^+ is required to be within $\pm 100\,\text{MeV}/c^2$ of the nominal Λ_c^+ or D^+ mass. The decay of the B^0 and Λ_b^0 (referred to as 'H_b^0') is subsequently reconstructed by requiring a pion and the H_c^+ to make a common vertex displaced from the PV. The H_b^0 trajectory is required to intersect the PV by the $\chi^2_{IP} < 25$ and $\cos(\theta_{PV}) > 0.999$ selections. The B^0 (Λ_b^0) candidate mass is required to be less than $7000\,\text{MeV}/c^2$ and greater than 4750 (5200) MeV/c^2, to allow for the evaluation of background decays in the remaining mass range.

The trigger requirements are included in the pre-selection. These are also identical for both channels. The requirements are designed to select candidates which are consistent with a multi-body decay with hadronic final-state particles and a vertex which is displaced from the PV. At the hardware L0 trigger stage, the event is selected if either the L0Hadron line is triggered by the candidate decay products, or any L0 line is triggered by other non-signal candidate particles in the event. At least one of the candidate decay products is required to trigger the HLT1TrackAllL0 line. The candidate decay is then required to trigger on at least one of the HLT2Topo$n$Body lines. These trigger lines are described in Sect. 3.6.

Selection variable	Criteria
H_c^+ system	
$(\pi^+ + \pi^+/p^+ + K^-)\ p_{\mathrm{T}}$	$> 1.8\,\mathrm{GeV}$
vertex DOCA$_{max}$	$< 0.5\,\mathrm{mm}$
vertex χ^2/DOF	< 10
flight distance χ^2	> 36
$\cos(\theta_{PV})$	> 0
H_b^0 system	
$(\pi^+ + \pi^+/p^+ + K^- + \pi^-)\ p_{\mathrm{T}}$	$> 5\,\mathrm{GeV}$
vertex χ^2/DOF	< 10
χ^2_{IP}	< 25
lifetime	> 0.2 ps
$\cos(\theta_{PV})$	> 0.999
flight distance χ^2	> 36
K,p^+ and π	
track χ^2/DOF	< 4
p_{T}	> 100 MeV
P	> 1000 MeV
χ^2_{IP}	> 4
At least 1 track with	
P	> 10000 MeV
p_{T}	> 1700 MeV
track χ^2/DOF	< 2.5
χ^2_{IP}	> 16
At least 2 tracks with	
P	> 5000 MeV
p_{T}	> 500 MeV
track χ^2/DOF	< 3
Trigger Requirements	
L0	L0Hadron TOS or any L0 line TIS
HLT1	HLT1TrackAllL0 TOS
HLT2	HLT2Topo$n$Body TOS
$\overline{B}^0 \to D^+\pi^-$ only	
$M_{K^-\pi^+\pi^+}$	$1769.6 < M_{K^-\pi^+\pi^+} < 1969.6\,\mathrm{MeV}/c^2$
$M_{K^-\pi^+\pi^+\pi^-}$	$4750 < M_{K^-\pi^+\pi^+\pi^-} < 7000\,\mathrm{MeV}/c^2$
$\Lambda_b^0 \to \Lambda_c^+\pi^-$ only	
$M_{K^-p^+\pi^+}$	$2186.5 < M_{K^-p^+\pi^+} < 2386.5\,\mathrm{MeV}/c^2$
$M_{K^-p^+\pi^+\pi^-}$	$5200 < M_{K^-p^+\pi^+\pi^-} < 7000\,\mathrm{MeV}/c^2$
$\overline{B}_s^0 \to D_s^+\pi^-$ only	
$M_{K^-p^+\pi^+}$	$1940 < M_{K^-K^+\pi^+} < 1990\,\mathrm{MeV}/c^2$
$M_{K^-p^+\pi^+\pi^-}$	$4750 < M_{K^-p^+\pi^+\pi^-} < 7000\,\mathrm{MeV}/c^2$

Table 5.2: The pre-selection criteria used to select $\overline{B}^0 \to D^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}_s^0 \to D_s^+\pi^-$ decays. The $\overline{B}_s^0 \to D_s^+\pi^-$ decays were used in [18] to train a BDT, which is reused in the analysis reported here.

Table 5.3: Summary of the offline selection criteria. The proton PID cut criteria are applied only for protons with p_{T} and η within the range $\mathbb{R}(p_{\mathrm{T}}, \eta)$, defined in Table 5.4. The BDT selection is described in Sect. 5.4.2

Selection	Criteria	
	$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	$\overline{B}^0 \to D^+ \pi^-$
BDT	BDT > 0.66	
$M_{H_c^+}$	$2265 < M_{\Lambda_c^+} < 2305$ MeV/c^2	$1844 < M_{D^+} < 1890$ MeV/c^2
$M_{H_b^0}$	$5350 < M_{\Lambda_b^0} < 6000$ MeV/c^2	$5000 < M_{\overline{B}^0} < 5800$ MeV/c^2
π from H_b^0	$\mathrm{DLL}_K < 0$	
π from H_c^+	$\mathrm{DLL}_K < 5$	
K	$\mathrm{DLL}_K > 0$	
p^+	($\mathrm{DLL}_p > 5$ & $\mathrm{DLL}_{pK} > 0$ & $(p_{\mathrm{T}},\eta) \in \mathbb{R}(p_{\mathrm{T}}, \eta)$) or $((p_{\mathrm{T}},\eta) \ni \mathbb{R}(p_{\mathrm{T}}, \eta))$	N/A
H_c^+ flight distance w.r.t. B^0 vertex	N/A	D^+ flight distance $\chi^2 > 2$

5.4.2 Offline Selection

A final offline selection is applied to the candidates which pass the trigger and pre-selection requirements. This selection is designed to remove as many background candidates as possible whilst retaining the maximum number of signal candidates. The invariant mass requirements on the Λ_c^+ and D^+ systems are tightened from the pre-selection. The Λ_b^0 and B^0 mass-ranges are defined such that both the H_b^0 signal peak and partially reconstructed background decays are seen. These backgrounds are included so that their invariant mass distributions can be evaluated. For the $\overline{B}^0 \to D^+\pi^-$ decay a cut is applied to the flight distance χ^2 of the D^+ decay vertex with respect to the B^0 vertex. This exploits the fact that the D^+ travels a short distance before it decays. No such cut is placed on the Λ_c^+ decay vertex as its lifetime is too short for it to be distinguishable from the Λ_b^0 vertex.

PID cuts are applied to all final-state particles to suppress background from similar four-body H_b^0 decays with different final-state particles. For example, a $\mathrm{DLL}_K < 0$ cut is applied on the π from the H_b^0 decay (referred to as the 'bachelor' pion from hereon) to remove $H_b^0 \to H_c^+ K^-$ decays. The proton PID cut is applied only for protons within a kinematic range $\mathbb{R}(p_{\mathrm{T}}, \eta)$, defined in Table 5.4. Protons outside of this range have no PID cut applied. Fig. 5.2 illustrates which protons of the signal $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay have the cut applied. The range defines the boundaries of the kinematic distribution of the protons from the proton PID calibration dataset, which is described in Appendix B.3. Applying the PID cut only in this range ensures that the efficiency of the selection can be calculated precisely (see Sect. 5.5.3).

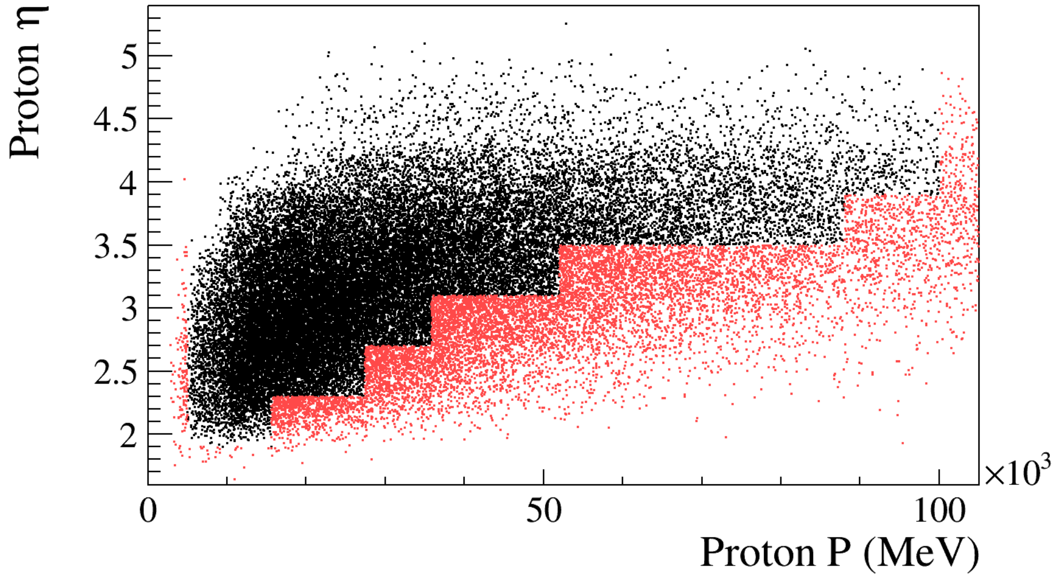


Figure 5.2: A plot of the (P,η) distribution of selected proton candidates from $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays. The black data points lie within the PID selection range $\mathbb{R}(p_{\mathrm{T}}, \eta)$, those in red (grey in B&W) do not. The plot is made from selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in data, with a $\pm 40\,\mathrm{MeV}/c^2$ mass window applied around the Λ_b^0 mass.

Table 5.4: Summary of the proton kinematic range $\mathbb{R}(p_{\mathrm{T}}, \eta)$, within which the proton PID selection is applied. No cuts are applied to protons outside of $\mathbb{R}(p_{\mathrm{T}}, \eta)$.

proton P(GeV/c)	proton η
5< P <15.6	1.9< η
15.6< P <27.4	2.3< η
27.4< P <35.8	2.7< η
35.8< P <52.0	3.1< η
52.0< P <88.0	3.5< η
88.0< P <100.0	3.9< η

BDT selection

Background candidates that were not formed from single H_b^0 decays, referred to as 'combinatorial background', are suppressed using a Boosted Decision Tree (BDT) classifier. The output of the classifier (referred to as '$\mathrm{BDT}_{\mathrm{out}}$' from hereon) has a range from -1 to 1, where candidates with $\mathrm{BDT}_{\mathrm{out}}$ closer to 1 are defined as being more signal-like and those closer to -1 more background-like. The BDT classifier is produced by a multivariate selection algorithm which exploits correlations between different input variables to distinguish between signal and background candidates. A more detailed description of BDT's can be found in reference [62].

The BDT configuration remains unchanged from the f_s/f_d analysis [18]. It was trained using preselected $\overline{B}_s^0 \to D_s^+\pi^-$ events from the full 2011 dataset. The dataset was split

Table 5.5: Kinematic variables used to train the BDT.

B_s^0, D_s^+	K^+, K^-, π^+, π^-
$\cos(\theta_{PV})$	χ^2_{IP}
χ^2_{IP}	p_{T}
FD χ^2 (w.r.t. PV)	
FD χ^2 (w.r.t. PV)	

into two samples, the 'training' sample was used to train the BDT, the 'testing' sample was used to evaluate the BDT output. In each sample, the invariant D_s^+ π^- mass distribution was fitted with a Gaussian signal and exponential background PDF. Signal events were selected by assigning weights $W_{M(D_s^+\pi^-)} = (PDF_{\mathcal{S}}/PDF_{\mathcal{S}+\mathcal{B}})_{M(D_s^+\pi^-)}$ to decays within a D_s^+ π^- invariant mass range of 5310-5430 MeV/c^2, where $PDF_{\mathcal{S}}(PDF_{\mathcal{B}})$ is the signal (background) PDF value for a candidate decay with invariant mass $M(D_s^+\pi^-)$. Background events were selected from the mass region $M(D_s^+\pi^-) > 5445\,\mathrm{MeV}/c^2$, where only combinatorial background events remain. The BDT was trained with the kinematic variables listed in Table 5.5. The most discriminating variables were found to be the $\cos(\theta_{PV})$ and χ^2_{IP} of the B_s^0, and the individual p_{T} of the final-state particles. The optimum BDT output cut was found to be $BDT_{\mathrm{out}} > 0.66$, as this yielded the maximum value of the signal significance, defined as $N_{\mathcal{S}}/\sqrt{N_{\mathcal{S}} + N_{\mathcal{B}}}$, where $N_{\mathcal{S}}(N_{\mathcal{B}})$ is the number of signal (background) events. The signal significance after the $BDT_{\mathrm{out}} > 0.66$ selection was found to be 24.4 ± 0.7 for the training sample and 23.6 ± 0.7 for the testing sample, the consistency of these values indicates that the BDT was not overtrained.

5.5 Efficiencies

The efficiencies with which the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ decays are reconstructed and selected are determined from the simulated signal event samples listed in Table 5.1. The $\overline{B}^0 \to D^+\pi^-$ efficiencies listed in Tables 5.6 to 5.13 have already been calculated in [18].

5.5.1 Acceptance Efficiencies

Tables 5.6 and 5.7 show the acceptance efficiency in each p_{T} and η bin for both $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ decay modes. The ratio of efficiencies is also shown. A clear trend of increasing efficiency with p_{T} and η is seen for both channels. This is a result of the H_b^0 momentum being greater for higher p_{T} and η bins. For an increase in the H_b^0 momentum, the component of the momentum of the decay products which is parallel to the H_b^0 flight direction increases whilst the perpendicular component remains unchanged. As a result the angles between the trajectories of the decay products become smaller in the $p - p$ collision frame of reference. This means that if one final-state particle is within the acceptance, the others are also more likely to be. The result is an increase in acceptance efficiency with increasing p_{T} (H_b^0) and $\eta(H_b^0)$.

A small difference between the Λ_b^0 and B^0 decay efficiencies is seen at the lower p_{T} and η bins, with the B^0 decay having a lower efficiency than the Λ_b^0 decay. This is due to small differences in the kinematics of the Λ_c^+ and D^+ decays, caused by the different masses of the final-state particles and the different resonant structures of the decays (discussed in [59] and [61]). The differences vanish in the higher p_{T} and η bins, where the kinematics of the H_b^0 decays are more strongly influenced by the high momentum of the H_b^0.

Table 5.6: The acceptance efficiencies of $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays, in bins of $p_{\rm T}(H_b^0)$.

$p_{\rm T}(H_b^0)$ (GeV/c)	$\epsilon(\overline{B}^0 \to D^+\pi^-)(\%)$	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)(\%)$	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
1.5 - 3.6	23.5 ± 0.1	25.1 ± 0.1	0.937 ± 0.002
3.6 - 4.5	29.7 ± 0.1	31.1 ± 0.1	0.955 ± 0.003
4.5 - 5.3	32.3 ± 0.1	33.9 ± 0.1	0.951 ± 0.003
5.3 - 6.0	34.7 ± 0.1	35.8 ± 0.1	0.969 ± 0.003
6.0 - 6.5	36.2 ± 0.1	37.4 ± 0.1	0.968 ± 0.004
6.5 - 7.0	37.4 ± 0.1	38.5 ± 0.1	0.971 ± 0.004
7.0 - 7.5	38.5 ± 0.2	39.5 ± 0.2	0.973 ± 0.005
7.5 - 8.0	39.3 ± 0.2	40.7 ± 0.2	0.966 ± 0.005
8.0 - 8.5	40.2 ± 0.2	41.3 ± 0.2	0.972 ± 0.005
8.5 - 9.0	41.2 ± 0.2	42.1 ± 0.2	0.98 ± 0.005
9.0 - 9.5	41.4 ± 0.2	42.8 ± 0.2	0.966 ± 0.005
9.5 - 10.0	42.4 ± 0.2	43.3 ± 0.2	0.981 ± 0.005
10.0 - 10.7	42.9 ± 0.2	44.0 ± 0.2	0.976 ± 0.005
10.7 - 11.5	43.8 ± 0.2	44.7 ± 0.2	0.98 ± 0.005
11.5 - 12.2	44.1 ± 0.2	45.2 ± 0.2	0.975 ± 0.006
12.2 - 13.0	44.7 ± 0.2	45.6 ± 0.2	0.98 ± 0.006
13.0 - 14.3	45.6 ± 0.2	46.3 ± 0.2	0.985 ± 0.005
14.3 - 16.0	46.5 ± 0.2	47.4 ± 0.2	0.98 ± 0.004
16.0 - 20.2	47.6 ± 0.2	47.7 ± 0.1	0.997 ± 0.004
20.2 - 40.0	48.6 ± 0.2	48.9 ± 0.1	0.995 ± 0.004

Table 5.7: The acceptance efficiencies of $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays, in bins of $\eta(H_b^0)$.

$\eta(H_b^0)$	$\epsilon(\overline{B}^0 \to D^+\pi^-)(\%)$	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)(\%)$	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
2.00 - 2.60	15.9 ± 0.1	14.5 ± 0.1	0.916 ± 0.004
2.60 - 2.75	28.0 ± 0.1	25.9 ± 0.1	0.926 ± 0.005
2.75 - 2.90	32.1 ± 0.1	30.0 ± 0.1	0.934 ± 0.004
2.90 - 3.05	35.6 ± 0.1	33.7 ± 0.1	0.948 ± 0.004
3.05 - 3.20	38.8 ± 0.1	36.9 ± 0.1	0.953 ± 0.003
3.20 - 3.35	41.1 ± 0.1	39.6 ± 0.1	0.961 ± 0.003
3.35 - 3.50	43.2 ± 0.1	41.8 ± 0.1	0.968 ± 0.003
3.50 - 3.65	44.4 ± 0.1	43.3 ± 0.1	0.974 ± 0.003
3.65 - 4.00	45.3 ± 0.1	44.7 ± 0.1	0.988 ± 0.002
4.00 - 5.00	41.3 ± 0.1	41.5 ± 0.1	1.003 ± 0.002

5.5.2 Combined reconstruction and pre-selection efficiencies

The efficiencies with which $H_b^0 \to H_c^+\pi^-$ decays within the LHCb acceptance are reconstructed and pass the pre-selection requirements (which includes the trigger requirements) is shown in Tables 5.8 and 5.9. The $\overline{B}^0 \to D^+\pi^-$ to $\Lambda_b^0 \to \Lambda_c^+\pi^-$ efficiency ratio fluctuates around 1.15 across all bins, with no trend apparent. The $\Lambda_b^0 \to \Lambda_c^+\pi^-$ efficiency is lower than the $\overline{B}^0 \to D^+\pi^-$ efficiency because the pions from the Λ_c^+ decay typically have a lower p_{T} than the pions from the D^+ decay. The mean p_{T} before reconstruction of the former is 0.8 GeV/c compared to 1.2 GeV/c for the latter. Decay products with a low p_{T} are less likely to satisfy the pre-selection requirements, which contain multiple explicit and implicit selections based on the p_{T} of the final-state particles. Three trends are apparent in the efficiencies. First, the efficiency increases with p_{T} (H_b^0). This is due to the p_{T} of the final-state particles increasing with p_{T} (H_b^0). As such they are more likely to satisfy the pre-selection requirements. Second, the efficiency increases with decreasing $\eta(H_b^0)$, for $\eta(H_b^0) > 2.9$. η is inversely correlated with p_{T}, hence the final-state particles of decays with lower $\eta(H_b^0)$ will have higher p_{T}, resulting in increased pre-selection efficiency. Third, for $\eta(H_b^0) < 2.9$ the efficiency decreases with $\eta(H_b^0)$. The low $\eta(H_b^0)$ decay products have low P, as P is correlated with η. Particles with low momenta are more likely to be swept outside of the LHCb acceptance by the magnet. This is especially the case for low $\eta(H_b^0)$ decays, which are closer to the outer acceptance boundary than high $\eta(H_b^0)$ decays. This trend dominates the other η trend described above in the four lowest $\eta(H_b^0)$ bins.

Table 5.8: The combined reconstruction and pre-selection efficiencies for the $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay channels, and their ratio, in bins of p_{T} (H_b^0).

$p_{\mathrm{T}}(H_b^0)$ (GeV/c)	$\epsilon(\overline{B}^0 \to D^+\pi^-)$ (%)	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$(%)	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
1.5 - 3.6	0.59 ± 0.01	0.54 ± 0.01	1.096 ± 0.033
3.6 - 4.5	1.35 ± 0.03	1.15 ± 0.02	1.168 ± 0.037
4.5 - 5.3	1.80 ± 0.04	1.58 ± 0.03	1.139 ± 0.035
5.3 - 6.0	2.37 ± 0.06	2.16 ± 0.04	1.098 ± 0.033
6.0 - 6.5	3.24 ± 0.09	2.83 ± 0.06	1.146 ± 0.038
6.5 - 7.0	3.67 ± 0.10	3.46 ± 0.07	1.061 ± 0.035
7.0 - 7.5	4.37 ± 0.11	4.17 ± 0.08	1.047 ± 0.034
7.5 - 8.0	5.54 ± 0.14	4.60 ± 0.09	1.202 ± 0.038
8.0 - 8.5	6.40 ± 0.16	5.25 ± 0.10	1.218 ± 0.038
8.5 - 9.0	6.72 ± 0.17	6.18 ± 0.12	1.087 ± 0.035
9.0 - 9.5	7.83 ± 0.20	6.81 ± 0.14	1.149 ± 0.037
9.5 - 10.0	8.68 ± 0.23	7.69 ± 0.16	1.129 ± 0.037
10.0 - 10.7	9.14 ± 0.22	8.28 ± 0.15	1.103 ± 0.033
10.7 - 11.5	10.22 ± 0.24	9.63 ± 0.17	1.061 ± 0.031
11.5 - 12.2	11.72 ± 0.30	10.18 ± 0.21	1.152 ± 0.038
12.2 - 13.0	12.63 ± 0.33	10.73 ± 0.22	1.177 ± 0.039
13.0 - 14.3	13.19 ± 0.30	12.19 ± 0.21	1.082 ± 0.031
14.3 - 16.0	14.80 ± 0.34	12.92 ± 0.24	1.145 ± 0.034
16.0 - 20.2	16.06 ± 0.32	14.06 ± 0.23	1.142 ± 0.029
20.2 - 40.0	15.11 ± 0.39	14.91 ± 0.29	1.014 ± 0.033

Table 5.9: The combined reconstruction and pre-selection efficiencies for the $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay channels, and their ratio, in bins of $\eta(H_b^0)$.

$\eta(H_b^0)$	$\epsilon(\overline{B}^0 \to D^+\pi^-)$ (%)	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$ (%)	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
2.00 - 2.60	2.89 ± 0.05	2.44 ± 0.03	1.186 ± 0.025
2.60 - 2.75	4.51 ± 0.09	3.81 ± 0.06	1.182 ± 0.031
2.75 - 2.90	4.62 ± 0.09	3.99 ± 0.06	1.156 ± 0.028
2.90 - 3.05	4.74 ± 0.09	4.06 ± 0.06	1.167 ± 0.027
3.05 - 3.20	4.58 ± 0.08	3.89 ± 0.06	1.179 ± 0.027
3.20 - 3.35	4.36 ± 0.08	3.80 ± 0.05	1.148 ± 0.027
3.35 - 3.50	4.29 ± 0.08	3.65 ± 0.05	1.176 ± 0.028
3.50 - 3.65	4.05 ± 0.08	3.63 ± 0.05	1.117 ± 0.028
3.65 - 4.00	3.53 ± 0.05	3.06 ± 0.03	1.151 ± 0.021
4.00 - 5.00	1.66 ± 0.03	1.43 ± 0.02	1.161 ± 0.023

5.5.3 Offline selection efficiencies

The efficiencies for reconstructed and pre-selected candidates passing the offline selection are shown in Table 5.10 and Table 5.11. All ratios are close to 1.0 for all bins, as expected from the similarity of the selection criteria for both channels. The individual $H_b^0 \to H_c^+\pi^-$ efficiencies increase with p_{T} and decrease with η. This is primarily due to the BDT favouring decays with high p_{T} and low η, as these are more easily distinguishable from combinatorial background events. The PID efficiency also increases for higher p_{T} (lower η) decays, as the Cherenkov cones produced by the decay products in the RICH system fall in regions of the HPD plane with less photoelectron background (see Sect. 3.4.1).

The PID efficiencies are calculated from data using the kaon, pion and proton PID calibration samples described in Appendix B. These are divided into bins of momentum, p_{T} and charged track multiplicity ($nTracks$), as the PID efficiency is dependent on these variables. Each simulated particle which has a DLL selection applied is assigned a weight corresponding to the efficiency of the selection when applied to the calibration sample in the appropriate $(P, p_{\mathrm{T}}, nTracks)$ bin. The weights of each DLL selection are multiplied together to produce an event weight corresponding to the full PID selection. The overall PID efficiency is then taken as the mean of these weights. The proton PID calibration sample is only used to calculate efficiencies within the $\mathbb{R}(p_{\mathrm{T}}, \eta)$ range shown in Table 5.4, as there are not enough protons in the calibration sample outside of $\mathbb{R}(p_{\mathrm{T}}, \eta)$ to calculate the weights accurately. Since no PID cuts are applied to protons outside of $\mathbb{R}(p_{\mathrm{T}}, \eta)$, their assigned efficiency weight is 1.0 ± 0.0 by definition.

5.5.4 Combined efficiencies

Tables 5.12 and 5.13 show the combined, acceptance, reconstruction pre-selection and offline selection efficiencies for both channels and their ratio. The values of this ratio are between 1.0 and 1.2 for all bins, with no strong trend in either binning scheme.

Table 5.10: Offline selection efficiencies for $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays in bins of $p_{\mathrm{T}}(H_b^0)$.

$p_{\mathrm{T}}(H_b^0)$ (GeV/c)	$\epsilon(\overline{B}^0 \to D^+\pi^-)$ (%)	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$ (%)	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
1.5 - 3.6	41.53 ± 1.06	38.45 ± 0.81	0.926 ± 0.031
3.6 - 4.5	46.45 ± 1.11	44.70 ± 0.89	0.962 ± 0.030
4.5 - 5.3	49.77 ± 1.04	48.06 ± 0.85	0.966 ± 0.026
5.3 - 6.0	52.80 ± 1.03	50.73 ± 0.86	0.961 ± 0.025
6.0 - 6.5	51.49 ± 1.09	51.47 ± 0.96	1.000 ± 0.028
6.5 - 7.0	52.22 ± 1.08	52.10 ± 0.92	0.998 ± 0.027
7.0 - 7.5	52.47 ± 1.06	52.21 ± 0.91	0.995 ± 0.027
7.5 - 8.0	55.33 ± 0.99	53.72 ± 0.93	0.971 ± 0.024
8.0 - 8.5	53.35 ± 0.97	54.46 ± 0.89	1.021 ± 0.025
8.5 - 9.0	56.80 ± 1.00	55.48 ± 0.89	0.977 ± 0.023
9.0 - 9.5	56.02 ± 1.04	56.19 ± 0.98	1.003 ± 0.025
9.5 - 10.0	56.43 ± 1.05	56.28 ± 0.98	0.997 ± 0.025
10.0 - 10.7	57.33 ± 0.94	57.06 ± 0.87	0.995 ± 0.022
10.7 - 11.5	58.74 ± 0.89	58.51 ± 0.83	0.996 ± 0.021
11.5 - 12.2	57.94 ± 0.95	58.88 ± 0.94	1.016 ± 0.023
12.2 - 13.0	60.35 ± 0.97	58.37 ± 0.95	0.967 ± 0.022
13.0 - 14.3	56.97 ± 0.86	60.51 ± 0.83	1.062 ± 0.022
14.3 - 16.0	58.69 ± 0.86	59.61 ± 0.88	1.016 ± 0.021
16.0 - 20.2	56.03 ± 0.72	60.07 ± 0.77	1.072 ± 0.019
20.2 - 40.0	54.45 ± 0.94	58.14 ± 0.94	1.068 ± 0.025

Table 5.11: Offline selection efficiencies for $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays in bins of $\eta(H_b^0)$.

$\eta(H_b^0)$	$\epsilon(\overline{B}^0 \to D^+\pi^-)$ (%)	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$ (%)	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
2.00 - 2.60	61.77 ± 0.68	60.76 ± 0.56	0.984 ± 0.014
2.60 - 2.75	60.27 ± 0.82	59.12 ± 0.74	0.981 ± 0.018
2.75 - 2.90	60.80 ± 0.80	59.79 ± 0.68	0.983 ± 0.017
2.90 - 3.05	61.35 ± 0.75	60.06 ± 0.68	0.979 ± 0.016
3.05 - 3.20	59.52 ± 0.75	58.98 ± 0.68	0.991 ± 0.017
3.20 - 3.35	58.06 ± 0.75	56.83 ± 0.67	0.979 ± 0.017
3.35 - 3.50	55.55 ± 0.79	55.41 ± 0.70	0.997 ± 0.019
3.50 - 3.65	50.79 ± 0.78	53.45 ± 0.70	1.052 ± 0.021
3.65 - 4.00	47.21 ± 0.56	48.46 ± 0.52	1.026 ± 0.016
4.00 - 5.00	39.39 ± 0.55	38.52 ± 0.53	0.978 ± 0.019

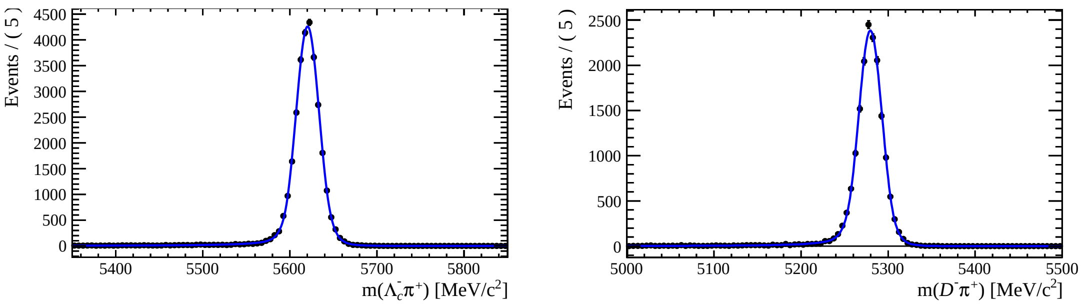
Table 5.12: The combined efficiencies of the $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay channels, and their ratio, in bins of $p_{\mathrm{T}}(H_b^0)$.

$p_{\mathrm{T}}(H_b^0)$ (GeV/c)	$\epsilon(\overline{B}^0 \to D^+\pi^-)$ (%)	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$ (%)	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
1.5 - 3.6	0.058 ± 0.002	0.052 ± 0.002	1.105 ± 0.053
3.6 - 4.5	0.186 ± 0.007	0.160 ± 0.005	1.163 ± 0.054
4.5 - 5.3	0.289 ± 0.010	0.258 ± 0.007	1.122 ± 0.049
5.3 - 6.0	0.434 ± 0.015	0.393 ± 0.010	1.107 ± 0.047
6.0 - 6.5	0.604 ± 0.023	0.545 ± 0.016	1.109 ± 0.052
6.5 - 7.0	0.716 ± 0.027	0.694 ± 0.019	1.031 ± 0.047
7.0 - 7.5	0.882 ± 0.032	0.860 ± 0.023	1.025 ± 0.046
7.5 - 8.0	1.203 ± 0.041	1.005 ± 0.027	1.198 ± 0.051
8.0 - 8.5	1.371 ± 0.047	1.181 ± 0.032	1.161 ± 0.051
8.5 - 9.0	1.573 ± 0.055	1.442 ± 0.038	1.091 ± 0.048
9.0 - 9.5	1.814 ± 0.064	1.638 ± 0.045	1.108 ± 0.049
9.5 - 10.0	2.077 ± 0.074	1.872 ± 0.052	1.109 ± 0.050
10.0 - 10.7	2.248 ± 0.072	2.077 ± 0.051	1.083 ± 0.044
10.7 - 11.5	2.627 ± 0.083	2.516 ± 0.059	1.044 ± 0.041
11.5 - 12.2	2.992 ± 0.104	2.707 ± 0.073	1.105 ± 0.049
12.2 - 13.0	3.403 ± 0.118	2.853 ± 0.080	1.193 ± 0.053
13.0 - 14.3	3.423 ± 0.108	3.412 ± 0.079	1.003 ± 0.039
14.3 - 16.0	4.035 ± 0.127	3.651 ± 0.089	1.105 ± 0.044
16.0 - 20.2	4.284 ± 0.121	4.029 ± 0.086	1.063 ± 0.038
20.2 - 40.0	3.995 ± 0.147	4.235 ± 0.112	0.943 ± 0.043

Table 5.13: The combined efficiencies of the $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay channels, and their ratio, in bins of $\eta(H_b^0)$.

$\eta(H_b^0)$	$\epsilon(\overline{B}^0 \to D^+\pi^-)$ (%)	$\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$ (%)	$\epsilon(\overline{B}^0 \to D^+\pi^-)/\epsilon(\Lambda_b^0 \to \Lambda_c^+\pi^-)$
2.00 - 2.60	0.259 ± 0.006	0.235 ± 0.004	1.102 ± 0.029
2.60 - 2.75	0.704 ± 0.019	0.631 ± 0.013	1.116 ± 0.038
2.75 - 2.90	0.843 ± 0.021	0.766 ± 0.015	1.100 ± 0.035
2.90 - 3.05	0.980 ± 0.024	0.867 ± 0.016	1.13 ± 0.034
3.05 - 3.20	1.006 ± 0.025	0.889 ± 0.017	1.131 ± 0.035
3.20 - 3.35	1.001 ± 0.025	0.888 ± 0.017	1.128 ± 0.035
3.35 - 3.50	0.996 ± 0.026	0.874 ± 0.018	1.140 ± 0.037
3.50 - 3.65	0.890 ± 0.025	0.862 ± 0.018	1.033 ± 0.036
3.65 - 4.00	0.745 ± 0.016	0.671 ± 0.011	1.110 ± 0.030
4.00 - 5.00	0.271 ± 0.007	0.228 ± 0.005	1.191 ± 0.038

Figure 5.3: Plots of the $H_c^+\pi^-$ mass spectra of simulated $H_b^0 \to H_c^+\pi^-$ events, both are fitted with the DCB function (5.3).



(a) $\Lambda_b^0 \to \Lambda_c^+\pi^-$ (b) $\overline{B}^0 \to D^+\pi^-$

5.6 Fit models

The $H_b^0 \to H_c^+\pi^-$ yields are extracted by performing an unbinned extended maximum-likelihood fit to the $H_c^+\ \pi^-$ invariant mass spectra. The H_b^0 mass resolution is improved by fixing ('constraining') the H_c^+ invariant mass to the world average H_c^+ mass [1], instead of obtaining it from the $K^-p/\pi^+\pi^+$ invariant mass. From hereon, references to the $H_c^+\ \pi^-$ invariant masses have the H_c^+ mass constrained. The $\Lambda_b^0 \to \Lambda_c^+\pi^-$ fit model is described in the following section. The $\overline{B}^0 \to D^+\pi^-$ yields are taken from the LHCb f_s/f_d analysis [18]. The $\overline{B}^0 \to D^+\pi^-$ fit model is also described in this section for completeness.

5.6.1 Signal shape

The $H_b^0 \to H_c^+\pi^-$ signal shape PDF's are obtained by fitting the $H_c^+\pi^-$ invariant mass spectra of the simulated signal event samples after application of the full selection. The fit PDF function is a Double Crystal Ball Function (DCB), which is composed of two CB functions, as described in Appendix C. The DCB function is the same as used in the LHCb f_s/f_d analysis, defined as:

$$DCB(m, \bar{m}, \sigma, \alpha_1, \alpha_2, n_1, n_2, f) = CB(m, \bar{m}, \sigma, \alpha_1, n_1) + CB(m, \bar{m}, \sigma, \alpha_2, n_2). \qquad (5.3)$$

The power-law components of the two CB PDFs in (5.3) are required to have $\alpha_1 > 0$ and $\alpha_2 < 0$, such that one PDF fits the mass distribution below $\bar{m}$ (to account for the 'radiative tail', see Appendix C) while the other fits the events above $\bar{m}$ (to account for non-Gaussian detector resolution effects).

The fitted DCB parameters are summarised in Table 5.14. All parameters are consistent for both signal decay modes (except $\bar{m}$, which is different for Λ_b^0 and B^0). Fig. 5.3 shows the fitted invariant mass distributions of the simulated signal events. The DCB shape provides a good fit for both decay modes. The signal fit was also performed in

Table 5.14: The parameters of the CB PDFs for $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays, obtained from simulated signal events after application of the complete selection.

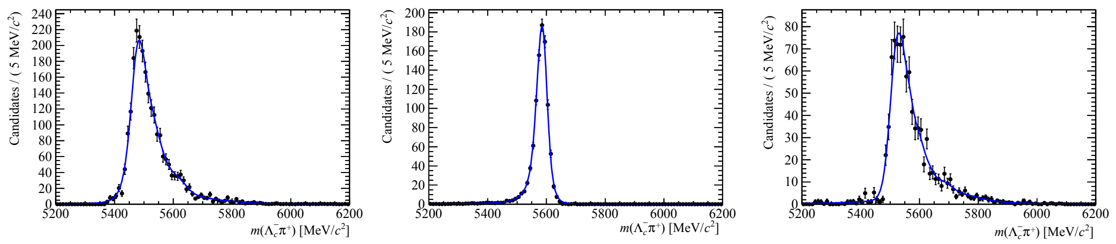
	$\Lambda_b^0 \to \Lambda_c^+\pi^-$	$\overline{B}^0 \to D^+\pi^-$
$\bar{m}$	5620.4 ± 0.1	5279.6 ± 0.1
σ	13.3 ± 0.1	13.4 ± 0.1
α_1	1.74 ± 0.05	1.60 ± 0.05
α_2	-1.60 ± 0.08	-1.74 ± 0.12
n_1	1.38 ± 0.07	1.57 ± 0.10
n_2	9.58 ± 2.31	8.19 ± 2.54
f	0.5	0.5

bins of p_{T} and η and the resulting fit parameters were found to be consistent across all bins, showing that the signal mass shape does not change with $p_{\mathrm{T}}(H_b^0)$ or $\eta(H_b^0)$. In the fit to the data, the σ parameter is left free because the detector resolution in simulation ($\sigma_{\mathrm{sim}} = 13.3 \pm 0.1\,\mathrm{MeV}/c^2$) is slightly better than that seen in the data ($\sigma_{\mathrm{data}} = 15.8 \pm 0.1\,\mathrm{MeV}/c^2$).

5.6.2 Backgrounds from misidentified decays

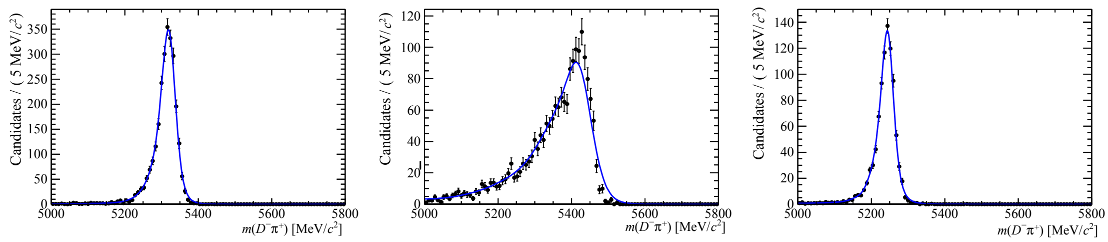
Four-body H_b^0 decays with similar kinematic properties to $H_b^0 \to H_c^+\pi^-$, but with different final-state particles can be misidentified as $H_b^0 \to H_c^+\pi^-$ decays. For the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ signal channel, misidentified backgrounds arise from $\overline{B}^0 \to D^+\pi^-$, $\overline{B}_s^0 \to D_s^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+K^-$ decays. For the $\overline{B}^0 \to D^+\pi^-$ channel the backgrounds are from $\Lambda_b^0 \to \Lambda_c^+\pi^-$, $\overline{B}_s^0 \to D_s^+\pi^-$ and $\overline{B}^0 \to D^+K^-$ decays. The $H_c^+\ \pi^-$ invariant mass distributions of these backgrounds are distorted by the misidentification, as one of the final-state particles is assigned an incorrect mass hypothesis and the misidentification rate of the PID cuts is sensitive to the P and η of the misidentified particle. The $H_c^+\pi^-$ mass distributions of these backgrounds overlap with the signal peaks. It is therefore important to use an accurate description for their contribution in the fit of their mass spectra in the data.

The invariant mass PDF of each background is built from the simulated event samples shown in Table 5.1. The events are reconstructed under the $H_b^0 \to H_c^+\pi^-$ mass hypothesis and then the full $H_b^0 \to H_c^+\pi^-$ selection is applied. The effect of the PID selection on the $H_c^+\pi^-$ invariant mass distribution is obtained by assigning to each event a PID selection efficiency weight, obtained from the data-driven PID reweighting technique described in Sect. 5.5.3. The resulting $H_c^+\pi^-$ mass distributions, shown in Fig. 5.4 and Fig. 5.5, are fitted with a PDF consisting of a superposition of Gaussian kernels, one for each datapoint, which is then simplified by a smoothing algorithm [19]. Where possible, a Gaussian constraint is applied to the number of events (referred to as the yield from hereon) for each background, where a Gaussian PDF with the yield as the variable, is



(c) $\overline{B}^0 \to D^+\pi^-$ (d) $\Lambda_b^0 \to \Lambda_c^+ K^-$ (e) $\overline{B}_s^0 \to D_s^+\pi^-$

Figure 5.4: Λ_c^+ π^- invariant mass distributions of background decays, obtained from simulated decay samples which are reconstructed and selected as $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays. The distributions are fitted with 1D kernel estimation PDFs.



(a) $\overline{B}_s^0 \to D_s^+\pi^-$ (b) $\Lambda_b^0 \to \Lambda_c^+\pi^-$ (c) $\overline{B}^0 \to D^+K^-$

Figure 5.5: D^+ π^- invariant mass distributions of background decays, obtained from simulated decay samples which are reconstructed and selected as $\overline{B}^0 \to D^+\pi^-$ decays. The distributions are fitted with 1D kernel estimation PDFs.

added to the $H_b^0 \to H_c^+\pi^-$ fit. The mean and width of the Gaussian is set to the expected yield and its uncertainty, both of which are calculated using data-driven techniques. This Gaussian constraint introduces a penalty term to the fit, which reduces the likelihood value of the fit when the yield deviates from its expected value. The calculation of the expected yield differs for each background.

The signal modes of each channel are also backgrounds in the other channel, i.e. $\overline{B}^0 \to D^+\pi^-$ is a background to $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and vice-versa. To address this, the $\overline{B}^0 \to D^+\pi^-$ yield is measured first, where the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ background yield is left free in the fit. The $\overline{B}^0 \to D^+\pi^-$ yield is then used to estimate and constrain the background $\overline{B}^0 \to D^+\pi^-$ yield in the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ fit. This approach is possible because the Λ_b^0 is not a significant background in the $\overline{B}^0 \to D^+\pi^-$ fit, thus avoiding having to use complex techniques such as iteration or matrix solutions for these backgrounds.

$\overline{B}^0 \to D^+\pi^-$ misidentified as $\Lambda_b^0 \to \Lambda_c^+\pi^-$

For this background a pion from the D^+ is misidentified as a proton. Its expected yield is calculated from the selected $\overline{B}^0 \to D^+\pi^-$ yield, obtained from the LHCb f_s/f_d analysis [18]. The $\overline{B}^0 \to D^+\pi^-$ yield is then divided by the D^+ PID and mass window selection efficiency, to obtain the yield before the application of these selections. The final expected background yield is then obtained by multiplying the yield before D^+ selections with the efficiency for applying the Λ_c^+ mass window PID selection to D^+ candidates reconstructed under the Λ_c^+ mass hypothesis. The mass window and PID selection efficiencies are calculated using the simulated $\overline{B}^0 \to D^+\pi^-$ event sample. This is the dominant misidentified background to the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay, as the $\overline{B}^0 \to D^+\pi^-$ branching fraction of $(2.68 \pm 0.13) \times 10^{-3}$ is the largest of the three backgrounds decays.

$\overline{B}_s^0 \to D_s^+\pi^-$ misidentified as $\Lambda_b^0 \to \Lambda_c^+\pi^-$

In this background the K^+ from the D^+ is misidentified as a proton. The expected yield is calculated with the same technique used to obtain the misidentified $\overline{B}^0 \to D^+\pi^-$ yield in Sect. 5.6.2. The $\overline{B}_s^0 \to D_s^+\pi^-$ yields from the 2011 f_s/f_d analysis are corrected for the D_s^+ selection efficiency and then multiplied by the Λ_c^+ selection efficiency for D_s^+ candidates reconstructed under the Λ_c^+ mass hypothesis.

$\Lambda_b^0 \to \Lambda_c^+K^-$ misidentified as $\Lambda_b^0 \to \Lambda_c^+\pi^-$

The decay $\Lambda_b^0 \to \Lambda_c^+K^-$ can be misidentified as $\Lambda_b^0 \to \Lambda_c^+\pi^-$ if the bachelor kaon from the Λ_b^0 decay is misidentified as a pion. The expected yield of this background, $N_{\Lambda_b^0\to\Lambda_c^+K^-}$, is calculated from the fitted $\Lambda_b^0 \to \Lambda_c^+\pi^-$ signal yield, $N_{\Lambda_b^0\to\Lambda_c^+\pi^-}$, as:

$$N_{\Lambda_b^0\to\Lambda_c^+K^-} = \frac{N_{\Lambda_b^0\to\Lambda_c^+\pi^-}}{\epsilon_{\pi\to\pi}} \times \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+K^-)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+\pi^-)} \times \epsilon_{K\to\pi}, \tag{5.4}$$

where $\epsilon_{\pi\to\pi}(\epsilon_{K\to\pi})$ is the efficiency of the bachelor PID selection for $\Lambda_b^0 \to \Lambda_c^+\pi^-$ ($\Lambda_b^0 \to \Lambda_c^+K^-$) decays. The ratio of the $\Lambda_b^0 \to \Lambda_c^+K^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ branching fractions has been calculated to be $(7.3 \pm 0.2)\%$ in [63].

$\Lambda_b^0 \to \Lambda_c^+\pi^-$ misidentified as $\overline{B}^0 \to D^+\pi^-$

For this background the proton from $\Lambda_b^0 \to \Lambda_c^+\pi^-$ is misidentified as a pion. The Λ_c^+ π^- mass distribution of this background peaks above the B^0 mass (see Fig. 5.5), in a region where only combinatorial background events reside. Because of this, the yield is left free in the fit, as this background can be unambiguously fitted in the data.

$\overline{B}_s^0 \to D_s^+\pi^-$ misidentified as $\overline{B}^0 \to D^+\pi^-$

The expected yield of this background is calculated by taking the $\overline{B}_s^0 \to D_s^+\pi^-$ signal yields from the 2011 f_s/f_d analysis, correcting for the D_s^+ selection efficiency and then multiplying by the D^+ selection efficiency for D_s^+ candidates reconstructed under the D^+ mass hypothesis.

$\overline{B}^0 \to D^+K^-$ misidentified as $\overline{B}^0 \to D^+\pi^-$

The expected $\overline{B}^0 \to D^+K^-$ yield, $N_{\overline{B}^0\to D^+K^-}$, is calculated from the fitted $\overline{B}^0 \to D^+\pi^-$ yield, $N_{\overline{B}^0\to D^+\pi^-}$, via:

$$N_{\overline{B}^0\to D^+K^-} = \frac{N_{\overline{B}^0\to D^+\pi^-}}{\epsilon_{\pi\to\pi}} \times \frac{\mathcal{B}(\overline{B}^0 \to D^+K^-)}{\mathcal{B}(\overline{B}^0 \to D^+\pi^-)} \times \epsilon_{K\to\pi}, \tag{5.5}$$

where $\epsilon_{\pi\to\pi}(\epsilon_{K\to\pi})$ is the efficiency of the bachelor PID selection on $\overline{B}^0 \to D^+\pi^-$ ($\overline{B}^0 \to D^+K^-$) decays. The ratio between the $\overline{B}^0 \to D^+K^-$ and $\overline{B}^0 \to D^+\pi^-$ branching fractions has been measured to be $(8.2 \pm 0.3)\%$ [18].

5.6.3 Backgrounds from partially reconstructed decays

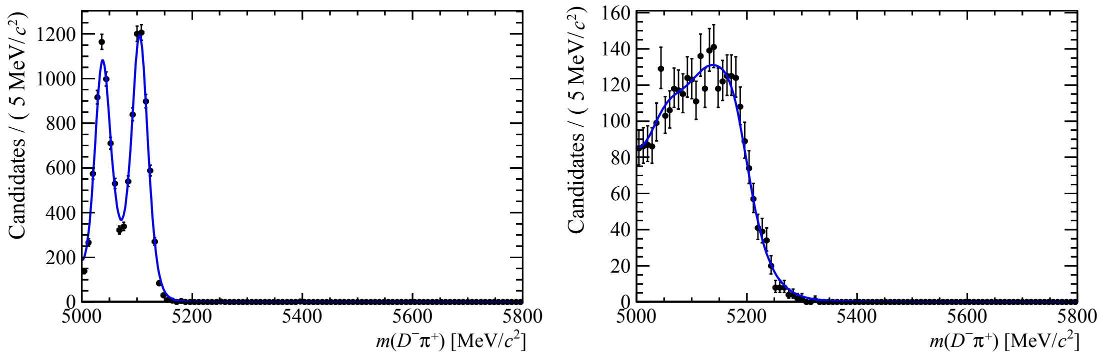
H_b^0 decays that have more than four final-state particles can be reconstructed as $H_b^0 \to H_c^+\pi^-$ decays, where the additional particles are not reconstructed. These backgrounds have a negligible influence on the $H_b^0 \to H_c^+\pi^-$ signal yield because their invariant $H_c^+\pi^-$ mass spectra reside far below the $H_b^0 \to H_c^+\pi^-$ signal mass shapes, due to the missing four-momenta of the unreconstructed final-state particles.

$\Lambda_b^0 \to \Lambda_c^+\pi^-$ backgrounds

There are multiple potential sources of partially reconstructed backgrounds to $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decays in the $\Lambda_c^+\ \pi^-$ invariant mass range, such as $\Lambda_b^0 \to \Sigma_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\rho^-$. The Σ_c^+ decays to $\Lambda_c^+\pi^0$ and the ρ to $\pi^-\pi^0$. For both modes the π^0 is not reconstructed. The PDF used to fit these backgrounds is a bifurcated Gaussian: a Gaussian with different widths above and below the mean, denoted as σ_L and σ_R. For the binned fits, the parameters of the bifurcated Gaussian are fixed to those obtained from the integrated fit. Attempts were made to fit the background with PDF's constructed from simulated $\Lambda_b^0 \to \Sigma_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\rho^-$ decay samples, but the bifurcated shape was found to provide a better overall fit. This may be due to the presence of additional partially reconstructed backgrounds, such as the Λ_b^0 decaying to π^- and higher resonant Λ_c^+ and Σ_c^+ states.

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

The dominant partially reconstructed backgrounds to the $\overline{B}^0 \to D^+\pi^-$ decay are $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0 \to D^+\rho^-$, where $D^{*+} \to D^+\ \pi^0$ and $\rho^- \to \pi^-\ \pi^0$. The π^0 is not reconstructed for either mode. 1D kernel estimation [19] PDF's are constructed for these backgrounds using simulated samples of both decays, reconstructed as $\overline{B}^0 \to D^+\pi^-$ and with the full $\overline{B}^0 \to D^+\pi^-$ selection applied. The resulting background mass distributions and their fitted PDF's are shown in Fig. 5.6. The mass distribution of the $\overline{B}^0 \to D^{*+}\pi^-$ has a double peak structure, which is a result of the π^0 produced by the D^{*+} decay be-



(a) $\overline{B}^0 \to D^{*+}\pi^-$ (b) $\overline{B}^0 \to D^+\rho^-$

Figure 5.6: D^+ π^- invariant mass distributions of the $\overline{B}^0 \to D^{*+}\pi^-$ and $\overline{B}^0 \to D^+\rho^-$ decays. These are obtained from simulated decay samples which are reconstructed and selected as $\overline{B}^0 \to D^+\pi^-$ decays. The distributions are fitted with 1D kernel estimation PDF's.

ing preferentially emitted either parallel or antiparallel to the D^{*+} flight direction. This results from the conservation of helicity in the decay.

5.6.4 Combinatorial background

The combinatorial background consists of events where a H_b^0 candidate decay is constructed from particles that did not originate from a single H_b^0 decay. For example, real H_c^+ decays which are combined with a pion from elsewhere in the event to reconstruct H_b^0 decay candidates. The invariant mass distribution of the combinatorial background is modeled with an exponential PDF. In the integrated fit the exponential coefficient $\alpha_{\rm exp}$ is left free. For the binned fits $\alpha_{\rm exp}$ is fixed to the value obtained from the integrated fit.

Table 5.15: Summary of the PDF functions and their parameterisations for each signal and background component in the $H_b^0 \to H_c^+\pi^-$ mass fit. The PDF's obtained from simulated events are 1D kernel estimation PDF's [19].

Component	PDF function	PDF parameterisation
$\Lambda_b^0 \to \Lambda_c^+\pi^-$ fit		
$\Lambda_b^0 \to \Lambda_c^+\pi^-$	Double Crystal Ball	α, n fixed from simulation, $\bar{m}, \sigma$ and yield left free.
$\overline{B}^0 \to D^+\pi^-$	Obtained from simulated events	Yield constrained to expected value.
$\overline{B}_s^0 \to D_s^+\pi^-$	Obtained from simulated events	Yield constrained to expected value.
$\Lambda_b^0 \to \Lambda_c^+K^-$	Obtained from simulated events	Yield constrained to expected value.
Partially reconstructed backgrounds	Bifurcated Gaussian	Mean and widths in binned fits fixed to values obtained from the integrated fit. Yield left free.
Combinatorial background	Exponential	Exponential coefficient fixed to value obtained from the integrated fit. Yield left free.
$\overline{B}^0 \to D^+\pi^-$ fit		
$\overline{B}^0 \to D^+\pi^-$	Double Crystal Ball	α, n fixed from simulation, $\bar{m}, \sigma$ and yield left free.
$\Lambda_b^0 \to \Lambda_c^+\pi^-$	Obtained from simulated events	Yield left free.
$\overline{B}_s^0 \to D_s^+\pi^-$	Obtained from simulated events	Yield constrained to expected value.
$\overline{B}^0 \to D^+K^-$	Obtained from simulated events	Yield constrained to expected value.
$\overline{B}^0 \to D^{*+}\pi^-$	Obtained from simulated events	Yield left free.
$\overline{B}^0 \to D^+\rho^-$	Obtained from simulated events	Yield left free.
Combinatorial background	Exponential	Exponential coefficient fixed to value obtained from the integrated fit. Yield left free.

5.6.5 Summary of the fit model

Table 5.15 summarises the treatment of each signal and background component in the $H_b^0 \to H_c^+\pi^-$ mass fit.

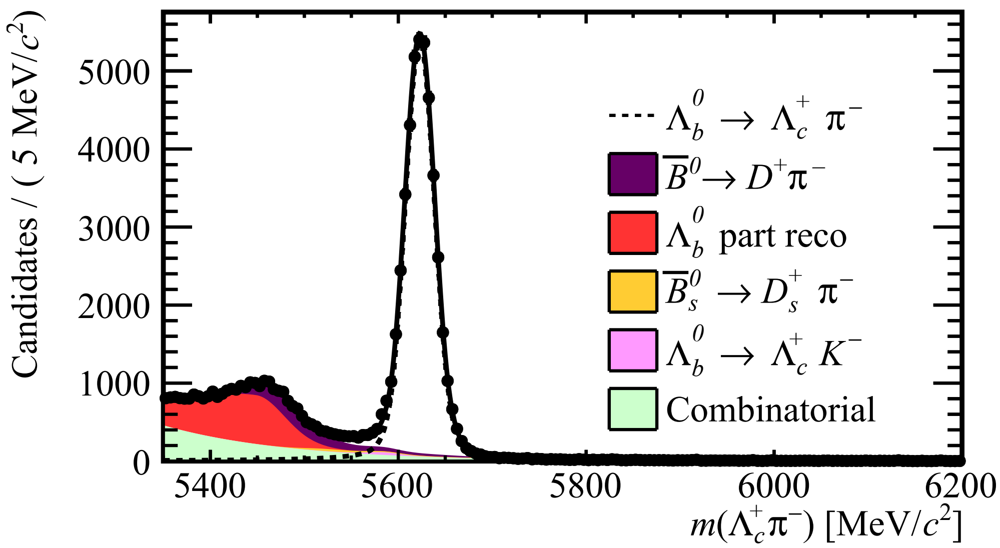


Figure 5.7: The fitted invariant $\Lambda_c^-\ \pi^+$ mass distribution of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in the full 2011 dataset. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labeled in the legend. The 'Λ_b^0 part reco' component refers to the partially reconstructed background.

5.7 Fit Results

The results of the integrated and binned $H_b^0 \to H_c^+\pi^-$ fits to the 2011 data are shown in Sect. 5.7.1, Sect. 5.7.2 and Appendix D.1. The integrated fit is used to check that the signal and background PDF's provide a good fit to the data and to obtain values for the combinatorial and partially reconstructed background PDF parameters, which are fixed in the binned fits. The binned fits are used to extract the dependence of the $H_b^0 \to H_c^+\pi^-$ yields on p_{T} and η.

5.7.1 Integrated fit

The fitted $H_c^+\pi^-$ invariant mass distributions of selected $H_b^0 \to H_c^+\pi^-$ decay candidates, integrated over all η, p_{T} bins, are shown in Fig. 5.7 and Fig. 5.8. The fitted values of the free parameters of the $H_b^0 \to H_c^+\pi^-$ fits are shown in Table 5.16 and Table 5.17. The fit model PDF's, described in Sect. 5.6, provide a good fit to the data across the entire mass range for both signal channels. The backgrounds from misidentified decays are more prominent for the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ mode than for $\overline{B}^0 \to D^+\pi^-$. This is a result of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ PID selection being looser than for $\overline{B}^0 \to D^+\pi^-$, which increases the rate at which misidentification backgrounds are selected in the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ mode.

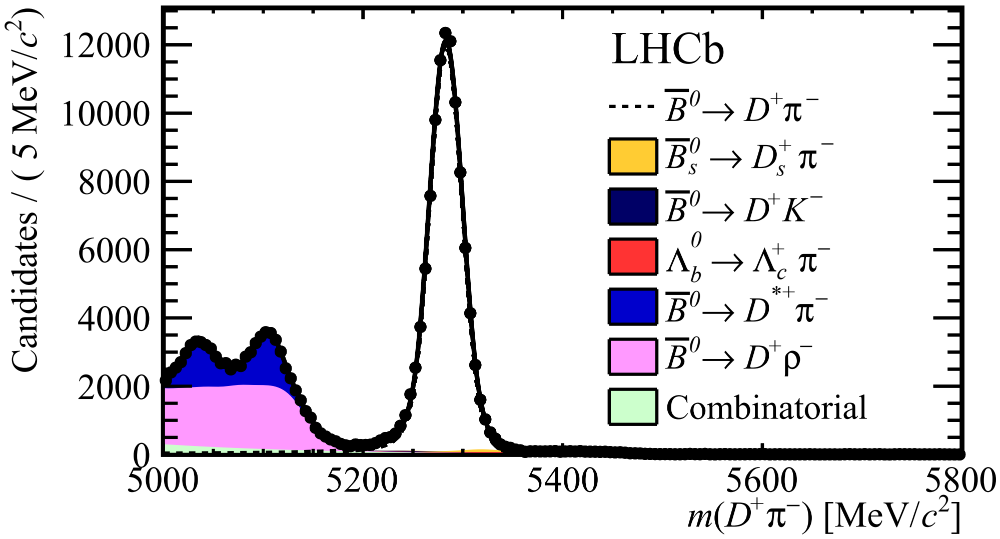


Figure 5.8: The fitted D^+ π^- invariant mass distribution of selected $\overline{B}^0 \to D^+\pi^-$ events in the full 2011 dataset. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labeled in the legend. Taken from [18].

5.7.2 Binned fit results

The fitted $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ yields in bins of p_T and η are shown in Table 5.18 and Table 5.19. The $\overline{B}^0 \to D^+\pi^-$ yield is consistently higher than the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ yield across all bins. Plots of the binned $\Lambda_b^0 \to \Lambda_c^+\pi^-$ fits are shown in Appendix D.1.

Table 5.16: Fitted values of the free parameters of the $\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$ PDF. The errors shown are statistical.

Fit Parameter	Value
Yields	
$\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$	44852 ± 228
$\overline{B}^0 \rightarrow D^+\pi^-$	5086 ± 159
$\overline{B}_s^0 \rightarrow D_s^+\pi^-$	663 ± 29
$\Lambda_b^0 \rightarrow \Lambda_c^+K^-$	431 ± 13
Partially reconstructed backgrounds	13395 ± 768
Combinatorial background	14517 ± 885
PDF shape parameters	
$\bar{m}$	5623.1 ± 0.1 MeV/c^2
σ	15.8 ± 0.1 MeV/c^2
$\alpha_{\rm exp}$	-0.0081 ± 0.0003
$\bar{m}_{partreco}$	5455 ± 2 MeV/c^2
σ_L	69.9 ± 7.8 MeV/c^2
σ_R	26.5 ± 1.7 MeV/c^2

Table 5.17: Fitted values of the free parameters of the $\overline{B}^0 \rightarrow D^+\pi^-$ PDF. The errors shown are statistical. Taken from [18].

Fit Parameter	Value
Yields	
$\overline{B}^0 \rightarrow D^+\pi^-$	106197 ± 344
$\overline{B}_s^0 \rightarrow D_s^+\pi^-$	977 ± 22
$\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$	2063 ± 83
$\overline{B}^0 \rightarrow D^+K^-$	288 ± 5
$\overline{B}^0 \rightarrow D^+\rho^-$	52617 ± 863
$B^0 \rightarrow D^{*-}\pi^+$	24018 ± 600
Combinatorial background	9539 ± 591
PDF shape parameters	
$\bar{m}$	5283.0 ± 0.1 MeV/c^2
σ	16.7 ± 0.1 MeV/c^2
$\alpha_{\rm exp}$	-0.0063 ± 0.0003

Table 5.18: The fitted yields of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ channels in bins of $p_{\rm T}$ (H_b^0). The $\overline{B}^0 \to D^+\pi^-$ yields are taken from [18].

$p_{\rm T}(H_b^0)$ (GeV/c)	$\Lambda_b^0 \to \Lambda_c^+\pi^-$	$\overline{B}^0 \to D^+\pi^-$
1.5 - 3.6	2562 ± 52	4333 ± 68
3.6 - 4.5	2618 ± 53	4763 ± 71
4.5 - 5.3	2805 ± 54	5356 ± 75
5.3 - 6.0	2903 ± 56	5422 ± 75
6.0 - 6.5	2372 ± 50	4648 ± 70
6.5 - 7.0	2394 ± 50	4899 ± 72
7.0 - 7.5	2400 ± 51	5069 ± 74
7.5 - 8.0	2399 ± 50	5041 ± 73
8.0 - 8.5	2288 ± 49	5263 ± 75
8.5 - 9.0	2299 ± 49	5221 ± 75
9.0 - 9.5	2240 ± 49	5175 ± 75
9.5 - 10.0	2006 ± 46	4884 ± 73
10.0 - 10.7	2527 ± 52	6503 ± 83
10.7 - 11.5	2555 ± 53	6419 ± 84
11.5 - 12.2	1917 ± 46	5124 ± 74
12.2 - 13.0	1848 ± 46	5054 ± 74
13.0 - 14.3	2329 ± 51	6941 ± 87
14.3 - 16.0	1998 ± 48	6563 ± 85
16.0 - 20.2	2417 ± 54	8621 ± 98
20.2 - 40.0	1454 ± 43	5227 ± 77

Table 5.19: The fitted yields of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\overline{B}^0 \to D^+\pi^-$ channels in bins of $\eta(H_b^0)$. The $\overline{B}^0 \to D^+\pi^-$ yields are taken from [18].

$\eta(H_b^0)$	$\Lambda_b^0 \to \Lambda_c^+\pi^-$	$\overline{B}^0 \to D^+\pi^-$
2.00 - 2.60	4262 ± 70	12182 ± 113
2.60 - 2.75	3117 ± 59	8212 ± 92
2.75 - 2.90	3823 ± 65	9748 ± 102
2.90 - 3.05	4357 ± 69	10746 ± 107
3.05 - 3.20	4615 ± 71	11277 ± 109
3.20 - 3.35	4720 ± 72	11327 ± 110
3.35 - 3.50	4542 ± 70	10554 ± 107
3.50 - 3.65	4046 ± 66	9337 ± 101
3.65 - 4.00	7255 ± 89	15952 ± 133
4.00 - 5.00	5483 ± 77	11121 ± 114

5.8 Evaluation of uncertainties

Various statistical and systematic uncertainties are associated with the measurement of the efficiency corrected yield ratio and the subsequent fit to $\mathcal{R}(p_{\mathrm{T}}, \eta)$. Statistical uncertainties arise from the fit to the data and the limited size of the simulation samples. Systematic uncertainties result from the configuration of the fit model, the methods used to calculate the efficiencies and potential differences between data and simulation. Sect. 5.8.1 describes each uncertainty and how it is evaluated. Sect. 5.8.3 evaluates the effects of the uncertainties on the fit to $\mathcal{R}(p_{\mathrm{T}}, \eta)$.

5.8.1 Description of uncertainties

Tail parameters of the DCB PDF

The α_1, α_2, n_1 and n_2 parameters of the DCB PDF (5.3) used to fit the $H_b^0 \to H_c^+\pi^-$ signal invariant mass distributions, referred to as the 'tail parameters', are obtained from simulated events. If the tail parameters are left free in the integrated fits to the data, their values are seen to vary by up to 10%. These variations are at least partly caused by the signal PDF fitting the background events above and below the $H_b^0 \to H_c^+\pi^-$ mass peak. As such they are interpreted as conservative limits on how much the tail parameters can vary between data and simulated events. The effect of the tail parameters on the signal yield ratio is evaluated by varying each parameter by $\pm 10\%$ simultaneously for both $H_b^0 \to H_c^+\pi^-$ fits.

σ parameter of the DCB PDF

In the default signal fit model the widths of the two CB functions in Eq. (5.3) are set to be equal. The systematic effect of this choice is evaluated by letting the σ of each CB to vary independently in the fit for both $H_b^0 \to H_c^+\pi^-$ fits. This fit model is referred to as 'DCB2Sigmas' from hereon. For the binned fits the ratio between the two widths is fixed to that obtained from the integrated fit, as the fit was found to be unstable in some bins if both widths were allowed to be free.

Exponential background coefficient

In the binned fit the exponential coefficient is fixed to that obtained from the integrated fit. The assumption being that there is no bin dependent variation in the combinatorial background shape. This assumption is verified by allowing the background coefficient to vary in both $H_b^0 \to H_c^+\pi^-$ binned fits. This fit configuration is referred to as 'comb-free'.

Partially reconstructed background in $\Lambda_b^0 \to \Lambda_c^+\pi^-$

The effect on the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ yield of the choice of fit model for the partially reconstructed background shapes in the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ fit is evaluated by fitting these backgrounds with PDF's constructed from simulated $\Lambda_b^0 \to \Sigma_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\rho^-$ event samples, instead of the bifurcated Gaussian PDF used in the default fit. This fit configuration is referred to as 'simulated part-reco in $\Lambda_b^0 \to \Lambda_c^+\pi^-$'.

$\overline{B}^0 \to D^{*+}\pi^-$ PDF in the $\overline{B}^0 \to D^+\pi^-$ fit

The $\overline{B}^0 \to D^{*+}\pi^-$ background PDF is obtained from simulated events. An alternative PDF model for this background was constructed, consisting of two DCB functions of equal width and different means. The change in $\overline{B}^0 \to D^+\pi^-$ yield is evaluated when the fit is performed using this model. This systematic is referred to as 'Alternative $\overline{B}^0 \to D^{*+}\pi^-$ PDF'.

PID efficiency

There are two main sources of systematic uncertainty associated with the reweighting method used to calculate the PID efficiency. The first is the choice of bin boundaries and widths. Varying these will result in a change in the assigned PID efficiencies in each bin. Candidates most affected by this uncertainty are those which lie close to the kinematic bin boundaries and those where the rate of change of efficiency with the kinematic variable is large across a single bin. The second source of uncertainty is the choice of variables which are reweighted. If a variable has an effect on the PID performance, but is not binned in the PID efficiency tables, then the calculated PID efficiency would be incorrect if the signal channel and PID calibration sample have different distributions of this variable. Both sources of uncertainty are evaluated by calculating the relative difference between the 'true' PID efficiency and that obtained by applying the reweighting method for simulated signal and PID calibration event samples. The true efficiency is obtained by applying the PID selection directly in the simulation. This uncertainty is subject to statistical fluctuations due to the limited size of the simulated signal and PID calibration samples from which it is calculated.

L0Hadron trigger efficiency

The efficiency of the hardware level hadronic trigger, encoded in the 'L0Hadron' line, has been seen to differ by up to 2% for kaons and pions with similar kinematic distributions in data [18]. The cause of this difference is not presently understood. In the simulation their efficiencies are identical. The proton trigger response has not been studied extensively due to the limited kinematic range of the high purity proton calibration sample

(Appendix B.3). The proton L0Hadron efficiency is conservatively estimated to differ from the pion efficiency by no more than 5%. The final states of the two $H_b^0 \to H_c^+\pi^-$ decay modes differ by one mode having a proton and the other a pion instead. Any potential systematic effect caused by different L0Hadron efficiencies for protons and pions is quantified by multiplying the fraction of events where the proton exclusively triggers the L0Hadron line by $\pm5\%$. The resulting value is the maximum change in $\mathcal{R}$ that can be caused by this systematic.

BDT Selection

Systematic uncertainties can enter into $\mathcal{R}$ if there are differences between the data and the simulation for the selection variables that do not cancel in the ratio. The BDT output variable is especially sensitive to these differences because the output is determined from multiple kinematic variables and their correlations. The BDT systematic uncertainty is evaluated by calculating the change in $\mathcal{R}$ when the BDT selection is varied from $BDT_{\text{out}} > 0.66$ to $BDT_{\text{out}} > 0.2$ and $BDT_{\text{out}} > 0.8$. These cut values are chosen because they result in the maximum change in the selection efficiency (up to $\pm25\%$) whilst retaining a signal significance (described in Sect. 5.4.2) within $\pm5\%$ of the default selection. This ensures that the precision of the signal fit is comparable to the default fit. This uncertainty is subject to statistical fluctuations caused by adding or removing events from the data and simulation samples when the BDT selection is tightened or loosened.

Statistical uncertainties

Statistical uncertainties are associated with both yield and efficiency components of $\mathcal{R}$ due to the limited size of the data and simulated samples of the $H_b^0 \to H_c^+\pi^-$ decays in each bin. The data yield uncertainties are obtained from the fit results, shown in Table 5.18 and Table 5.19. The efficiency uncertainties are taken as the binomial errors of the combined efficiencies, shown in Table 5.12 and Table 5.13.

Bin migration

A potential uncertainty may arise from the p_{T} and η resolution smearing the p_{T} and η distributions of the H_b^0, resulting in the migration of events across p_{T}/η bins for events with which have H_b^0 p_{T} or η close to the p_{T}/η bin boundaries. The H_b^0 p_{T}/η resolution is estimated to be the same as the mass resolution, $\Delta m/m = 0.3\%$ (from Table 5.16), because the precision of p_{T},η and m are all derived from the precision of measuring the four-momentum vectors of the final state particles. Given this resolution, the number of events that are sensitive to bin migration is negligible, as a result this systematic is not evaluated in the presented analysis.

5.8.2 Binned uncertainty values

Tables 5.20 and 5.21 show the change in $\mathcal{R}$ that is caused by each evaluated uncertainty. The PID efficiency, BDT selection variation and simulation sample size provide the dominant contributions to the systematic uncertainty. The signal PDF tail parameter α and n systematics are combined into a single value, denoted as 'Tail Parameters', assigned as the largest positive and negative changes of these systematics. This is done to account for the correlations between these parameters, as varying one parameter results in a compensating change in the other tail parameters. The remaining systematic uncertainties are treated as being uncorrelated and are combined by adding the positive and negative values in quadrature for each bin. The combined uncertainties are of a similar size to the statistical uncertainties on the fitted yields. The magnitudes of the systematic uncertainties are subject to greater fluctuations across the p_{T}, eta bins than the data statistical uncertainties. This is primarily caused by the BDT and PID systematics, both of which are sensitive to statistical fluctuations in the data and simulation samples from which they are calculated.

5.8.3 Uncertainties of the fit to $\mathcal{R}(p_{\mathrm{T}}, \eta)$

The $\mathcal{R}$ distributions are fitted by functions of the form $\mathcal{R}(p_{\mathrm{T}}) = a + \exp(b + c \times p_{\mathrm{T}})$ and $\mathcal{R}(\eta) = a + b \times (\eta - 3.20)$, where a, b and c are all free fit parameters. The functions are discussed in more detail in Sect. 5.9. For each systematic uncertainty, the change in the fit parameters is calculated by simultaneously varying the value of $\mathcal{R}$ in every bin by the uncertainties shown in Table 5.20 and Table 5.21. This ensures that systematic effects that are correlated across bins are accounted for. The statistical uncertainties are obtained from the fit to $\mathcal{R}$, with the uncertainty of each bin set to the statistical error only.

As is done in Sect. 5.8.2, the signal PDF tail parameter systematic is assigned as the greatest increase (decrease) in the variable resulting from one of the tail parameter systematics. The remaining systematics are treated as being uncorrelated and are added in quadrature. The resulting uncertainties are shown in Table 5.22. The dominant uncertainties are the statistical, BDT and PID systematic uncertainties.

An additional systematic uncertainty, 'Bin centre', is evaluated in Table 5.22. This systematic assesses the change in the fit function when the p_{T},η bin centres are assigned as the mean Λ_b^0 or B^0 p_{T},η values, instead of the midpoint between the B^0 and Λ_b^0 mean values as is done in the default fit.

Interpreting the fit uncertainties

The relative uncertainties on the fit model parameters are greater than those for the binned values of $\mathcal{R}$ shown in Table 5.20 and Table 5.21. This is a result of the correlations

Table 5.20: Values of the individual and combined uncertainties on $\mathcal{R}$ in bins of p_{T} (H_b^0).

	% change in $\mathcal{R}(p_{\mathrm{T}})$ in bins of $\mathrm{p_T(H_b^0)}$ GeV/c									
	1.5 - 3.6	3.6 - 4.5	4.5 - 5.3	5.3 - 6.0	6.0 - 6.5	6.5 - 7.0	7.0 - 7.5	7.5 - 8.0	8.0 - 8.5	8.5 - 9.0
Signal PDF tail parameters										
$\alpha_1 - 10\%$	0.3	0.1	0.1	0.2	0.0	0.1	0.1	0.2	0.1	0.2
$\alpha_1 + 10\%$	-0.2	-0.1	0.0	-0.1	0.1	-0.0	-0.0	-0.1	0.0	-0.1
$\alpha_2 - 10\%$	0.0	-0.0	0.0	0.0	-0.0	0.1	0.1	-0.0	0.0	-0.0
$\alpha_2 + 10\%$	-0.0	0.0	-0.0	0.0	0.1	-0.0	-0.0	-0.0	0.0	0.0
$n_1 - 10\%$	0.1	0.0	-0.0	0.0	-0.0	-0.0	0.0	-0.0	-0.0	0.0
$n_1 + 10\%$	-0.0	0.0	0.1	0.0	0.1	0.1	0.0	-0.0	0.1	0.0
$n_2 - 10\%$	0.0	0.0	0.0	0.0	-0.0	0.0	0.0	-0.0	0.0	-0.0
$n_2 + 10\%$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.0	0.0	0.0
Uncorrelated Uncertainties										
Tail parameters (High)	0.3	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.2
Tail parameters (Low)	-0.2	-0.1	-0.0	-0.1	-0.0	-0.0	-0.0	-0.1	-0.0	-0.1
DCB2sigmas Signal PDF	-0.1	0.0	-0.1	0.1	-0.0	0.1	0.0	0.1	0.1	0.1
Combinatorial Background	-0.5	-0.4	-0.1	-0.3	-0.4	0.5	0.4	0.5	0.7	0.5
$B^0 \to D^{*-}\pi^+$ PDF in $\overline{B}^0 \to D^+\pi^-$	0.2	0.1	0.0	0.1	0.1	-0.0	0.1	-0.1	0.1	0.1
$\Lambda_b^0 \to \Sigma_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\rho^-$ in $\Lambda_b^0 \to \Lambda_c^+\pi^-$	0.1	0.1	0.2	0.0	0.0	0.1	0.1	0.0	0.1	-0.1
BDT > 0.8	6.0	3.3	-2.4	2.9	0.2	-2.0	0.1	1.9	0.4	4.2
BDT > 0.2	-3.6	3.5	-0.0	-5.0	-6.7	-1.9	0.9	1.4	-4.7	-4.2
Trigger	± 0.2	± 0.2	± 0.3	± 0.3	± 0.4	± 0.3	± 0.4	± 0.4	± 0.4	± 0.5
PID efficiency	0.7	-3.5	-2.0	1.6	-1.0	-2.3	0.8	-2.4	-2.9	0.4
Simulation sample size	± 5.3	± 5.5	± 4.9	± 4.7	± 5.2	± 4.7	± 4.6	± 5.2	± 5.1	± 4.8
Combined Systematic Uncertainty (High)	7.4	6.4	4.0	5.3	4.5	4.4	4.4	4.8	4.1	5.9
Combined Systematic Uncertainty (Low)	- 5.5	- 5.5	- 5.1	- 6.5	- 8.1	- 5.7	- 4.2	- 4.7	- 6.8	- 5.9
Data Statistical Uncertainty	± 4.2	± 4.2	± 4.0	± 4.1	± 4.5	± 4.4	± 4.2	± 4.1	± 4.0	± 4.0

	% change in $\mathcal{R}(p_{\mathrm{T}})$ in bins of $\mathrm{p_T(H_b^0)}$ GeV/c									
$\mathrm{p_T(H_b^0)}$ GeV/c	9.0 - 9.5	9.5 - 10.0	10.0 - 10.7	10.7 -11.5	11.5 - 12.2	12.2 -13.0	13.0 - 14.3	14.3 - 16.0	16.0 - 20.2	20.2 - 40.0
Signal PDF tail parameters										
$\alpha_1 - 10\%$	0.2	0.2	0.3	0.3	0.4	0.6	0.4	0.6	0.3	0.3
$\alpha_1 + 10\%$	0.0	0.1	-0.2	-0.1	-0.1	-0.4	-0.2	-0.2	-0.2	-0.1
$\alpha_2 - 10\%$	0.1	0.1	-0.0	-0.1	0.0	-0.1	-0.0	-0.0	-0.1	0.1
$\alpha_2 + 10\%$	0.0	0.0	0.0	-0.0	0.0	0.1	-0.0	0.1	-0.0	-0.1
$n_1 - 10\%$	0.0	-0.0	0.0	-0.0	0.1	0.1	0.0	0.1	-0.0	-0.1
$n_1 + 10\%$	0.1	0.1	-0.0	-0.0	-0.0	-0.1	-0.0	-0.0	-0.0	-0.0
$n_2 - 10\%$	0.0	0.0	-0.0	-0.0	0.0	-0.0	-0.0	0.0	-0.0	-0.0
$n_2 + 10\%$	0.0	0.0	-0.0	-0.0	0.0	0.0	-0.0	0.0	-0.0	-0.1
Uncorrelated Uncertainties										
Tail parameters (High)	0.2	0.2	0.3	0.3	0.4	0.6	0.4	0.6	0.3	0.3
Tail parameters (Low)	0.0	-0.0	-0.2	-0.1	-0.1	-0.4	-0.2	-0.2	-0.2	-0.1
DCB2sigmas Signal PDF	-0.0	0.1	-0.0	-0.0	0.0	-0.0	0.1	0.1	0.2	0.3
Combinatorial Background	-0.1	-0.1	0.5	0.2	0.8	-0.1	0.4	0.3	0.0	-0.0
$B^0 \to D^{*-}\pi^+$ PDF in $\overline{B}^0 \to D^+\pi^-$	-0.1	-0.1	0.1	0.0	0.1	0.0	-0.1	-0.0	-0.1	0.4
$\Lambda_b^0 \to \Sigma_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\rho^-$ in $\Lambda_b^0 \to \Lambda_c^+\pi^-$	0.3	0.2	0.5	0.3	-0.1	0.1	0.6	1.1	1.0	2.2
BDT > 0.8	0.1	-0.4	2.6	1.3	2.7	3.7	-0.5	3.1	0.1	-0.6
BDT > 0.2	1.1	-1.6	-1.7	-1.6	0.6	-1.3	-0.9	0.2	0.9	1.6
Trigger	± 0.4	± 0.4	± 0.4	± 0.4	± 0.4	± 0.4	± 0.3	± 0.3	± 0.2	± 0.1
PID efficiency	-0.6	-4.7	0.4	0.2	-4.5	-3.4	-1.8	-2.8	-0.5	-1.8
Simulation Sample Size	± 4.9	± 5.0	± 4.4	± 4.1	± 4.9	± 5.3	± 3.9	± 4.4	± 3.8	± 4.3
Combined Systematic Uncertainty (High)	4.4	4.3	4.7	3.9	5.0	5.6	3.7	5.0	3.5	4.9
Combined Systematic Uncertainty (Low)	- 4.3	- 6.5	- 4.2	- 4.0	- 6.0	- 5.5	- 4.1	- 4.6	- 3.2	- 4.4
Data Statistical Uncertainty	± 4.2	± 4.3	± 3.8	± 3.6	± 4.1	± 4.1	± 3.5	± 3.6	± 3.2	± 4.0

Table 5.21: Values of the individual and combined uncertainties on $\mathcal{R}$ in bins of $\eta(H_b^0)$.

	% change in $\mathcal{R}(\eta)$ in bins of $\eta(\mathrm{H_b^0})$ GeV/c									
	2.00 - 2.60	2.60 - 2.75	2.75 - 2.90	2.90 - 3.05	3.05 - 3.20	3.20 - 3.35	3.35 - 3.50	3.50 - 3.65	3.65 - 4.00	4.00 - 5.00
Signal PDF tail parameters										
$\alpha_1 - 10\%$	0.6	0.6	0.4	0.4	0.3	0.3	0.2	0.1	0.1	0.0
$\alpha_1 + 10\%$	-0.3	-0.3	-0.2	-0.2	-0.1	-0.1	-0.1	-0.0	0.0	0.0
$\alpha_2 - 10\%$	-0.0	0.1	0.0	0.1	0.0	-0.0	0.0	-0.0	-0.0	-0.0
$\alpha_2 + 10\%$	0.0	-0.1	-0.0	-0.1	0.0	0.0	-0.0	-0.0	0.0	0.0
$n_1 - 10\%$	0.1	0.1	0.1	0.1	0.1	0.0	0.0	-0.0	-0.0	-0.1
$n_1 + 10\%$	-0.1	-0.1	-0.1	-0.0	0.0	-0.0	0.0	0.0	0.0	0.0
$n_2 - 10\%$	-0.0	0.0	-0.0	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0
$n_2 + 10\%$	0.0	-0.0	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0
Uncorrelated Uncertainties										
Tail parameters (High)	0.6	0.6	0.4	0.4	0.3	0.3	0.2	0.1	0.1	0.0
Tail parameters (Low)	-0.3	-0.3	-0.2	-0.2	-0.1	-0.1	-0.1	-0.0	-0.0	-0.1
DCB2sigmas Signal PDF	0.2	0.1	-0.0	0.0	0.1	0.1	0.0	-0.0	0.1	-0.1
Combinatorial Background	-0.1	-0.1	0.5	0.2	0.8	-0.1	0.4	0.3	0.0	-0.0
$B^0 \to D^{*-}\pi^+$ PDF in $\overline{B}^0 \to D^+\pi^-$	-0.0	0.0	0.0	-0.0	-0.0	0.0	0.0	0.0	0.1	0.5
$\Lambda_b^0 \to \Sigma_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\rho^-$ in $\Lambda_b^0 \to \Lambda_c^+\pi^-$	0.0	0.0	-0.1	0.3	-0.0	0.0	0.3	0.3	0.2	0.1
BDT > 0.8	0.2	-1.3	-1.1	0.2	1.3	1.5	-0.4	-1.0	-1.3	2.3
BDT > 0.2	0.5	0.9	0.9	0.1	1.4	1.2	-0.8	-2.6	-1.3	0.8
Trigger	± 0.4	± 0.3	± 0.4	± 0.4	± 0.4	± 0.3	± 0.3	± 0.3	± 0.3	± 0.3
PID efficiency	-1.7	-4.6	-2.3	-3.0	-1.0	-1.5	-1.4	-1.4	-0.3	3.3
Simulation Sample Size	± 2.9	± 3.8	± 3.5	± 3.4	± 3.5	± 3.5	± 3.7	± 3.6	± 3.0	± 3.8
Combined Systematic Uncertainty (High)	2.8	3.4	3.3	3.1	3.6	3.5	3.1	3.2	2.4	5.2
Combined Systematic Uncertainty (Low)	- 3.2	- 5.8	- 4.0	- 4.3	- 3.1	- 3.3	- 3.5	- 4.4	- 3.0	- 3.0
Data Statistical Uncertainty	± 2.6	± 3.2	± 3.0	± 3.0	± 2.9	± 2.9	± 3.0	± 3.2	± 2.4	± 3.0

Table 5.22: Systematic uncertainties on the $\mathcal{R}$ fit parameters. The $H_b^0 \to H_c^+\pi^-$ signal fit parameters are correlated, from these the greatest variations in the fit parameter is taken and added in quadrature with the remaining uncorrelated systematics to get the final errors, listed at the bottom of the table. The 'Λ_b^0 part reco' component refers to the partially reconstructed Λ_b^0 background systematic.

	$p_{\rm T}$ bin fit function $\mathcal{R}(p_{\rm T}) = a + \exp(b + c \times p_{\rm T})$			η bin fit function $\mathcal{R}(\eta) = a + b \times (\eta - 3.20)$	
	a	b	c	a	b
Signal fit parameters					
$\alpha_1 - 10\%$	0.2 %	0.1 %	-0.4 %	0.3 %	-1.6 %
$\alpha_1 + 10\%$	-0.4 %	-0.2 %	-0.0 %	-0.1 %	1.0 %
$\alpha_2 - 10\%$	0.1 %	-0.1 %	0.2 %	0.0 %	-0.2 %
$\alpha_2 + 10\%$	-0.3 %	-0.1 %	-0.2 %	-0.0 %	0.1 %
$n_1 - 10\%$	-0.2 %	-0.0 %	-0.1 %	0.0 %	-0.6 %
$n_1 + 10\%$	-0.1 %	-0.2 %	0.0 %	-0.0 %	0.3 %
$n_2 - 10\%$	-0.0 %	-0.1 %	0.0 %	0.0 %	-0.0 %
$n_2 + 10\%$	-0.2 %	-0.1 %	-0.1 %	0.0 %	-0.1 %
Uncorrelated parameters					
Tail parameters (High)	0.2 %	0.1 %	0.2 %	0.3 %	1.0 %
Tail parameters (Low)	-0.5 %	-0.2 %	-0.4 %	-0.1 %	-1.6 %
DCB2Sigmas Signal PDF	0.7 %	0.5 %	0.1 %	0.1 %	-0.5 %
Combinatorial Background	-3.6 %	0.1 %	-3.1 %	0.2 %	0.2 %
$\overline{B}^0 \to D^{*+}\pi^-$ PDF in $\overline{B}^0 \to D^+\pi^-$	1.3 %	0.0 %	0.9 %	0.1 %	1.0 %
Λ_b^0 part reco	5.9 %	2.9 %	2.2 %	0.1 %	0.7 %
BDT > 0.8	2.9 %	-5.8 %	3.0 %	-0.0 %	2.1 %
BDT > 0.2	9.1 %	11.4 %	3.5 %	0.0 %	-4.7 %
Trigger	1.2 %	1.1 %	1.0 %	0.3 %	0.0 %
PID efficiency	1.2 %	1.0 %	3.1 %	-1.3 %	12.5 %
Simulation sample size	± 15.2 %	± 11.9 %	± 8.3 %	± 0.9 %	± 9.1 %
Bin centre	±0.4 %	±0.4 %	±0.1 %	±0.1 %	±1.3 %
Total Systematic Uncertainty (High)	19.2 %	15.2 %	13.5 %	1.0 %	15.6 %
Total Systematic Uncertainty (Low)	-15.7 %	-10.1 %	-12.3 %	-1.7 %	-10.6 %
Statistical Uncertainty in Data	± 13.4 %	± 5.8 %	± 8.8 %	± 0.6 %	± 5.9 %

between the fit variables and that each variable has a different effect on the output of the fitted function at different values of $p_{\rm T}$,η. As an example of the effect of the correlations, the function of $\mathcal{R}$ that results from the default $p_{\rm T}$ binned fit is:

$$\mathcal{R}(p_{\rm T})_{\rm default} = 0.16 + \exp(-0.43 - 0.09 \times p_{\rm T}(\,{\rm GeV})), \tag{5.6}$$

while the fit to the $\mathcal{R}$ values which are varied by the $BDT_{\rm out} > 0.2$ systematic is:

$$\mathcal{R}(p_{\rm T})_{BDT_{\rm out}>0.2} = 0.18 + \exp(-0.47 - 0.09 \times p_{\rm T}(\,{\rm GeV})). \tag{5.7}$$

Although the a and b variables of Eqn's (5.6) and (5.7) differ by $\mathcal{O}(10\%)$, the output values of $\mathcal{R}(p_{\rm T})$ differ by much less, i.e. $\mathcal{R}(p_{\rm T} = 0)_{BDT_{\rm out}>0.2}$ and $\mathcal{R}(p_{\rm T} = 0)_{\rm default}$ differ

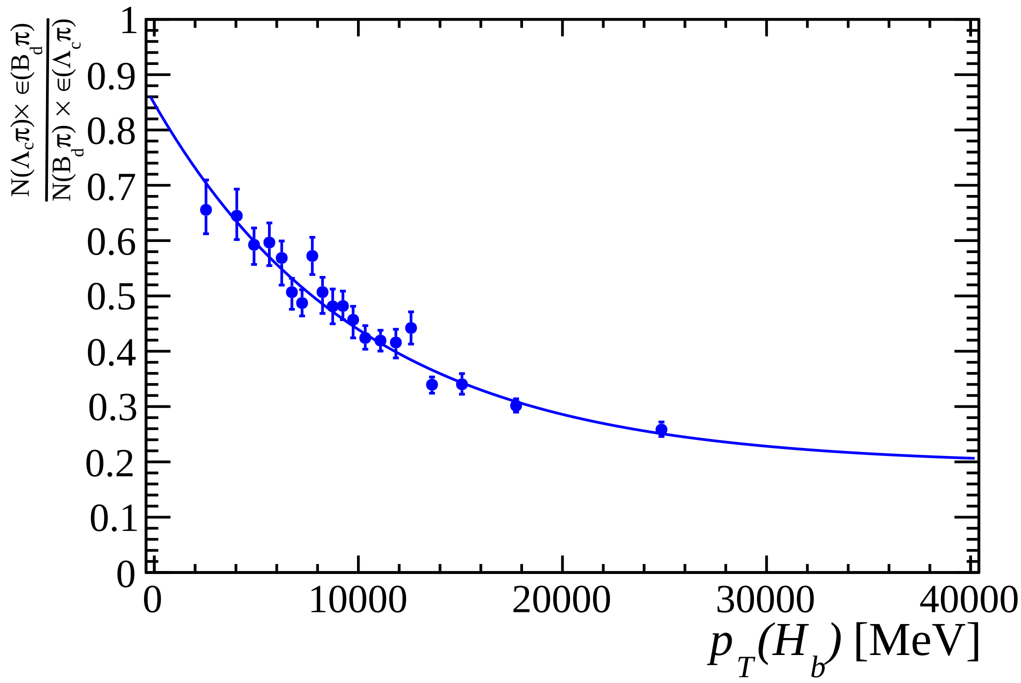


Figure 5.9: The $\mathcal{R}(p_{\mathrm{T}})$ distribution, fitted with a combined exponential and constant PDF. The $\mathcal{R}$ value error bars show the combined statistical and systematic errors of the efficiency-corrected yields. The p_{T} value errors are the standard error on the mean p_{T} (too small to be visible).

by 2.0%. At $p_{\mathrm{T}} = 10$ GeV/c the difference is 1.4%. This is a result of the change in one variable being compensated for by the change in another variable.

The $\mathcal{R}(\eta)$ fit function is parameterised to remove the correlation between the two function variables. Although uncertainties induce a variation in the b variable of up to 15%, their effect on the output of $\mathcal{R}(\eta)$ tends to be negligible, as it is much less sensitive to relative variations in the b variable than the a variable. The PID systematic results in the greatest variation in a and b in the η binned fit. The default η fit function is

$$\mathcal{R}(\eta)_{\mathrm{default}} = 0.46 + 0.08 \times (\eta - 3.20), \tag{5.8}$$

while that of the (dominant) PID systematic fit function is

$$\mathcal{R}(\eta)_{BDT_{\mathrm{out}}>0.2} = 0.46 + 0.09 \times (\eta - 3.20). \tag{5.9}$$

The difference in output between (5.8) and (5.9) is 4.9% at $\eta = 2$, 1.8% at $\eta = 3$ and 2.0% at $\eta = 5$, much smaller than the 12.5% variation of b.

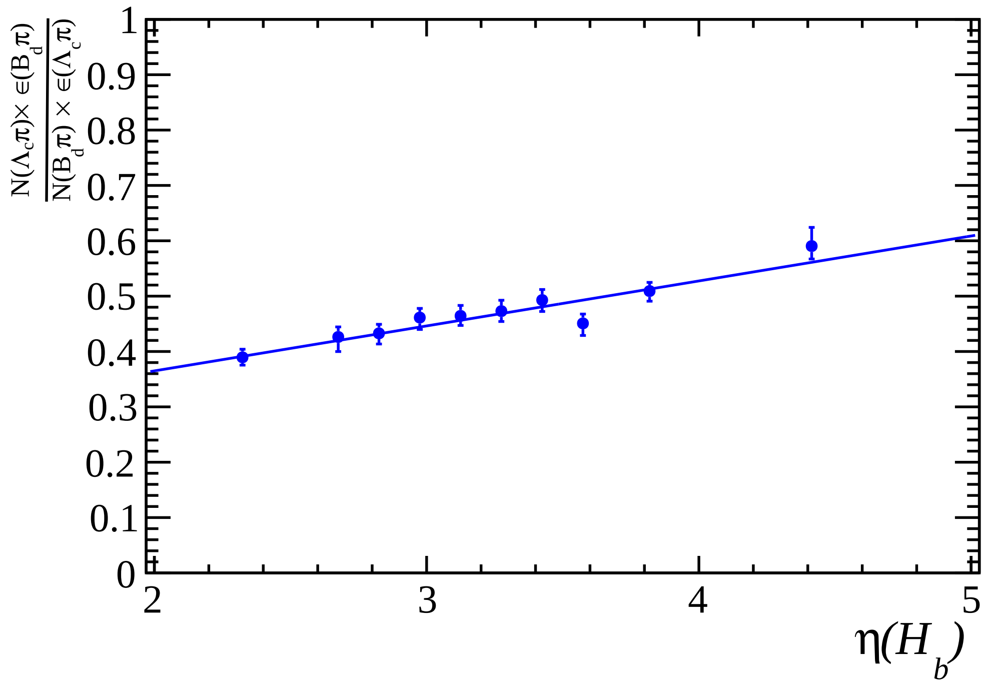


Figure 5.10: The $\mathcal{R}(\eta)$ distribution, fitted with a linear PDF. The $\mathcal{R}$ value error bars show the combined statistical and systematic errors of the efficiency-corrected yields. The η value errors are the standard error on the mean η (too small to be visible).

5.9 Results

Figures 5.9 and 5.10 show the fitted $\mathcal{R}$ distributions. The combined systematic and statistical uncertainties are shown in the plots. The fit itself is performed considering only the statistical uncertainties. The treatment of the systematic uncertainties is described in Sect. 5.8.2. The central value of each bin is evaluated as $(\overline{p_{\mathrm{T}}}, \overline{\eta}(\Lambda_b^0) + \overline{p_{\mathrm{T}}}, \overline{\eta}(B^0))/2$, where $\overline{p_{\mathrm{T}}}, \overline{\eta}(H_b^0)$ is the mean H_b^0 p_{T}, η in the bin. The p_{T}, η central value uncertainties are taken as the standard error on the mean. The systematic effect of the choice of bin centre value is assessed in Sect. 5.8.2.

The distributions are fitted with different PDF's. The p_{T} dependence is best described by an exponential function combined with a constant term, $\mathcal{R}(p_{\mathrm{T}}) = a + \exp(b + c \times p_{\mathrm{T}})$, where a, b and c are all free fit parameters. The η dependence is fitted with a linear function of the form $\mathcal{R}(\eta) = a + b \times (\eta - c)$, where a and b are free fit parameters. The c term performs a linear translation to the η variable. It is assigned a value that results in the removal of the correlations between a and b, allowing for a clearer interpretation of the uncertainties.

The results of both fits are

$$\mathcal{R}(p_{\mathrm{T}}) = (0.17 \pm 0.02^{+0.03}_{-0.03}) + \exp\left\{(-0.44 \pm 0.04^{+0.05}_{-0.07}) + (-0.09 \pm 0.01^{+0.01}_{-0.01}) \times p_{\mathrm{T}}(\,\mathrm{GeV})\right\}, \tag{5.10}$$

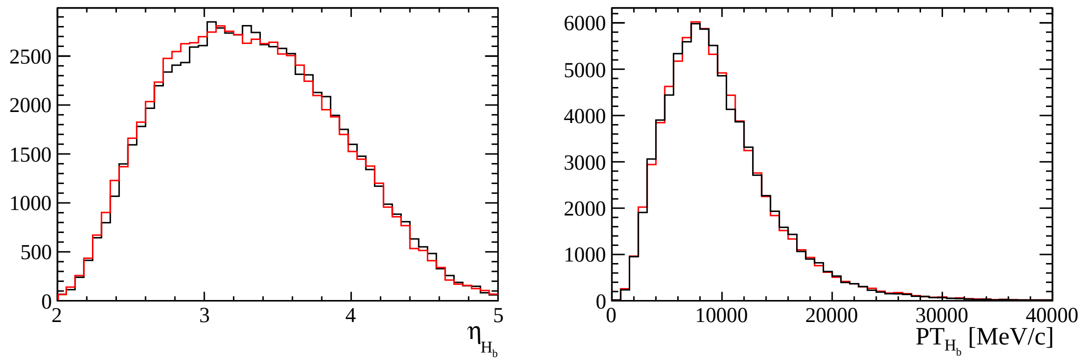


Figure 5.11: The $p_{\rm T}$ and η distributions of the Λ_b^0 (black) and B^0 (red, gray in B & W) for simulated $\overline{B}^0 \to D^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events.

and

$$\mathcal{R}(\eta) = (0.46 \pm 0.01^{+0.01}_{-0.01}) + (0.08 \pm 0.01^{+0.01}_{-0.01}) \times (\eta - 3.20), \tag{5.11}$$

where the stated uncertainties are of the form $\pm(\text{statistical})^{+\text{upper systematic}}_{-\text{lower systematic}}$. Both functions provide good fits to the data, the χ^2/DOF value of the $p_{\rm T}(\eta)$ fit is 16.40/17 (7.63/8), which correspond to a fit p-value of 0.50 (0.47).

The simulation does not predict any $p_{\rm T}$ or η dependence for $f_{\Lambda_b^0}/f_d$, as shown in Fig. 5.11 where the $p_{\rm T}$ and η distributions of the B^0 and Λ_b^0 are identical.

5.10 Conclusions

The $p_{\rm T}$ dependence of $f_{\Lambda_b^0}/f_d$ is exponential with a plateau at high $p_{\rm T}$, suggesting that the probability for a b quark hadronising to an Λ_b^0 baryon is greater than zero across the entire spectrum of the b quark's $p_{\rm T}$. This contrasts with the purely exponential dependence used in the HFAG review [17], where $f_{\Lambda_b^0}/f_d \to 0$ as $p_{\rm T} \to \infty$. A first measurement of the η dependence of $f_{\Lambda_b^0}/f_d$ is also performed, the dependence is seen to be linear. These conclusions can aid the development of QCD models describing b quark hadronisation [64] and the Pythia [48] simulation framework.

The scope of the presented analysis can be extended by measuring the correlations between the $p_{\rm T}$ and η dependencies by binning in both $p_{\rm T}$ and η, and also incorporating the 2 fb^{-1} 2012 dataset to improve statistical precision. This could be done using either the hadronic $H_b^0 \to H_c^+\pi^-$ or the semileptonic $H_b^0 \to H_c^+\,\mu^-\,\nu_\mu$ decay modes.

Chapter 6

Search for the rare decay $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$

This chapter describes the search for the decays $B^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-\mu^+\mu^-$ using LHCb data collected in 2011 with $\sqrt{s} = 7\,\text{TeV}$, corresponding to an integrated luminosity of 1.0 fb^{-1}. The preliminary results of this search were presented at the XLVIIth Recontres de Moriond session devoted to QCD and High Energy Interactions, La Thuile, 10-17 March 2012 [65]. The final results were published in the Physical Review Letters journal in May 2013 [39]. The author contributed to all parts of the analysis, except the selection development and the $B^0 \to J/\psi K^{*0}$ S-wave analysis. These are detailed for reference and completeness only.

6.1 Introduction

The non-resonant variants of the decays $B^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-\mu^+\mu^-$ have yet to be observed by experiment. They are FCNC processes which are heavily suppressed in the SM. Any observed enhancement in their branching fractions would be indicative of physics beyond the SM. The dominant SM decay mechanism of $B^0_s \to \mu^+\mu^-\mu^+\mu^-$ is mediated via the J/ψ and $\phi(1020)$ resonances, as shown in Fig. 6.1(a). The branching fraction of this decay channel, referred to as the 'resonant' channel, is calculated as the product of the $B^0_s \to J/\psi\phi$, $J/\psi \to \mu^+\mu^-$ and $\phi \to \mu^+\mu^-$ branching fractions [1], resulting in a value of $\mathcal{B}(B^0_s \to J/\psi\,(\to \mu^+\mu^-)\,\phi\,(\to \mu^+\mu^-)) = (2.3 \pm 0.9) \times 10^{-8}$. The main non-resonant SM decay channel is $B^0_{(s)} \to \mu^+\mu^-\gamma\,(\to \mu^+\mu^-)$, shown in Fig. 6.1(b), where one opposite-sign muon pair is produced via an electroweak loop diagram and the other pair is produced by a virtual photon. The branching fraction for this channel is expected to be less than 10^{-10} [66]. The resonant channel is excluded from the search for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$.

The decay rates of $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ can be enhanced by new physics processes, some

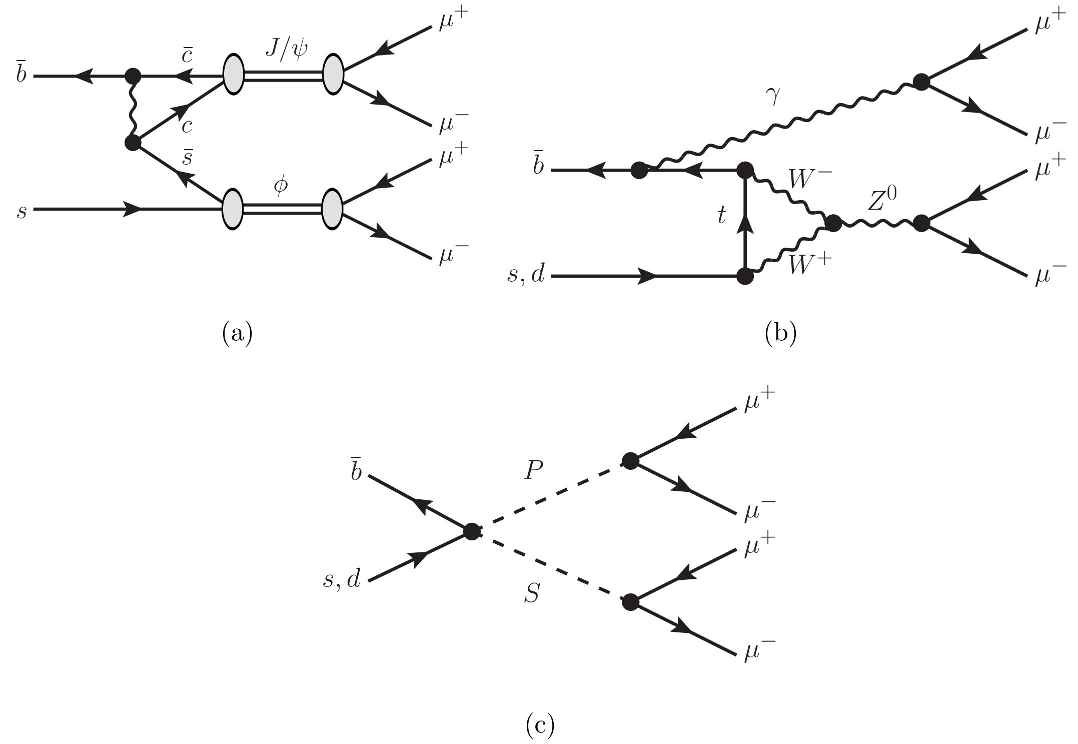


Figure 6.1: Feynman diagrams for the $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to \mu^+\mu^-\mu^+\mu^-$ decays. The resonant $B_s^0 \to J/\psi\phi$ SM channel (a), the non-resonant SM channel (b) and the sgoldstino mediated supersymmetric channel (c) are shown.

of which can involve novel couplings that have yet to be probed by experiment. One such process is the supersymmetric $B_{(s)}^0 \to S\,(\to \mu^+\mu^-)\,P\,(\to \mu^+\mu^-)$[1] decay channel shown in Fig. 6.1(c), where the decay is mediated by scalar S and pseudoscalar P sgoldstinos via the type I and II couplings described in Sect. 2.6.1. The phenomenology of this decay model is detailed in [67, 68]. The 214.3 MeV/c^2 resonance hinted at by the HyperCP collaboration [34] can be interpreted as the P sgoldstino in the $B_{(s)}^0 \to SP$ decay.

6.1.1 Analysis strategy

The search for $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$, referred to as the 'signal channel' is performed by selecting candidate $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ decay events in the 2011 dataset. The dataset is described in Sect. 6.2 and the selection process is described in Sect. 6.4. The number of selected signal $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ candidates with a four-muon invariant mass close to the B_s^0 and B^0 masses is counted and compared with background expectations. The $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ branching fractions are measured relative to that of the well measured decay $B^0 \to J/\psi\,(\to \mu^+\mu^-)\,K^{*0}\,(\to K^+\pi^-)$, referred to as the 'normalisation channel'. The normalisation process is detailed in Sect. 6.6. The relative differences between the reconstruction and selection efficiencies of the signal and normalisation channels are cor-

[1] abbreviated as $B_{(s)}^0 \to SP$ from hereon

rected using simulated samples of the channels, which are described in Sect. 6.3. The systematic uncertainties associated with the normalisation process are discussed in Sect. 6.6.3 and the results of the search are presented in Sect. 6.7.

6.2 Dataset

The search for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ is conducted on data collected by the LHCb detector at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ between March and October 2011, corresponding to an integrated luminosity of 1.0 fb^{-1}. Most of the data was delivered by the LHC with a 50 ns bunch spacing scheme and 1380 circulating bunches, with 1296 colliding bunches in LHCb. The luminosity in LHCb was continuously leveled in order not to exceed $3 \times 10^{32} cm^{-2}s^{-1}$ for the first part of data taking (370 pb^{-1}), and $3.5 - 4 \times 10^{32} cm^{-2}s^{-1}$ for the remaining part of the data. The average number of pp interactions per bunch crossing was in the range of 1.4 to 1.5.

6.3 Simulation samples

Three sets of simulated events are used in the present analysis: the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$, $B^0_{(s)} \to SP$ and $B^0 \to J/\psi K^{*0}$ decay channels, with each sample containing $\sim 500k$ events. A simulated event consists of a pp interaction at $\sqrt{s} = 7\,\text{TeV}$, which is forced to produce a single instance of the signal decay. A brief description of the simulation is given in Sect. 3.7.1. For the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays the matrix element $|\mathcal{M}|$ is set to unity, such that the kinematics of the final state muons are distributed according to the phase space of the decays. These are referred to as the 'phase-space' samples and they provide a model-independent measure of the efficiencies for reconstructing and selecting the signal channel. The MSSM $B^0 \to SP$ and $B^0_s \to SP$ simulation samples are generated with matrix elements taken from [67,68]. The mass of the P sgoldstino is set to 214.3 MeV/c^2, corresponding to the mass of the HyperCP resonance [34] and the mass of S is set to 2.5 GeV/c^2. The widths of the sgoldstinos are set to 0.1 MeV/c^2. The MSSM samples are used to measure the sensitivity of the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ search to the HyperCP resonance.

6.3.1 Comparison with data

The accuracy with which the simulation reproduces the distributions in data is verified using $B^0 \to J/\psi K^{*0}$ events. The data and simulation distributions of the selection variables shown in Table 6.1 are compared for $B^0 \to J/\psi K^{*0}$ events and shown in Fig. 6.2 and Fig. 6.3. The variables are defined in Appendix A. There is good agreement between data and simulation for the p,$p_{\rm T}$ and χ^2_{IP} distributions of the π^-,K^+ and $\mu^{\pm}$, as well as for

Selection variable	Criteria
B^0 χ^2_{IP}	< 9
B^0 vertex $\chi^2 DOF$	< 30
μ χ^2_{IP}	> 16
μ DLL$_K$	< 0
μ DLL$_\mu$	> 0
K DLL$_K$	> 5
π DLL$_K$	< -5
$K^+\pi^-$ mass	$826 < M_{K\pi} < 966\,\mathrm{MeV}/c^2$
$\mu^+\mu^-$ mass	$3040 < m_{\mu^+\mu^-} < 3140\,\mathrm{MeV}/c^2$
$\mu^+\mu^-K^+$ mass	$5220 > M_{\mu^+\mu^-K^+} > 5340\,\mathrm{MeV}/c^2$
$K^+\pi^-_{M(\pi)\to M(K)}$ mass :	$950 > M_{K^+(\pi^-\to K^-)} > 1090\,\mathrm{MeV}/c^2$

Table 6.1: Selection criteria for the $B^0 \to J/\psi K^{*0}$ decay.

the χ^2_{IP} of the B^0. The B^0 vertex χ^2 distribution in the simulation does not resemble that in the data. However, the overall impact of this on the selection efficiency is negligible as the signal retention of the B^0 vertex $\chi^2 < 30$ cut used in the analysis is 96.9 ± 0.2 % in data and $96.3 \pm 0.6\%$ in the simulation.

The DLL distributions are not well reproduced by the simulation, in particular those for DLL_K. These disagreements would manifest themselves in the PID efficiencies extracted from the simulation, so to correct for this, the efficiencies are calculated using event-by-event weights as described in Sect. 6.6.1.

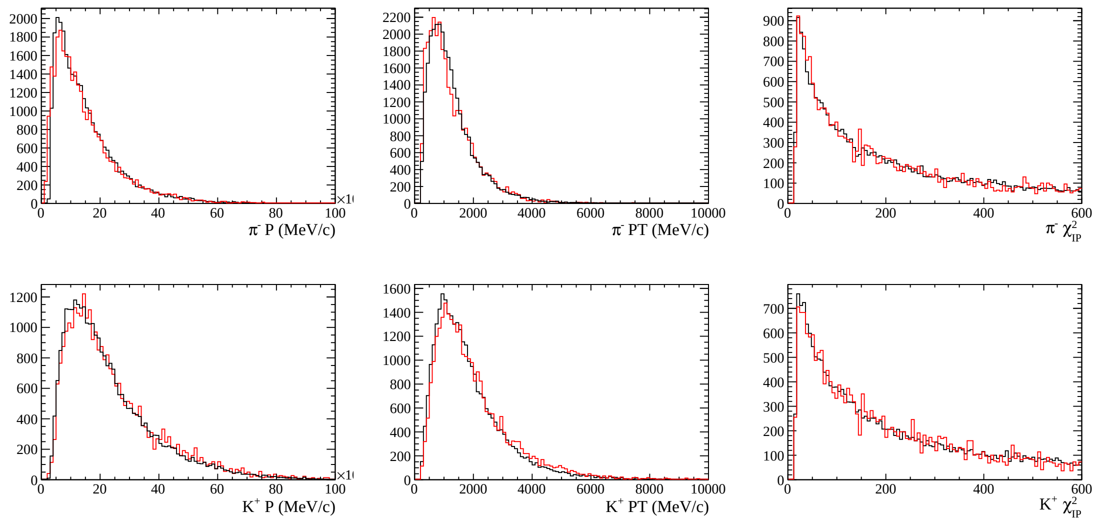


Figure 6.2: The distributions of the normalisation channel selection variables, taken from $B^0 \to J/\psi K^{*0}$ decays after application of the selection criteria in Table 6.1 in a B^0 invariant mass window of $5239.5 < M(K^+\pi^-\mu^+\mu^-) < 5319.5$ MeV/c^2. Simulated (red, grey in B&W) and data (black) events are compared.

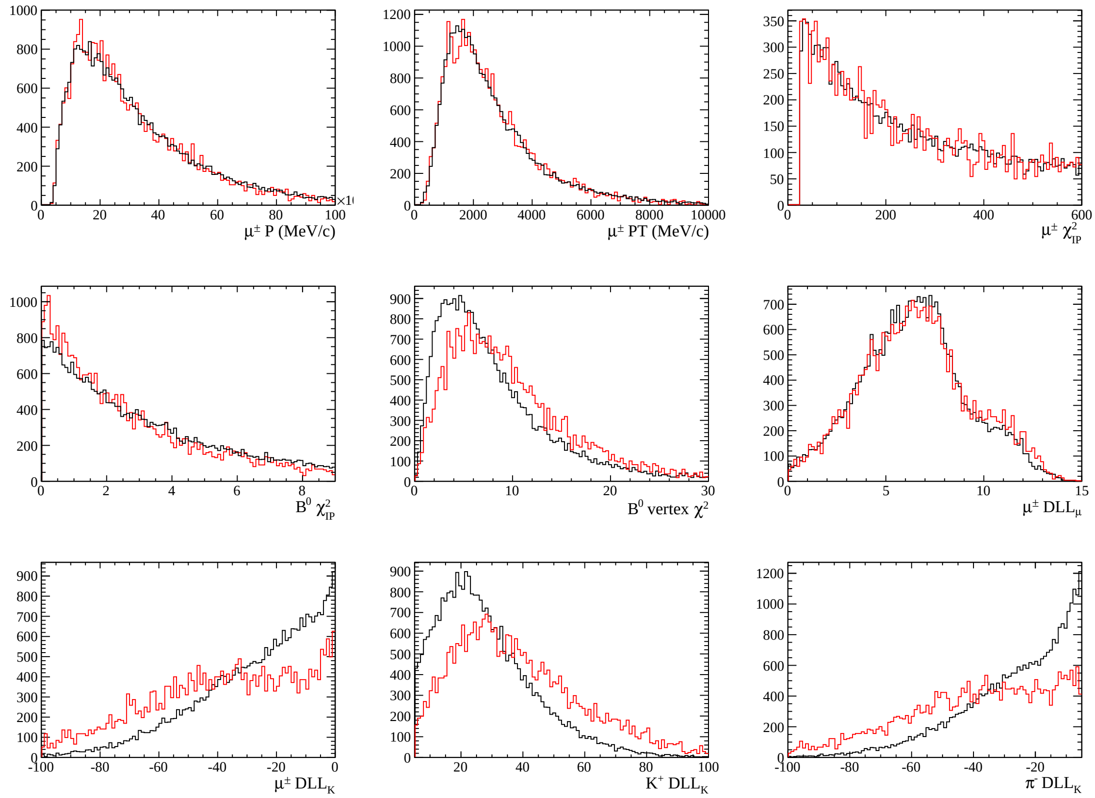


Figure 6.3: The distributions of the normalisation channel selection variables, taken from $B^0 \rightarrow J/\psi K^{*0}$ decays after application of the selection criteria in Table 6.1 in a B^0 invariant mass window of $5239.5 < M(K^+\pi^-\mu^+\mu^-) < 5319.5$ MeV/c^2. Simulated (red, grey in B&W) and data (black) events are compared.

6.4 Event selection

6.4.1 Stripping

Signal $B^0_{(s)} \rightarrow \mu^+\mu^-\mu^+\mu^-$ candidates are stripped from the LHCb dataset by selecting two pairs of oppositely charged muons which make a common vertex that is displaced from the PV. The stripping pre-selection is applied to the muons and reconstructed $B^0_{(s)}$ vertex with the criteria detailed in Table 6.2. These criteria are designed to select generic B decays and apply basic muon PID requirements. They are loosely based on those employed in the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ [69] analysis. This decay is a four-body B decay with final state muons. As such it has similar kinematic properties to the $B^0_{(s)} \rightarrow \mu^+\mu^-\mu^+\mu^-$ decays.

The $B^0 \rightarrow J/\psi K^{*0}$ normalisation decay candidates are reconstructed by initially building J/ψ and K^* candidates by requiring μ^+ μ^- and K^+ π^- pairs to make common vertices with an invariant mass consistent with the nominal J/ψ and K^* masses, respectively. The J/ψ and K^* candidates which make a common displaced vertex are used to reconstruct the B^0 meson, which is required to have a displaced decay vertex and a trajectory which

Selection variable	Criteria
μ ISMUON	True
μ p_{T}	$> 250\,\mathrm{MeV}/c$
μ χ^2_{IP}	> 9
μ track χ^2/DOF	< 5
$M_{\mu^+\mu^-\mu^+\mu^-}$	$4366 < M_{\mu^+\mu^-\mu^+\mu^-} < 6366\,\mathrm{MeV}/c^2$
B χ^2_{IP}	< 25
B vertex χ^2/DOF	< 9
B vertex DOCA$_{max}$	$< 0.3\,\mathrm{mm}$
B $\cos(\theta_{PV})$	> 0
B flight distance χ^2	> 100

Table 6.2: The selection criteria used in the stripping line for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$. Descriptions of the selection variables can be found in Appendix A.

Selection variable	Criteria
μ ISMUON	True
μ, K, π p_{T}	$> 250\,\mathrm{MeV}/c$
μ, K, π χ^2_{IP}	> 25
μ, K, π track χ^2/DOF	< 5
$M_{K\pi}$	$295.5 < M_{K\pi} < 1495.5\,\mathrm{MeV}/c^2$
K^* χ^2_{IP}	> 25
K^* vertex DOCA	$< 0.3\,\mathrm{mm}$
$M_{\mu^+\mu^-}$	$2096.9 < M_{\mu^+\mu^-} < 3196.9\,\mathrm{MeV}/c^2$
J/ψ vertex χ^2/DOF	< 9
J/ψ $\cos(\theta_{PV})$	> 0
J/ψ vertex DOCA	$< 0.3\,\mathrm{mm}$
$M_{K^+\pi^-\mu^+\mu^-}$	$4780 < M_{K^+\pi^-\mu^+\mu^-} < 5780\,\mathrm{MeV}/c^2$
B^0 flight distance χ^2	> 169
B^0 χ^2_{IP}	< 25
B^0 vertex χ^2/DOF	< 15

Table 6.3: The selection criteria used in the stripping line for $B^0 \to J/\psi K^{*0}$.

points back to the PV. The stripping selection criteria for $B^0 \to J/\psi K^{*0}$ were taken from the analysis described in [70], where this channel is also used. These are detailed in Table 6.3.

6.4.2 Trigger Requirements

Signal and normalisation channel events are required to be triggered (TIS and/or TOS) by at least one of the lines listed in Table 6.4 at each trigger level (L0, HLT1 and HLT2), the efficiencies of the individual lines, calculated from the simulated phase-space $B^0_s \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ events are also shown. At the L0 stage all major trigger lines are used to ensure maximum efficiency. The muon lines are the most efficient at selecting both channels. For the HLT1 stage the single muon, dimuon and charged

L0 efficiencies		
Trigger line	$B_s \to 4\mu$	$B_d \to K^* J/\psi$
L0DiMuon	93.1 ± 1.1 %	68.9 ± 1.0 %
L0Muon	96.6 ± 1.1 %	92.2 ± 1.2 %
L0Hadron	16.5 ± 0.4 %	22.8 ± 0.5 %
L0Photon	2.5 ± 0.1 %	2.7 ± 0.2 %
L0Electron	4.1 ± 0.2 %	4.3 ± 0.2 %
Combined L0	97.6 ± 1.2 %	94.0 ± 1.2 %
HLT1 efficiencies		
Hlt1SingleMuonHighPT	23.9 ± 0.5 %	22.0 ± 0.5 %
Hlt1TrackAllL0	87.6 ± 1.1 %	86.8 ± 1.2 %
Hlt1TrackMuon	96.9 ± 1.2 %	92.3 ± 1.3 %
Hlt1DiMuonHighMass	80.2 ± 1.0 %	80.7 ± 1.1 %
Hlt1DiMuonLowMass	96.8 ± 1.2 %	79.2 ± 1.1 %
Combined HLT1	99.4 ± 1.2 %	97.9 ± 1.3 %
HLT2 efficiencies		
Hlt2Topo2Body	70.2 ± 0.9 %	70.8 ± 1.1 %
Hlt2Topo3Body	79.5 ± 1.0 %	74.4 ± 1.1 %
Hlt2Topo4Body	57.6 ± 0.8 %	51.5 ± 0.8 %
Hlt2TopoMu2Body	82.1 ± 1.0 %	81.0 ± 1.2 %
Hlt2TopoMu3Body	84.8 ± 1.1 %	81.1 ± 1.2 %
Hlt2TopoMu4Body	60.7 ± 0.8 %	54.4 ± 0.9 %
Hlt2DiMuonDetached	93.8 ± 1.1 %	80.1 ± 1.2 %
Hlt2DiMuonDetachedHeavy	45.2 ± 0.7 %	92.0 ± 1.3 %
Combined HLT2	97.1 ± 1.2 %	96.2 ± 1.3 %
Combined L0+HLT1+HLT2	94.3 ± 1.1 %	88.2 ± 1.2 %

Table 6.4: Individual and combined efficiencies of the trigger lines used to select signal and normalisation channel decays, extracted from simulated phase-space $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ events after stripped and selection criteria are applied. The efficiencies are stated with respect to the events that have passed the lower previous trigger levels i.e. HLT2 efficiencies are given for events that have passed the L0 and HLT1 requirements.

track lines are used, with the single muon line 'Hlt1TrackMuon' being the most efficient line, having normalisation and signal channel efficiencies of 96.9 ± 1.1 % and 92.3 ± 1.3 % respectively. At the HLT2 stage the n-body topological and detached dimuon lines are used, where the latter provide the highest efficiency. The overall trigger efficiency is found to be 94.3 ± 1.1 % for $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ and 88.2 ± 1.2 % for $B^0 \to J/\psi K^{*0}$.

6.4.3 Signal channel selection

Candidate $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ decays that pass the trigger and stripping requirements are separated into non-resonant signal and resonant $B_s^0 \to J/\psi\phi$ candidate samples. The $B_s^0 \to J/\psi\phi$ sample is used as a proxy for the signal channel to develop the selection. For the signal channel all ϕ and J/ψ candidates are removed by requiring all opposite-sign

	Mass Windows (MeV/c^2)	
	B^0	B_s^0
Signal window	5240-5320	5326-5406
Blind window	5220-5340	5306-5426
Evaluation sidebands	4776-5220 and 5426-5966	
Training sidebands	4336-4776 and 5966-6366	

Table 6.5: The components of the four-muon mass range.

muon pairs to have an invariant mass outside 950-1090 MeV/c^2 and 3000-3200 MeV/c^2. For the $B_s^0 \to J/\psi\phi$ sample one dimuon pair is required to have an invariant mass of $3040 < M_{\mu^+\mu^-} < 3140\,\text{MeV}/c^2$ and the other a mass of $980 < M_{\mu^+\mu^-} < 1060\,\text{MeV}/c^2$. Candidates that do not satisfy either of the dimuon mass criteria are discarded.

The four-muon invariant mass range is split into four regions, shown in Table 6.5. The signal windows, corresponding to twice the width of the $B_{(s)}^0$ mass resolution, are used to select resonant and non-resonant $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ decays. Non-resonant events in the blind windows were not considered until after the selection was developed and the background evaluated, in order to avoid bias in the selection. The evaluation sideband is used to make an unbiased assessment of the combinatorial background that remains after the selection. Events in the training sideband of the $B_s^0 \to J/\psi\phi$ sample are used to provide a background sample with which to develop the selection algorithm.

The selection is developed to maximise the signal S to background B metric $S/\sqrt{S+B}$, where S and B are the number of $B_s^0 \to J/\psi\phi$ candidates in the B_s^0 signal window and the training sideband, respectively. The selection metric is maximised by sequentially varying the following cut variables:

- The χ^2_{IP} of the muons and reconstructed $B_{(s)}^0$, to ensure that the former are not consistent with originating from the PV and that the latter is.

- Muon DLL_μ and DLL_K, so that the muon candidates are consistent with the muon PID hypothesis and not the kaon hypothesis. The latter condition minimises the number of kaons that are mis-identified as muons after they decay into $\mu\nu_\mu$ upstream of the muon chambers.

- The $B_{(s)}^0$ vertex χ^2, to ensure that a good quality vertex is made by the four muons.

The optimal selection criteria are shown in Table 6.6. After application of these criteria one event remains in the training sideband and seven $B_s^0 \to J/\psi\phi$ candidates remain in the signal window. The expected $B_s^0 \to J/\psi\phi$ yield obtained from simulation and normalisation to $B^0 \to J/\psi K^{*0}$ is 5.5±2.3, consistent with the observed yield.

Selection variable	Criteria
$B^0_{(s)}$ χ^2_{IP}	< 9
$B^0_{(s)}$ vertex $\chi^2 DOF$	< 30
μ χ^2_{IP}	> 16
μ DLL$_K$	< 0
μ DLL$_\mu$	> 0
B^0 mass	$5326 < M_{\mu^+\mu^-\mu^+\mu^-} < 5406\,\text{MeV}/c^2$
B^0_s mass	$5240 < M_{\mu^+\mu^-\mu^+\mu^-} < 5320\,\text{MeV}/c^2$
non-resonant channel only	
μ^+ μ^- mass	$3000 < M_{\mu^+\mu^-} < 3200\,\text{MeV}/c^2$
μ^+ μ^- mass	$950 < M_{\mu^+\mu^-} < 1090\,\text{MeV}/c^2$
$B^0_s \to J/\psi\phi$ channel only	
μ^+ μ^- mass	$3040 < M_{\mu^+\mu^-} < 3140\,\text{MeV}/c^2$
μ^+ μ^- mass	$980 < M_{\mu^+\mu^-} < 1060\,\text{MeV}/c^2$

Table 6.6: Selection criteria for the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ channels.

6.4.4 Normalisation channel selection

The selection criteria for $B^0 \to J/\psi K^{*0}$, shown in Table 6.1, are identical to those of the signal channel, with the removal of the lower $M_{\mu^+\mu^-}$ cut and the addition of the following cuts:

- The $K^+\pi^-$ invariant mass is required to be within ±100 MeV/c^2 of the nominal K^{*0} mass.
- The $\mu^+\mu^-$ invariant mass is required to be within ±50 MeV/c^2 of the nominal J/ψ mass.
- DLL_K cuts are applied on the kaon and pion candidates such that they are consistent with their respective PID hypotheses. This substantially reduces the number of background events and removes duplicate candidates where the K and π mass hypotheses are exchanged.
- The background arising from $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ decays which are combined with a pion from elsewhere in the event is removed by excluding candidates with a $K^+\mu^+\mu^-$ invariant mass within $\pm60\,\text{MeV}/c^2$ of the B^+ mass.
- Background $B^0_s \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ decays, where one of the kaons is mis-identified as a pion, are removed by excluding events with a $K^+\pi^-_{M(\pi)\to M(K)}$ invariant mass of $\pm70\,\text{MeV}/c^2$ around the ϕ mass, where the pion is assigned a kaon mass hypothesis.

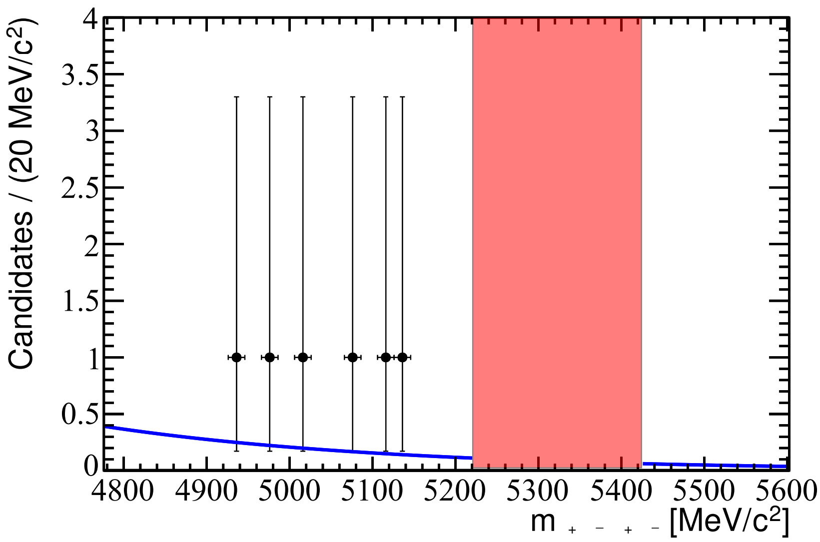


Figure 6.4: The non-resonant $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ evaluation sideband after selection, fitted with a single exponential PDF. The pink region denotes the blind mass windows, the events within are not used to perform the fit.

6.5 Background evaluation

After application of the selection criteria six non-resonant events are seen in the background evaluation sidebands, shown in Fig. 6.4. Peaking backgrounds, which arise from other four-body B^0 and B^0_s decays where the final state particles are all (mis)identified as muons are considered for the search for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$. Peaking backgrounds are estimated to have a negligible contribution to the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ signal window yields. Of these backgrounds the $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^*(\to K^+\pi^-)$ contribution is the largest, with an expected yield after selection of 0.44 ± 0.06 events, of which less than 0.1 fall within the B^0 or B^0_s signal windows. This follows from the fact that the mis-identification of the kaon and pion mass hypotheses shifts the invariant mass of the reconstructed B^0 meson far below the B^0 mass window.

Fit model	B^0	B^0_s
Single exponential	$0.377^{+0.233}_{-0.174}$	$0.295^{+0.218}_{-0.201}$
Double exponential	$0.377^{+0.232}_{-0.173}$	$0.295^{+0.219}_{-0.200}$
Linear	$0.383^{+0.178}_{-0.135}$	$0.285^{+0.132}_{-0.101}$

Table 6.7: The background expectations in the $B^0_{(s)}$ mass windows, using single exponential, double exponential and linear fit models. The single exponential model is used to extract the background expectations.

The dominant source of background is combinatorial, where a $B^0_{(s)}$ candidate vertex is made from four particles that did not originate from a single $B^0_{(s)}$ meson and are (mis)identified as muons. This background is evaluated by fitting the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ background evaluation sidebands with a single exponential PDF and extrapolating the resulting fit into the B^0 and B_s^0 signal windows. The resulting background yields are 0.4±0.2 in the B^0 window and 0.3±0.2 in the B_s^0 window. Background fits using double exponential[2] and linear PDFs give background expectations consistent with the single exponential fit, shown in Table 6.7. From this it can be concluded that the background fit model has a negligible impact on the expected background yield.

6.6 Normalisation

The $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ yield is converted to a branching fraction by normalising to $B^0 \to J/\psi K^{*0}$ using the following equation:

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-) = \mathcal{B}(B^0 \to J/\psi K^{*0}) \times \frac{(\epsilon_{rec\&strip}\epsilon_{sel|rec\&strip}\epsilon_{trig|sel})_{B^0\to J/\psi K^{*0}}}{(\epsilon_{rec\&strip}\epsilon_{sel|rec\&strip}\epsilon_{trig|sel})_{B^0_{(s)}\to\mu^+\mu^-\mu^+\mu^-}} \frac{N_{B^0_{(s)}\to\mu^+\mu^-\mu^+\mu^-}}{N_{B^0\to J/\psi K^{*0}}} \left(\frac{f_{d(s)}}{f_d}\right)^{-1} \kappa \quad (6.1)$$

where:

- $\mathcal{B}(B^0 \to J/\psi K^{*0})$ is the branching fraction of the normalisation channel, calculated as the product of the $B^0 \to J/\psi K^{*0}$, $J/\psi \to \mu^+\mu^-$ and $K^{*0} \to K^+\pi^-$ branching fractions, where the S-wave component of the non-resonant $B^0 \to J/\psi\ K^+\ \pi^-$ decay is removed [1, 71].

- $\epsilon_{rec\&strip}$,$\epsilon_{sel|rec\&strip}$ and $\epsilon_{trig|sel}$ are the sequential efficiencies with which the decay channel events are reconstructed and stripped, pass the selection criteria and then satisfy the trigger requirements, respectively. The efficiencies are calculated using the simulated events described in Sect. 6.3. For example, $\epsilon_{sel|rec\&strip}$ is calculated as the fraction of reconstructed and stripped events that pass the selection criteria

- $N_{B^0_{(s)}\to\mu^+\mu^-\mu^+\mu^-}$ and $N_{B^0\to J/\psi K^{*0}}$ are the yields of the signal and normalisation channels, respectively. $N_{B^0_{(s)}\to\mu^+\mu^-\mu^+\mu^-}$ is obtained by counting the number of events in the non-resonant $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ signal window. The $B^0 \to J/\psi K^{*0}$ yield is obtained from a fit to the $K^+\pi^-\mu^+\mu^-$ invariant mass distribution, described in Sect. 6.6.2.

- f_s/f_d is the relative production fraction for B^0 and B_s^0 mesons, measured by LHCb to be 0.256 ± 0.020 [18].

- The factor κ removes the non-resonant S-wave contribution to the $B^0 \to J/\psi K^{*0}$ yield and efficiencies. This is necessary as the simulation sample and $\mathcal{B}(B^0 \to$

[2] A superposition of two exponential functions with independent coefficients

channel	$\epsilon_{rec\&strip}$ / %	$\epsilon_{sel\|rec\&strip}$ / %	$\epsilon_{trig\|sel}$ / %	ϵ_{tot} / %
$B^0 \to J/\psi K^{*0}$	1.48 ± 0.01	21.58 ± 0.23	88.16 ± 1.22	0.282 ± 0.003
$B^0_s \to \mu^+\mu^-\mu^+\mu^-$	2.12 ± 0.01	17.95 ± 0.16	94.27 ± 1.14	0.359 ± 0.003
$B^0_s \to SP$	2.13 ± 0.01	18.31 ± 0.17	93.96 ± 1.14	0.366 ± 0.003
$\frac{B_d \to K^* J/\psi}{B^0_s \to \mu^+\mu^-\mu^+\mu^-}$	69.90 ± 0.42	120.18 ± 1.67	93.51 ± 1.72	78.55 ± 1.06
$B^0 \to \mu^+\mu^-\mu^+\mu^-$	2.13 ± 0.01	17.46 ± 0.16	93.74 ± 1.13	0.349 ± 0.003
$B^0 \to SP$	2.14 ± 0.01	18.06 ± 0.17	93.71 ± 1.14	0.361 ± 0.003
$\frac{B_d \to K^* J/\psi}{B^0 \to \mu^+\mu^-\mu^+\mu^-}$	69.55 ± 0.41	123.55 ± 1.71	94.04 ± 1.72	80.81 ± 1.08

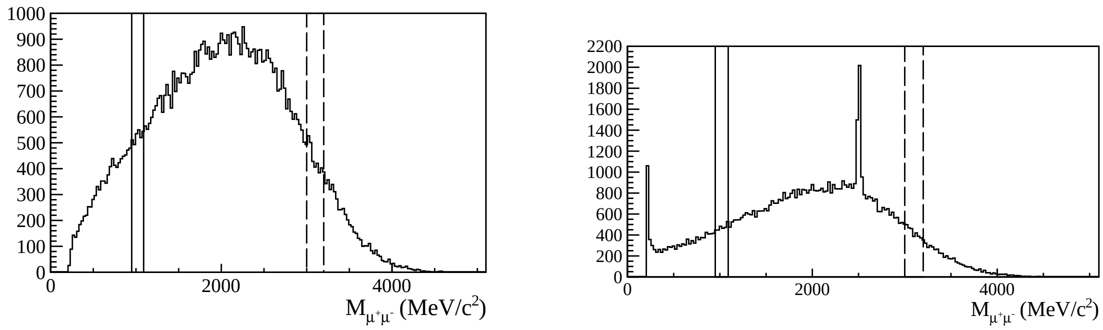
Table 6.8: Values for the efficiencies shown in equation (6.1), calculated using the simulation samples described in Sect. 6.3. The errors given are purely statistical.

$J/\psi K^{*0}$) take into account the resonant component only. An angular analysis detailed in Appendix E is performed to extract a value of $\kappa = 1.09 \pm 0.09$.

6.6.1 Efficiencies

All non-PID efficiencies are calculated using the simulated event samples described in Sect. 6.3. The PID efficiencies are calculated from data using the kaon, pion and muon PID calibration samples described in Appendix B. These data are divided into bins of momentum, p_{T} and charged track multiplicity ($nTracks$), as the PID efficiency is dependent on these variables. Each simulated particle which has a DLL cut applied is assigned a weight corresponding to the efficiency of the cut when applied to the track sample in the appropriate $(p, p_{\mathrm{T}}, nTracks)$ bin of the PID calibration sample. The weights are multiplied together to produce per-event weights and the overall PID efficiency is then taken as the mean of these per-event weights in a given bin. A similar procedure is used to correct for the track reconstruction efficiency when calculating $\epsilon_{rec\&strip}$ by applying weights that correspond to the ratio of the tracking efficiency between data and the simulation to each final state particle .

The individual and combined components of the signal and normalisation channel efficiencies are shown in Table 6.8. The ratios of the $B^0 \to J/\psi K^{*0}$ to the non-resonant $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ efficiencies are also shown. The combined reconstruction and stripping efficiency of the signal channel is greater than for $B^0 \to J/\psi K^{*0}$. This is mainly due to the soft kinematics of the pions, 25.1 % of which have momenta less than 3 GeV/c and are swept outside of the full LHCb acceptance by the magnet. For signal channel muons the equivalent fraction is 3.2 %. The lower selection efficiencies of the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ channels are caused by the PID selections, which are tighter than those for $B^0 \to J/\psi K^{*0}$ as eight cuts are applied (two for each muon), compared to just six for $B^0 \to J/\psi K^{*0}$. The trigger efficiency of the signal channel is slightly greater than that of the normalisation channel because the former has more final state muons with which to satisfy the trigger criteria, which are dominated by muon-dedicated lines.



(a) (b)

Figure 6.5: The $\mu^+\mu^-$ invariant mass distributions of the phase-space $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ (a) and MSSM $B_s^0 \to SP$ (b) simulation samples. The solid and dashed lines respectively indicate the boundaries of the ϕ and the J/ψ mass vetoes.

The overall MSSM $B_{(s)}^0 \to SP$ efficiencies are 2-3 % larger than for the phase-space $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ sample. This is because the J/ψ and ϕ mass vetoes remove a smaller fraction of events in the dimuon mass distributions in the $B_{(s)}^0 \to SP$ sample, where the $M_{\mu^+\mu^-}$ spectra are dominated by the hypothesised S and P resonances, neither of which fall into the veto regions shown by the vertical lines in Fig. 6.5.

6.6.2 $B^0 \to J/\psi K^{*0}$ fit and yield

The $K^+\pi^-\mu^+\mu^-$ invariant mass distribution of candidate $B^0 \to J/\psi K^{*0}$ events passing the selection is shown in Fig. 6.6. The distribution has three components, the $B^0 \to J/\psi K^{*0}$ and $\overline{B}_s^0 \to J/\psi K^{*0}$ mass peaks, and a combinatorial background slope. The mass peaks are fitted using Crystal Ball (CB) functions, defined in Appendix C. Each mass peak is fitted with a combination of two Crystal Ball functions:

$$PDF_{B_{(s)}^0} = f \times CB(M_{K^+\pi^-\mu^+\mu^-}, \overline{M}_{B_{(s)}^0}, \sigma_1, \alpha_1, n) + (1-f) \times CB(M_{K^+\pi^-\mu^+\mu^-}, \overline{M}_{B_{(s)}^0}, \sigma_2, \alpha_2, n) \quad (6.2)$$

where $0 < f < 1$, the variables $f, \sigma_1, \alpha_1, \sigma_2, \alpha_2$ and n are common to both B_s^0 and B^0 PDFs and are left free in the fit and the mean mass variables $\overline{M}_{B_{(s)}^0}$ are required to satisfy $\overline{M}_{B_s^0} - \overline{M}_{B^0} = 87.3\,\mathrm{MeV}/c^2$, the difference between the nominal B_s^0 and B^0 masses. The background shape is fitted with a single exponential PDF:

$$PDF_{bkg} = \exp(\beta \times M_{K^+\pi^-\mu^+\mu^-}) \quad (6.3)$$

The PDFs defined in (6.2) and (6.3) are combined to make the total fit PDF:

$$PDF_{tot} = PDF_{B^0}(M_{K^+\pi^-\mu^+\mu^-}) + PDF_{B_s^0}(M_{K^+\pi^-\mu^+\mu^-}) + PDF_{bkg}(M_{K^+\pi^-\mu^+\mu^-}) \quad (6.4)$$

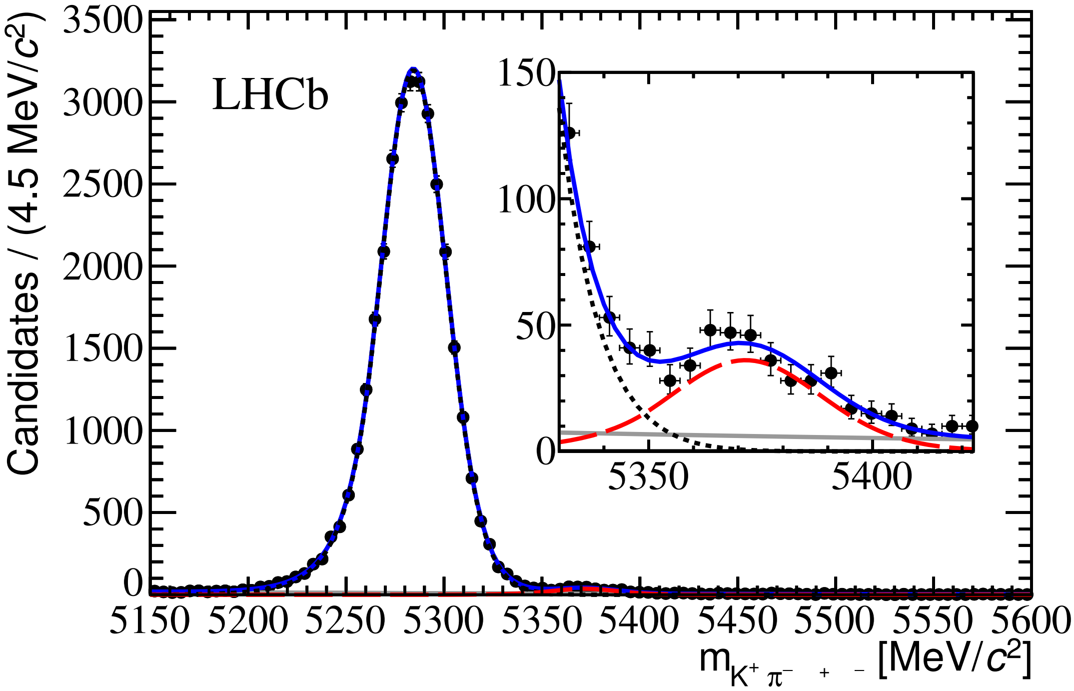


Figure 6.6: Invariant mass distribution of $K^+\pi^-\mu^+\mu^-$ candidates after selection. The B^0 and $\overline{B}{}^0_s$ signal PDFs are shown by short-dashed black and long-dashed red (grey in B&W) lines, respectively. The background PDF is shown in light grey. The total fit PDF is shown as a solid blue (dark grey in B&W) line. The inset shows the mass distribution centred around the $\overline{B}{}^0_s$ mass.

Parameter	Fit Value
α_1	1.41 ± 0.07
α_2	1.53 ± 0.14
n	15.5 ± 11.1
f	0.73 ± 0.13
$\overline{M}_{B^0}$	5284.3 ± 0.1 MeV/c^2
$\overline{M}_{B^0_s}$	5371.6 ± 0.1 MeV/c^2
σ_1	15.9 ± 0.6 MeV/c^2
σ_2	23.4 ± 1.8 MeV/c^2
β	$(-4.74 \pm 0.41) \times 10^{-3}(\text{MeV}/c^2)^{-1}$
N_{B^0}	31837 ± 183
$N_{B^0_s}$	363 ± 27
N_{bkg}	711 ± 53

Table 6.9: Values of the $B^0 \to J/\psi K^{*0}$ fit parameters for the PDF in (6.4).

The $B^0 \to J/\psi K^{*0}$ fit results are shown in Table 6.9 and Fig. 6.6. The respective yields of the $B^0 \to J/\psi K^{*0}$, $\overline{B}^0_s \to J/\psi K^{*0}$ and combinatorial background are 31837 ± 183, 363 ± 27 and 711 ± 53.

6.6.3 Uncertainties

Statistical and systematic uncertainties are associated with the individual elements of (6.1). The significant systematic uncertainties are mostly associated with the methods used to correct data-simulation disagreements.

A potential systematic uncertainty might arise from the difference between the four-body phase space of the $B^0 \to J/\psi K^{*0}$ and the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays. The $B^0_{(s)} \to SP$ decay in particular has a very different phase-space, due to the small mass difference between P and the μ^+ μ^- pair, which would have a very small opening angle. The LHCb reconstruction software would be able to reconstruct such a decay because although the trajectories of the two muons would be co-linear in the VELO, the two muons would be bent in opposite directions by the magnet, allowing both tracks to be reconstructed. Any systematic uncertainty is deemed to be negligible, as previous studies has been shown that the LHCb simulation very precisely describes the reconstruction and acceptance efficiencies of decays with small opening angles between the final state particles, such as $B^0 \to K^{*0}\phi$ [72], where $\phi \to K^+K^-$.

IP smearing

The track smearing used to correct the IP resolution in the simulation, is detailed in Sect. 3.7.1. The systematic uncertainty associated with this treatment is taken as the relative change in the ratio between the signal and normalisation channel efficiencies when the smearing scale is varied between 0, corresponding to no smearing, and 2.0, double the amount required to reproduce the data IP resolution. The variation is found to be $\pm 4.3\%$

Trigger efficiency

The trigger efficiency of the normalisation channel is calculated in data, using the 'TIS-TOS' method described in Sect. 3.6 and [73]. The resulting efficiency is $92.1 \pm 5.6\%$, compared to $88.2 \pm 1.2\%$ for the simulation. The relative difference between the two of 4.4% is assigned as the systematic uncertainty.

Track reconstruction efficiency

Hadronic interactions in the detector are not taken into account when calculating the corrections to the track reconstruction efficiency, as these are extracted using muons. An additional systematic uncertainty of 1.1-1.5% is associated with reconstructing a hadron [15]. Assigning the conservative end of this uncertainty to both the kaon and pion in the normalisation channel results in a combined uncertainty of 3.0%.

PID efficiency

The dominant uncertainty associated with the reweighting technique used to extract the PID efficiency arises from the choice of binning scheme used for the PID calibration samples. This is assessed by exchanging the PID weights of tracks which have either p, p_{T} or $nTracks$ values within 1/10 of a bin-width from the bin edge, with the weight of the adjacent bin. For example, two momentum bins are $5 < p < 9.3\,\mathrm{GeV}/c$ and $9.3 < p < 15.6\,\mathrm{GeV}/c$, if a muon in the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ simulation sample has a momentum of 9.4 GeV/c, the PID weight assigned to it will be that of the $5 < p < 9.3\,\mathrm{GeV}/c$ bin instead of the $9.3 < p < 15.6\,\mathrm{GeV}/c$ bin, and vice-versa if the muon has $p = 9.2\,\mathrm{GeV}/c$.

After re-binning the weights as described above, the $B^0_s \to \mu^+\mu^-\mu^+\mu^- / B^0 \to J/\psi K^{*0}$ efficiency ratio is 75.3%, compared to 78.6% without rebinning. The 4.1 % relative difference between the two numbers is taken as the PID systematic uncertainty.

S-wave correction

Details of the S-wave analysis are presented in Appendix E. A systematic uncertainty of 8.3 % is assigned to κ.

$\mathbf{f_s/f_d}$

The uncertainty associated with the B^0_s/B^0 meson production ratio is 7.8%, obtained from [18].

Source	Uncertainty [%]
$\mathcal{B}(B^0 \to J/\psi K^{*0})$	10.2
S-wave correction	8.3
f_d/f_s	7.8
Data-simulation differences	5.2
Trigger efficiency	4.4
PID selection efficiency	4.1
Simulation sample size	1.3
$B^0 \to J/\psi K^{*0}$ yield	0.6
Combined B_s^0 uncertainty	17.2
Combined B^0 uncertainty	15.4

Table 6.10: Uncertainties associated with calculating $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-)$. The combined uncertainties are calculated by adding the individual components in quadrature.

Statistical uncertainties

Statistical uncertainties are associated with the efficiencies ϵ due to the size of the simulated event samples from which they are calculated and the $B^0 \to J/\psi K^{*0}$ yield, the former is calculated as a binomial error to be 1.3% and the latter is obtained from the fit uncertainty as 0.6%.

Combined uncertainty

The uncertainties associated with calculating $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-)$ are summarised in Table 6.10. When combined in quadrature the overall systematic uncertainties associated with B^0 and B^0_s are 15.4% and 17.2%, respectively, the B^0_s uncertainty is larger due to the inclusion of the f_s/f_d uncertainty.

6.7 Results

The full $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ invariant mass distribution is shown in Fig. 6.7. One candidate is seen in the B^0 mass window, none in the B^0_s window. The $\mathrm{CL_s}$ method [74, 75] is used to assess whether the observations are consistent with that expected for background only (H_b) or signal with background (H_{s+b}) hypotheses. This is done by defining a test statistic Q:

$$Q = \frac{e^{-(s(\mathcal{B})+b)}(s(\mathcal{B})+b)^d}{e^{-(b)}(b)^d} \tag{6.5}$$

where $s(\mathcal{B})$ is the expected number of signal events ($N_{B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-}$) calculated using (6.1) for a given input value of $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-)$, b is the expected number of background events shown in Table 6.7 and obtained from the fit to the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ mass

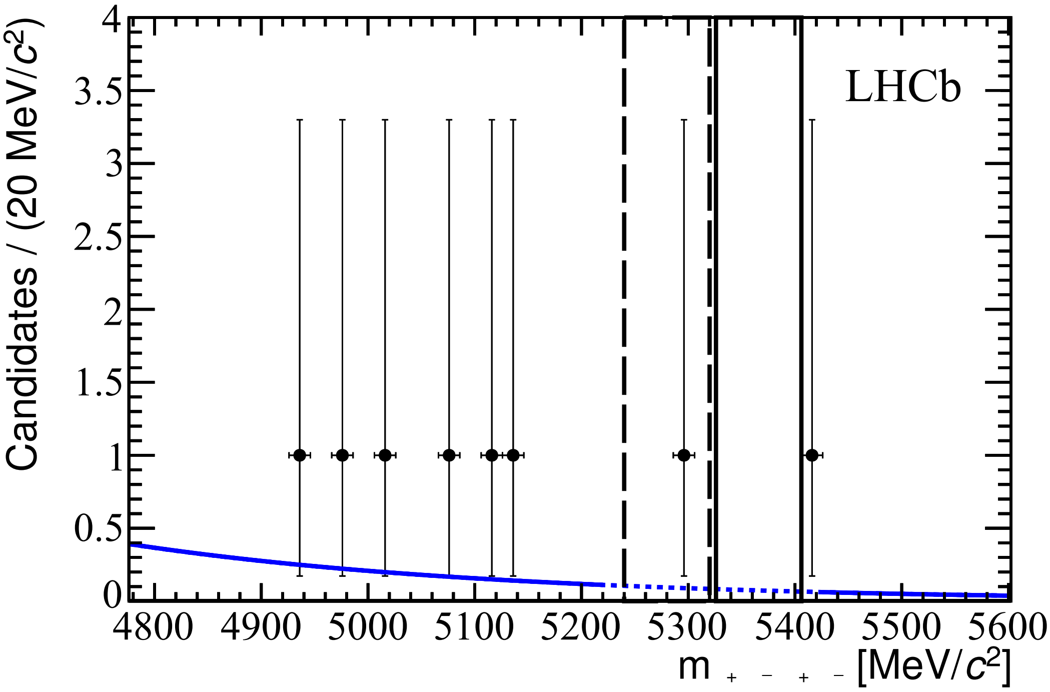


Figure 6.7: Invariant mass distribution of non-resonant $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ candidates. The solid (dashed) black lines indicate the boundaries of the B^0_s (B^0) signal window. The blue (grey in B&W) curve shows the single exponential PDF fit to the events in the background evaluation mass sidebands, indicated by the solid component of the fit, the dashed component corresponds to the blind windows.

sidebands and d is the number of observed events in the signal window. The confidence levels for H_b and H_{s+b} are defined as:

$$\mathrm{CL_{s+b}} = \int_{Q_{obs}}^{\infty} PDF_{s+b}(Q)dQ \qquad \mathrm{CL_b} = \int_{Q_{obs}}^{\infty} PDF_b(Q)dQ \tag{6.6}$$

where Q_{obs} is the value of the test statistic (6.5) when d is set to the observed number of events, $PDF_{s+b}(Q)$ and $PDF_b(Q)$ are probability distributions of the test statistic for the H_{s+b} and H_b hypotheses, each of which are generated from 10,000 instances of Q where d is fluctuated around a normal distribution with the mean set to b for H_b and $s(\mathcal{B}) + b$ for H_{s+b} and the widths set in accordance with their respective errors as shown in Table 6.10 for $s(\mathcal{B})$ and Table 6.7 for b.

The probability that the observed number of signal events is consistent with the background only hypothesis is given by 1-$\mathrm{CL_b}$. The 3 and 5σ significance thresholds of 1-$\mathrm{CL_b}$ define the criteria used for evidence and discovery of a decay channel, where 3σ corresponds to a one-sided (two-sided) probability of $1 - \mathrm{CL_b} = 2.70(1.35) \times 10^{-3}$ and 5σ to $1 - \mathrm{CL_b} = 5.73(2.87) \times 10^{-5}$. The observation of one $B^0 \to \mu^+\mu^-\mu^+\mu^-$ candidate yields $1 - \mathrm{CL_b} = 0.07$, well within the 3σ boundary and therefore consistent with background expectations, as is $B^0_s \to \mu^+\mu^-\mu^+\mu^-$, where no signal candidates are observed.

The ratio $\mathrm{CL_s} = \mathrm{CL_{s+b}}/\mathrm{CL_b}$ is used to set upper limits on the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ branching fractions. The 95%(90%) confidence level boundaries are defined as the values of $\mathcal{B}$ which yield $\mathrm{CL_s} = 0.05(0.1)$. For the non-resonant $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ signal models the limits are set at:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-\mu^+\mu^-) < 1.6\ (1.2) \times 10^{-8},$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 6.6\ (5.3) \times 10^{-9}.$$

The corresponding limits for the MSSM model with $B^0_{(s)} \to SP$ and the mass of $P(S)$ set to 214.3 MeV/c^2 (2.5 GeV/c^2), are

$$\mathcal{B}(B^0_s \to SP) < 1.6\ (1.2) \times 10^{-8},$$

$$\mathcal{B}(B^0 \to SP) < 6.3\ (5.1) \times 10^{-9}.$$

where both S and P decay into $\mu^+\ \mu^-$.

6.8 Conclusions

No evidence is found for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays in 1 fb^{-1} of LHCb data with $\sqrt{s} = 7\,\mathrm{TeV}$. The upper limits set on $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-)$ are ~ 2 orders of magnitude above the SM expectations. This is the case for both non-resonant and MSSM channels, where the decay is mediated via P and S sgoldstinos, with P being the 214.3 MeV/c^2 HyperCP resonance. When the mass of S is varied across the allowed phase space of the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decay the 95% confidence level limit on the branching fraction varies by $^{+6}_{-23}$% with respect to the limits of the default model with $m_S = 2.5\,\mathrm{GeV}/c^2$ (see Appendix F and [76]).

Subsequent updates for the search for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays will begin to probe the SM branching fractions. A further factor of three will be obtained after inclusion of the 2 fb^{-1} 2012 dataset. Future datasets include the expected 5 fb^{-1} of $\sqrt{s} = 14\,\mathrm{TeV}$ data which will be collected after the first consolidation of the LHC, and a further $\sim$45 fb^{-1} of $\sqrt{s} = 14\,\mathrm{TeV}$ data collected after the LHCb detector upgrade in 2018 [77, 78], assuming that the present efficiencies detailed in Sect. 6.6.1 are maintained.

Chapter 7

Conclusions

This thesis presents the results of three seperate analyses performed with data collected by the LHCb detector.

The photoelectron yield of the RICH detector is measured in Chapter 4 to be 15%(19%) less than that in the simulation for the C_4F_{10} (CF_4) radiator medium. The result is a slightly reduced PID performance, which is still sufficient to achieve the physics goals of LHCb.

The measurement of the p_{T} and η dependence of the $b \to \Lambda_b^0$ to $b \to B^0$ hadronisation ratio is presented in Chapter 5. The p_{T} dependence is seen to be of an exponential form with a plateau at high p_{T}, an improvement on the linear and purely exponential models used to fit the dependence in previous experiments. A linear dependence on η is also observed. These measurements can provide guidance to the development of QCD models and simulation frameworks that describe b quark hadronisation.

The search for rare $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays presented in Chapter 6 sets the world's first upper limits on their branching fractions. These limits begin to exclude phase-space regions of specific MSSM models involving S and P sgoldstinos.

Bibliography

[1] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. D **86** (2012) 010001.

[2] C. Lefevre, *LHC: the guide*, http://cds.cern.ch/record/1092437, Jan, 2008.

[3] LHCb collaboration, S. A. et. al., *LHCb : Technical Proposal*. Tech. Proposal. CERN, Geneva, 1998. CERN-LHCC-98-004.

[4] F. Alessio, R. Jacobsson, and R. Le Gac, *LHCb Operations Plots Webpage*, .

[5] LHC Collaboration. LHC Programme Coordination web pages, `http://lpc.web.cern.ch/lpc/`.

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

[7] P. Rodríguez Pérez, *The lhcb vertex locator performance and vertex locator upgrade*, Journal of Instrumentation **7** (2012) C2008, arXiv:1209.4845.

[8] S. e. a. Amato, *LHCb RICH: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.

[9] C. G. Blanks, *V^0 production ratios at LHCb and the alignment of its RICH detectors*. PhD thesis, Imperial College London, London, 2011. Presented 14 October, 2011.

[10] LHC Collaboration. LHCb panoramix displays, `http://lhcb-reconstruction.web.cern.ch/lhcb-reconstruction/panoramix/prplots/samples.htm`.

[11] M. e. a. Adinolfi, *Performance of the LHCb RICH detector at the LHC*, European Physical Journal C **73** (2013) 2431, arXiv:1211.6759.

[12] LHCb Collaboration, R. Aaij *et al.*, *Measurement of b-hadron branching fractions for two-body decays into charmless charged hadrons*, JHEP **1210** (2012) 037, arXiv:1206.2794.

[13] M. e. a. Adinolfi, *Performance of the Muon Identification at LHCb*, ArXiv e-prints (2013) arXiv:1306.0249.

[14] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[15] A. Jaeger, P. Seyfert, M. De Cian, J. van Tilburg, and S. Hansmann-Menzemer, *Measurement of the track finding efficiency*, Tech. Rep. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.

[16] R. Young and LHCb-Rich Collaboration, *Operating the Hybrid Photon Detectors in the LHCb RICH counters*, Nuclear Instruments and Methods in Physics Research A **639** (2011) 94.

[17] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of B-Hadron, C-Hadron, and tau-lepton properties as of early 2012*, arXiv:1207.1158.

[18] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of fragmentation functions f_s/f_d and the dependence on B meson kinematics*, arXiv:1301.5286. submitted to JHEP.

[19] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput.Phys.Commun. **136** (2001) 198, arXiv:hep-ex/0011057.

[20] J. Thomson, *Cathode rays*, The Electrician (1897), no. v. 39.

[21] A. A. Michelson and E. W. Morley., *On the relative motion of the earth and the luminiferous ether*, American Journal of Science (1887) .

[22] J. C. Maxwell, *On physical lines of force*, Philosophical Magazine **90** (2010), no. 1 supp 1 11.

[23] A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, Annalen der Physik **322** (1905) 132.

[24] D. Griffiths, *Introduction to Elementary Particles.* John Wiley & Sons, New York, USA, 1987.

[25] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory.* Addison-Wesley, Reading, 1995.

[26] C. Burgess and G. Moore, *The standard model: A primer*, .

[27] H. Yukawa, *On the Interaction of Elementary Particles. I* *, Progress of Theoretical Physics Supplement **1** (1955) 1.

[28] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[29] LHCb collaboration, R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys.Rev.Lett. (2013) arXiv:1304.6173.

[30] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K_S^0$ decays*, Phys.Lett. **B721** (2013) 24, arXiv:1211.6093.

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

[32] Planck Collaboration, P. Ade *et al.*, *Planck 2013 results. I. Overview of products and scientific results*, arXiv:1303.5062.

[33] J. Goldstone, A. Salam, and S. Weinberg, *Broken symmetries*, Phys. Rev. **127** (1962) 965.

[34] HyperCP collaboration, H. Park *et al.*, *Evidence for the decay $\Sigma^+ \to p\mu^+\mu^-$*, Phys.Rev.Lett. **94** (2005) 021801, arXiv:hep-ex/0501014.

[35] G. Valencia, *$\Sigma^+ \to p\mu^+\mu^-$: Standard Model or New Particle?*, arXiv:0805.3285.

[36] e865 Collaboration, H. Ma *et al.*, *A New measurement of the rare decay $K^+ \to \pi^+\mu^+\mu^-$* , Phys.Rev.Lett. **84** (2000) 2580, arXiv:hep-ex/9910047.

[37] X.-G. He, J. Tandean, and G. Valencia, *Light Higgs production in hyperon decay*, Phys.Rev. **D74** (2006) 115015, arXiv:hep-ph/0610274.

[38] Belle Collaboration, H. Hyun *et al.*, *Search for a Low Mass Particle Decaying into $\mu^+\mu^-$ in $B^0 \to K^{*0}X$ and $B^0 \to \rho^0 X$ at Belle*, Phys.Rev.Lett. **105** (2010) 091801, arXiv:1005.1450.

[39] LHCb collaboration, R. Aaij *et al.*, *Search for rare $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays*, Phys.Rev.Lett. **110** (2013) 211801, arXiv:1303.1092.

[40] Gargamelle Neutrino Collaboration, F. Hasert *et al.*, *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*, Phys.Lett. **B46** (1973) 138.

[41] UA1 Collaboration, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV*, Phys.Lett. **B122** (1983) 103.

[42] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, Phys.Lett. **B126** (1983) 398.

[43] LHCb Collaboration, R. Aaij *et al.*, *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s} = 7$ TeV *in the forward region*, Phys.Lett. **B694** (2010) 209, arXiv:1009.2731.

[44] P. M. Bjrnstad, *Performance of the lhcb vertex locator*, Journal of Instrumentation **6** (2011), no. 12 C12024.

[45] LHCb collaboration, R. Aaij *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 J.Phys. **15** (2013) 053021, arXiv:1304.4741.

[46] P. A. Čerenkov, *Visible radiation produced by electrons moving in a medium with velocities exceeding that of light*, Phys. Rev. **52** (1937) 378.

[47] R. W. Forty and O. Schneider, *Rich pattern recognition*, Tech. Rep. LHCb-98-040, CERN, Geneva, Apr, 1998.

[48] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 Physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[49] I. Belyaev *et al.*, *Handling of the generation of primary events in* Gauss*, the LHCb simulation framework*, Nuclear Science Symposium Conference Record (NSS/MIC) **IEEE** (2010) 1155.

[50] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[51] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in* Z *and* W *decays*, Eur.Phys.J. **C45** (2006) 97, arXiv:hep-ph/0506026.

[52] GEANT4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250; GEANT4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans.Nucl.Sci. **53** (2006) 270.

[53] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: design, evolution and experience*, J.Phys.Conf.Ser. **331** (2011) 032023.

[54] T. Ypsilantis and J. Seguinot, *Development of Ring Imaging Cherenkov Counters for Particle Identification*, vol. 352 of *NATO ASI Series.* Springer US, 1996. doi: 10.1007/978-1-4613-1147-8_29.

[55] M. Albrow, T. Coughlin, and J. Forshaw, *Central Exclusive Particle Production at High Energy Hadron Colliders*, Prog.Part.Nucl.Phys. **65** (2010) 149, arXiv:1006.1289.

[56] CDF Collaboration, T. Aaltonen *et al.*, *Measurement of Ratios of Fragmentation Fractions for Bottom Hadrons in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96-TeV*, Phys.Rev. **D77** (2008) 072003, arXiv:0801.4375.

[57] LHCb Collaboration, R. Aaij *et al.*, *Measurement of b-hadron production fractions in* 7 TeVpp *collisions*, Phys.Rev. **D85** (2012) 032008, arXiv:1111.2357.

[58] R. Barate et al., *A measurement of the semileptonic branching ratio* BR(*b-baryon* $\rightarrow$ $\mathrm{p}l\bar{\nu}X$) *and a study of inclusive* $\pi^{\pm}, \mathrm{K}^{\pm}, (\mathrm{p}, \bar{\mathrm{p}})$ *production in Z decays*, The European Physical Journal C - Particles and Fields **5** (1998), no. 2 205.

[59] E791 Collaboration, E. Aitala *et al.*, *Multidimensional resonance analysis of* $\Lambda_c^+ \rightarrow pK^-\pi^+$, Phys.Lett. **B471** (2000) 449, arXiv:hep-ex/9912003.

[60] CLEO Collaboration, G. Bonvicini *et al.*, *Dalitz plot analysis of the* $D^+ \rightarrow K^-\pi^+\pi^+$ *decay*, Phys.Rev. **D78** (2008) 052001, arXiv:0802.4214.

[61] BaBar Collaboration, P. del Amo Sanchez *et al.*, *Dalitz plot analysis of* $D_s^+ \rightarrow K^+K^-\pi^+$, Phys.Rev. **D83** (2011) 052001, arXiv:1011.4190.

[62] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research A **543** (2005) 577, arXiv:physics/0408124.

[63] LHCb collaboration, R. Aaij *et al.*, *Studies of beauty baryon decays to* D^0ph^- *and* $\Lambda_c^+h^-$ *final states*, arXiv:1311.4823.

[64] A. Berezhnoy and A. Likhoded, *The relative yields of heavy hadrons as function of transverse momentum at LHC experiments*, arXiv:1309.1979.

[65] LHCb collaboration, *Search for the rare decays* $b_s^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ *and* $b_d^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$, . LHCb-CONF-2012-010.

[66] D. Melikhov and N. Nikitin, *Rare radiative leptonic decays* $B_{d,s} \rightarrow l^+l^-\gamma$, Phys. Rev. D **70** (2004) 114028, arXiv:hep-ph/0410146.

[67] D. Gorbunov, *Light sgoldstino: Precision measurements versus collider searches*, Nucl.Phys. **B602** (2001) 213, arXiv:hep-ph/0007325.

[68] S. Demidov and D. Gorbunov, *Flavor violating processes with sgoldstino pair production*, Phys.Rev. **D85** (2012) 077701, arXiv:1112.5230.

[69] LHCb Collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay* $B^0 \rightarrow K^{*0}\mu^+\mu^-$, arXiv:1304.6325.

[70] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to J/\psi \overline{K}^{*0}$ branching fraction and angular amplitudes*, Phys. Rev. D **86** (2012) 071102, arXiv:1208.0738.

[71] Belle collaboration, K. Abe *et al.*, *Measurements of branching fractions and decay amplitudes in $B \to J/\psi K^*$ decays*, Phys.Lett. **B538** (2002) 11, arXiv:hep-ex/0205021.

[72] LHCb collaboration, R. Aaij *et al.*, *Measurement of polarization amplitudes and CP asymmetries in $B^0 \to \phi K^*(892)^0$*, arXiv:1403.2888.

[73] J. A. Hernando Morata, E. Lopez Asamar, D. Martinez Santos, H. Ruiz-Prez, and F. Teubert, *Measurement of trigger efficiencies and biases*, Tech. Rep. LHCb-2008-073. CERN-LHCb-2008-073, CERN, Geneva, March, 2010.

[74] A. L. Read, *Presentation of search results: the $CL_{\rm s}$ technique*, J.Phys. **G28** (2002) 2693.

[75] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl.Instrum.Meth. **A434** (1999) 435, arXiv:hep-ex/9902006.

[76] R. A. et. al., *Search for rare $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ decays*, LHCb-ANA-2012-070 (2012). Analysis note.

[77] LHCb Collaboration, R. Aaij *et al.*, *Implications of LHCb measurements and future prospects*, Eur.Phys.J. **C73** (2013) 2373, arXiv:1208.3355.

[78] LHCb collaboration, P. e. a. Collins, *The LHCb Upgrade*, arXiv:1108.1403.

[79] J. Gaiser, *Charmonium spectroscopy from radiative decays of the J/ψ and ψ', Ph.D. thesis, Appendix F*, . SLAC-R-255.

[80] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances.* PhD thesis, Institute of Nuclear Physics, Krakow, 1986. DESY-F31-86-02.

Appendix A

Glossary of selection variables

The variables listed below are commonly used in LHCb physics analyses to discriminate between signal and background decay channels.

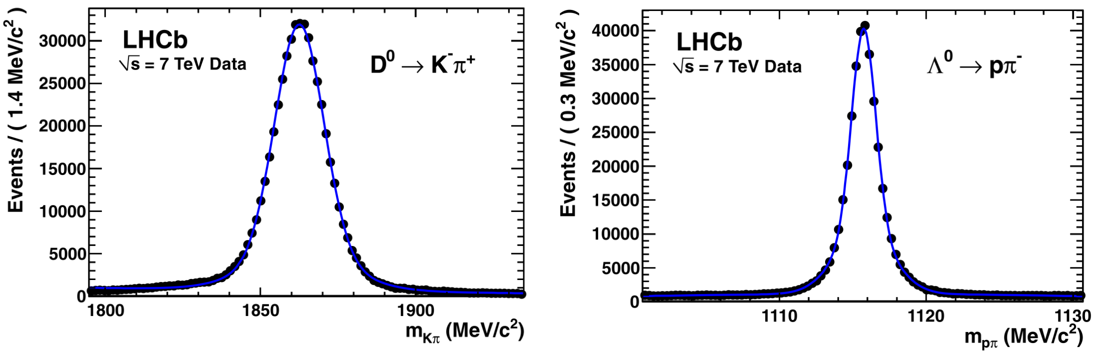
- Impact Parameter (IP): The distance of closest approach between a track and the PV.
- χ^2_{IP}: The consistency of the measured IP of a track with IP=0, given the uncertainties associated with measuring the track and vertex positions.
- Track χ^2: A measure of the consistency of a track trajectory with the tracking system hits from which it was reconstructed.
- p_{T}: Transverse momentum, the projection of a tracks momentum onto the X plane, perpendicular to the *pp* collision axis.
- ISMUON: A boolean variable stating whether a track is consistent with being a muon. Returns 'true' if there are a sufficient number of muon station hits within a defined field of interest around the track trajectory.
- vertex χ^2: The consistency of the position of a vertex with the trajectories of the tracks associated with it.
- vertex DOCA$_{max}$: The maximum value of the distance of closest approach of all track pairs that are associated with a vertex.
- θ_{PV}: Defined for reconstructed particles that have a decay vertex as the angle between the particle's momentum vector and the direction from the PV to the decay vertex.
- flight distance χ^2: The χ^2 value for the displacement between a decay vertex and the primary vertex being consistent with zero.

- $\mathrm{DLL}_{\alpha=p,K,\mu}$: The difference between the α and pion PID hypothesis likelihoods for a track; described in Sect. 3.4.1.

Appendix B

PID calibration datasets

The efficiency of the RICH PID selections in data is measured using calibration datasets containing very clean samples of kaons, pions, protons and muons which are identified by kinematic selections. To avoid bias no information from the RICH system is used. Kaons and pions are obtained from $D^0 \to K^-\pi^+$ decays, which are produced via $D^{*+} \to D^0\pi^+$ decays.



(a) (b)

Figure B.1: Invariant mass plots for (a) the tagged $D^0 \to K^-\pi^+$ and (b) the $\Lambda \to p\pi$ PID calibration samples in 2011, taken from [11].

B.1 Kaons and pions

Kaons and pions are selected from $D^0 \to K^-\pi^+$ decays, where the D^0 is produced via the $D^{*+} \to D^0\pi^+$ decay. These decays are selected kinematically by requiring that the $K\pi$ tracks have a high IP and make a good quality vertex which is displaced from and points back to the primary vertex. A tight cut is applied on the difference between the reconstructed D^0 and D^* masses to further reduce background. D^0 candidates where either the kaon, pion or both have been assigned the wrong K or π mass hypothesis are eliminated by applying a mass veto of 25 MeV/c^2 around the nominal D^0 mass for the reconstructed $K\pi$ invariant mass where the K and/or π hypotheses are swapped. These selections are summarised in Table B.1. The post-selection $K\pi$ invariant mass distribution is shown in Fig. B.1(a).

Table B.1: Selection criteria for the D^* tagged $D^0 \to K^-\pi^+$ decays used for the kaon and pion calibration samples. The pion originating from the D^* is referred to as the 'slow' pion).

Selection variable	Selection criteria
K,π P	$> 2\,\mathrm{GeV}/c$
K,π χ^2_{IP}	> 16
K,π, slow π track χ^2/DOF	< 5
$m_{K\pi}$	$m_{D^0} \pm 75\,\mathrm{MeV}/c^2$
D^0 PT	$> 1.5\,\mathrm{GeV}/c$
D^0 vertex χ^2	< 65
D^0 flight distance χ^2	> 49
D^0 DIRA	> 0.9999
D^0 χ^2_{IP}	> 30
slow π PT	$> 150\,\mathrm{MeV}/c$
D^* PT	$> 2.2\,\mathrm{GeV}/c$
D^* vertex χ^2	< 65
$m_{D^0,slow\pi} - m_{K\pi}$	130 - 155 MeV/c^2
$m_{K,\pi\to K}$	veto $m_{D^0} \pm 25\,\mathrm{MeV}/c^2$
$m_{K\to\pi,\pi}$	veto $m_{D^0} \pm 25\,\mathrm{MeV}/c^2$
$m_{K\to\pi,\pi\to K}$	veto $m_{D^0} \pm 25\,\mathrm{MeV}/c^2$

B.2 Muons

Muons are obtained from $B \to J/\psi(\to \mu^+\mu^-)X$ decays, where one of the two muon candidates is required to satisfy the 'IsMuon = true' condition (described in 3.5) to 'tag' the decay, while the other 'probe' muon is used for the calibration sample. The background for these decays is reduced by applying cuts that require the J/ψ decay vertex to be highly displaced from the primary vertex, its decay products to have a high impact parameter and the dimuon invariant mass to be consistent with the nominal J/ψ mass. These criteria are detailed in Table B.2.

B.3 Protons

Protons are selected from $\Lambda \to p\pi^-$ decays. The decays are selected by requiring the proton and pion to have a high IP and an invariant mass consistent with the Λ. The reconstructed Λ is required to have a long lifetime. Decays of $K^0_S \to \pi^+\pi^-$ where a pion is taken to be a proton are eliminated by applying a mass veto around the K^0_S mass for the $p\pi$ invariant mass where the proton is assigned a pion mass hypothesis. The selection

Table B.2: Selection criteria for the $B \to J/\psi(\to \mu^+\mu^-)X$ decays used for the muon calibration sample.

Selection variable	Selection criteria
J/ψ vertex χ^2	< 50
J/ψ flight distance χ^2	> 225
$m_{\mu^+\mu^-}$	$m_{J/\psi} \pm 150$ MeV/c^2
μ χ^2_{IP}	> 50
μ IP	$> 50\,\mu$m
μ track χ^2/DOF	< 3
μ PT	> 800 MeV/c
μ P	> 3 GeV/c
tag μ IsMuon	true
tag μ PT	> 1.5 GeV/c
tag μ P	> 6 GeV/c
tag μ IP	$> 120\,\mu$m

criteria are listed in Table B.3 and the p, π invariant mass distribution after selection is shown in Fig. B.1 (b).

Table B.3: Selection criteria for the $\Lambda \to p\pi^-$ decays used for the proton calibration samples.

Selection variable	Selection criteria
p, π χ^2_{IP}	> 25
p, π track $\chi 2/DOF$	< 5
$m_{p\pi}$	$m_\Lambda \pm 25$ MeV/c^2
$m_{p\to\pi,\pi}$	veto $m_{K^0_S} \pm 20$ MeV/c^2
Λ vertex z position	< 220 cm
Λ $c * \tau$	> 5 mm
Λ χ^2_τ	< 49
Λ vertex χ^2	< 16

Appendix C

The Crystal Ball PDF

A 'Crystal Ball' (CB) PDF [79, 80] consists of a Gaussian signal peak with a power-law tail either above or below the Gaussian mean. The CB shape is commonly used to fit particle signals in invariant mass spectra. The Gaussian component fits the region around to the mass peak itself and the decay component can fit either low-mass events caused by the decay products radiating photons (referred to as a 'radiative tail'), or high-mass events caused by non-Gaussian detector resolution effects.

The CB PDF is defined as:

$$CB(m, \bar{m}, \sigma, \alpha, n) = N \cdot \begin{cases} e^{-\frac{(m-\bar{m})^2}{2\sigma^2}}, & \text{for } \frac{m-\bar{m}}{\sigma} > \alpha \\ A \cdot \left(B - \frac{m-\bar{m}}{\sigma}\right)^{-n}, & \text{for } \frac{m-\bar{m}}{\sigma} \leq \alpha, \end{cases} \tag{C.1}$$

where $A = \left(\frac{n}{|\alpha|}\right)^n \cdot e^{-\frac{|\alpha|^2}{2}}$, $B = \frac{n}{\alpha} - |\alpha|$ and N is a normalisation factor which ensures the total area under the PDF is equal to 1. In the context of an invariant mass fit, m is the invariant mass, $\bar{m}$ is the mean invariant mass, σ is the mass resolution, $-\alpha$ defines the upper boundary of the CB decay function component, and n defines the size of the decay function.

Appendix D

Measuring the p_T and η dependence of $f_{\Lambda_b^0}/f_d$.

D.0.1 Comparison of data and simulation

The kinematic distributions of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay candidates are shown in Fig. D.1 and Fig. D.2. The variables used in the selection are compared for events in the data and simulated samples. Good agreement is seen for all variables except for the Λ_b^0 p_T, η and χ^2_{IP} distributions, the BDT output and the proton η. These result from the p_T and η dependences of the Λ_b^0 hadronisation rate not being well modeled in the simulation (one of the motivations of the analysis itself is to improve the description of this dependence). Although differences between some variables are seen in the integrated data sample, the binned samples show good agreement for these, as the act of binning in Λ_b^0 p_T or η minimises their effect on other variables via correlations. For example, the BDT output distribution (Fig. D.2(g)) is not modeled well because many of its input variables, such as the Λ_b^0 χ^2_{IP}, are correlated with the Λ_b^0 η (Fig. D.2(g)) and p_T which are not well modeled in the simulation. Figures D.3 and D.4 show the BDT output variable in bins of Λ_b^0 p_T (Fig. D.1(g)) and η (Fig. D.2(k)). Much better agreement between the data and simulation is seen for these events. This is because the binning constrains the Λ_b^0 p_T and η variable ranges, thereby reducing the dependence of the BDT input variables on the Λ_b^0 p_T and η distributions.

Comparisons of simulated and data $\overline{B}^0 \to D^+\pi^-$ decays have been performed in [18], where good agreement is seen. A comparison of another four-body B^0 decay, $B^0 \to J/\psi K^+\pi^-$, is described in Sect. 6.3.1, where good agreement is also seen.

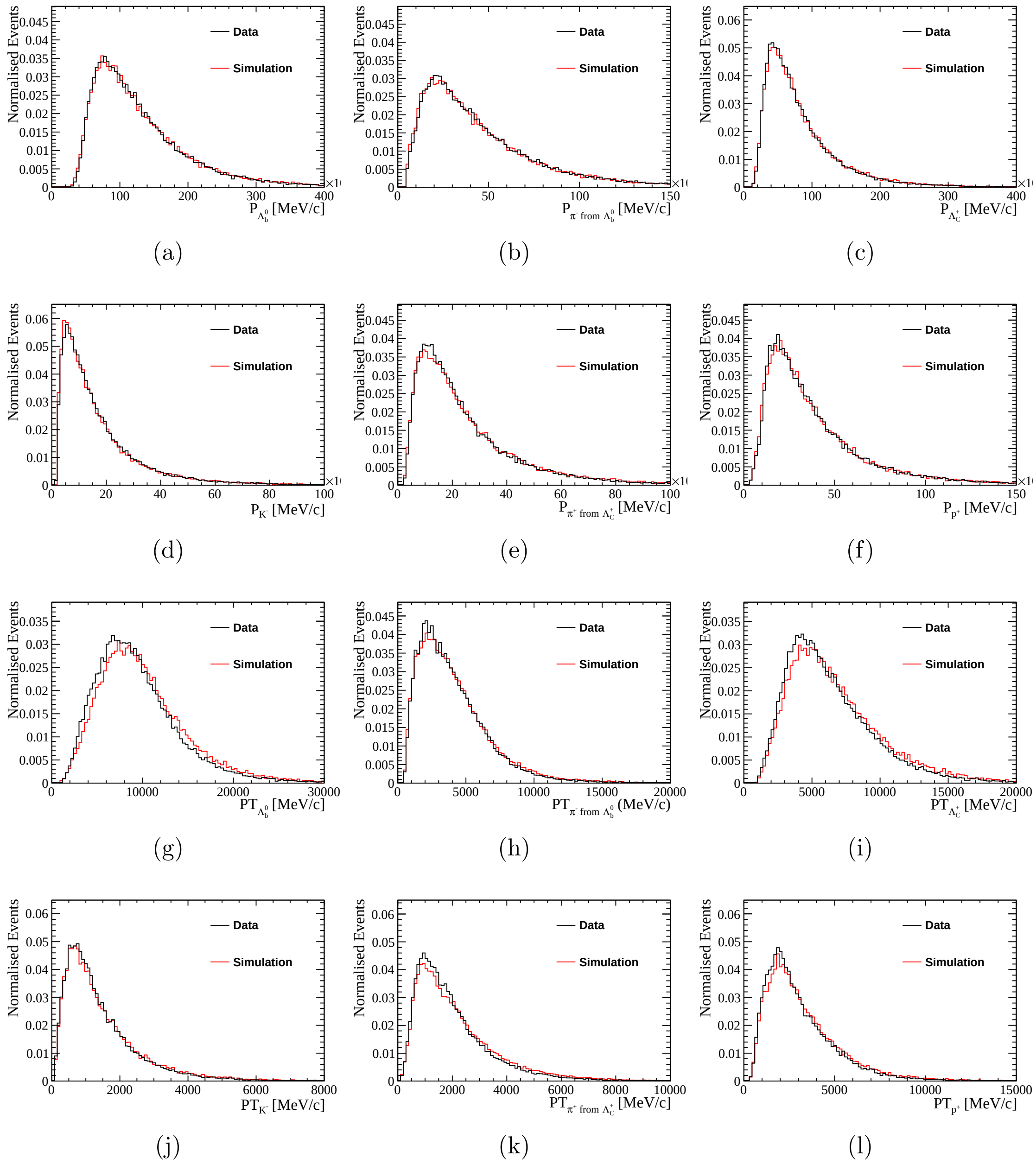


Figure D.1: The P and p_{T} distributions of all final state and intermediate particles of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay, after application of the full selection criteria. Events from simulation (red, grey in B&W) and data (black) are compared. The events in data are selected by requiring the Λ_b^0 invariant mass to be within $\pm 40\,\mathrm{MeV}/c^2$ of the world average Λ_b^0 mass.

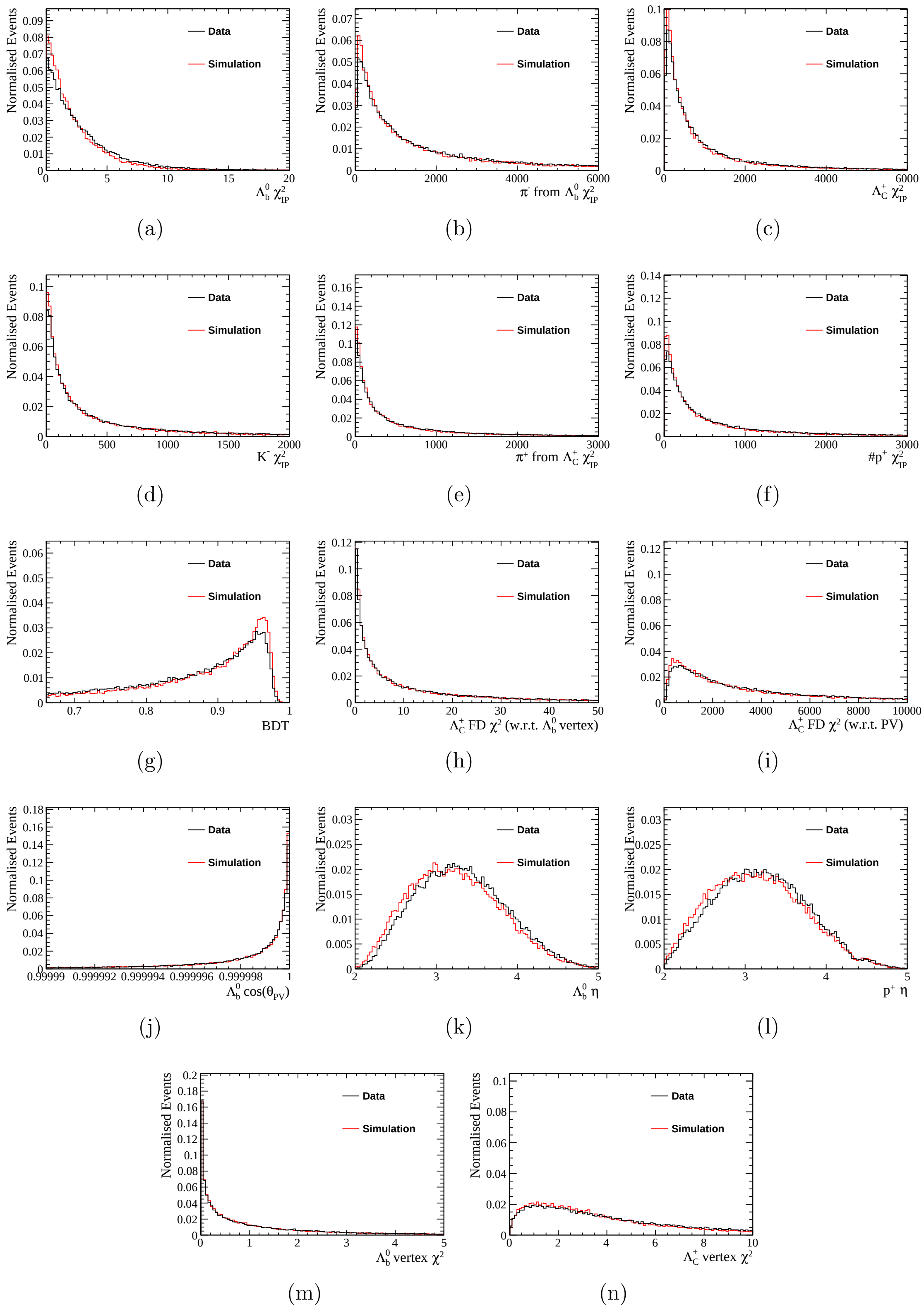
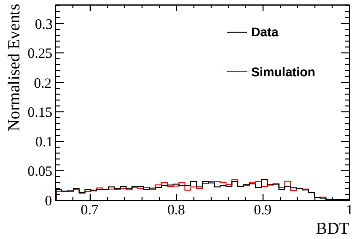
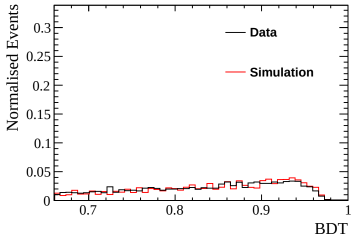


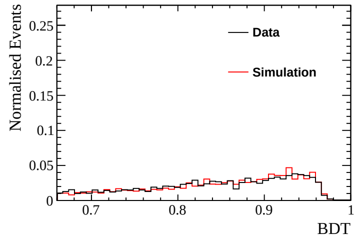
Figure D.2: The χ^2_{IP} distributions of all final state and intermediate particles of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ decay after application of the full selection criteria. Also shown are distributions of: the Λ_c^+ flight distance χ^2, the BDT output, the Λ_b^0 and Λ_c^+ vertex χ^2, and the η of the Λ_b^0 and proton. Events from simulation (red, grey in B&W) and data (black) are compared. The events in data are selected by requiring the Λ_b^0 invariant mass to be within $\pm 40\,\mathrm{MeV}/c^2$ of the world average Λ_b^0 mass.



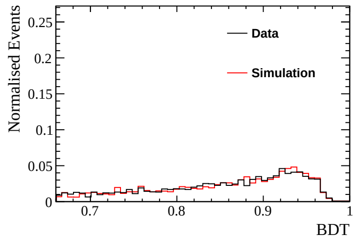
(a) $1.5 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 3.6$



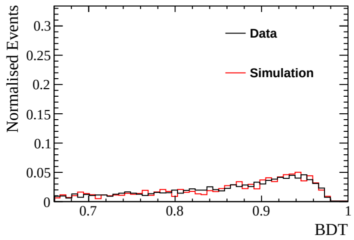
(b) $3.6 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 4.5$



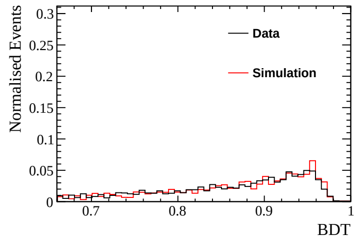
(c) $4.5 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 5.3$



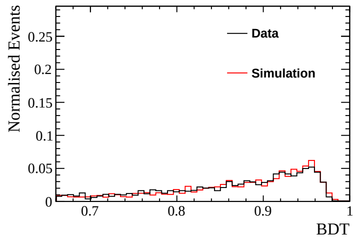
(d) $5.3 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 6.0$



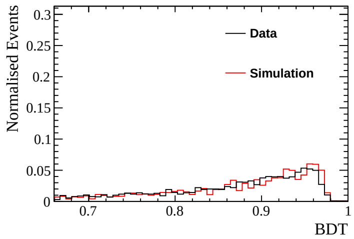
(e) $6.0 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 6.5$



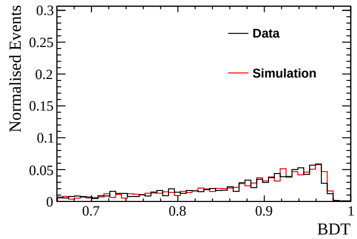
(f) $6.5 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 7.0$



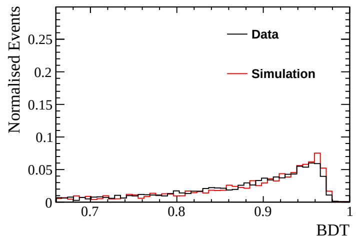
(g) $7.0 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 7.5$



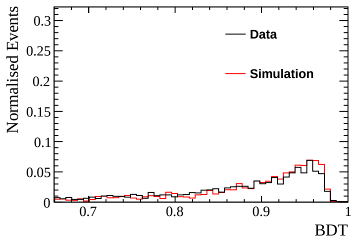
(h) $7.5 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 8.0$



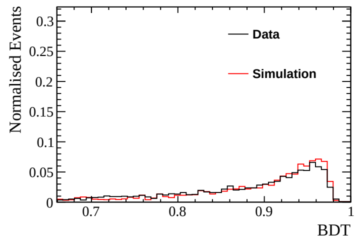
(i) $8.0 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 8.5$



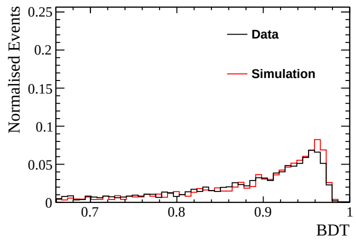
(j) $8.5 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 9.0$



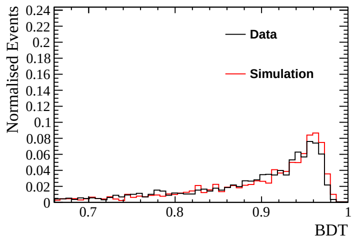
(k) $9.0 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 9.5$



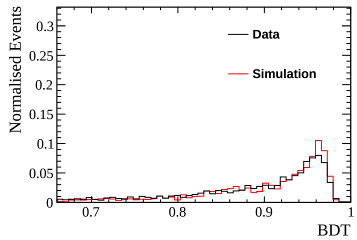
(l) $9.5 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 10.0$



(m) $10.0 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 10.7$



(n) $13.0 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 14.3$

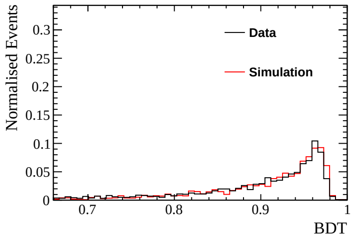


(o) $14.3 < p_{\mathrm{T}}\ (\Lambda_b^0)(\mathrm{GeV}/c) < 16.0$

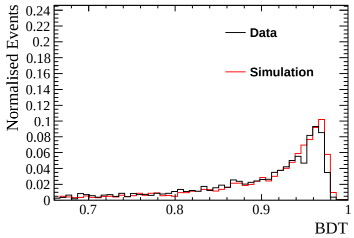
Figure D.3: The BDT output distribution in the binned Λ_b^0 signal samples. Events from simulation (red, grey in B&W) and data (black) are compared. The events in data are selected by requiring the Λ_b^0 invariant mass to be within $\pm 40\,\mathrm{MeV}/c^2$ of the world average Λ_b^0 mass.

D.1 $\Lambda_b^0 \to \Lambda_c^+\pi^-$ fit in bins of Λ_b^0 (p_{T}) and Λ_b^0 (η)

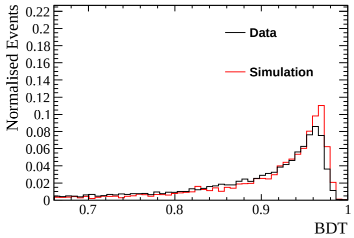
Figures D.5 to D.9 show the fitted $\Lambda_c^+\ \pi^-$ invariant mass distributions of each p_{T}, η bin. The PDF provides a good fit to the data for all bins. As p_{T} increases and η decreases, the yield of the misidentified background from $\overline{B}^0 \to D^+\pi^-$ is seen to increase relative to the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ signal, as more proton candidates fall outside of the kinematic region $\mathbb{R}(p_{\mathrm{T}}, \eta)$ where PID cuts are applied. The equivalent fits to the $\overline{B}^0 \to D^+\pi^-$ mode were performed in [18].



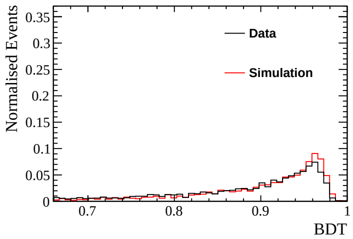
(a) 16.0< p_T (Λ_b^0)(GeV/c) <20.2



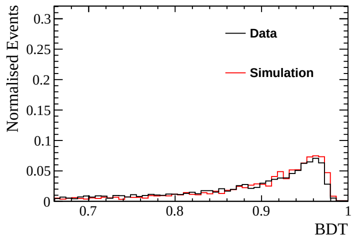
(b) 20.2< p_T (Λ_b^0)(GeV/c) <40.0



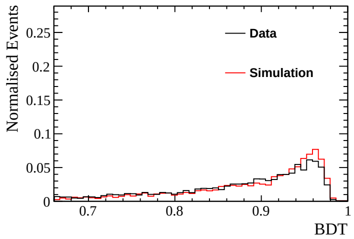
(c) 2.0< $\eta(\Lambda_b^0)$ <2.6



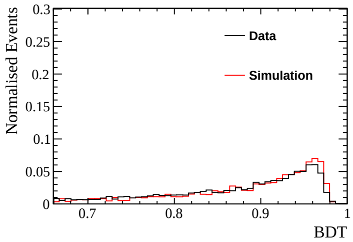
(d) 2.6< $\eta(\Lambda_b^0)$ <2.75



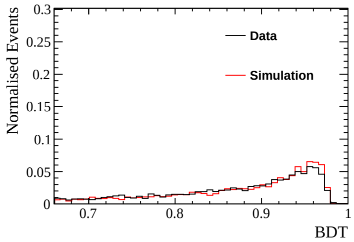
(e) 2.75< $\eta(\Lambda_b^0)$ <2.9



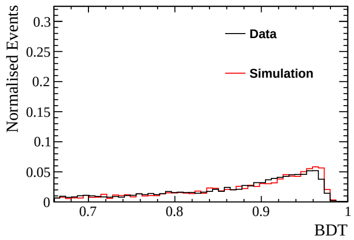
(f) 2.9< $\eta(\Lambda_b^0)$ <3.05



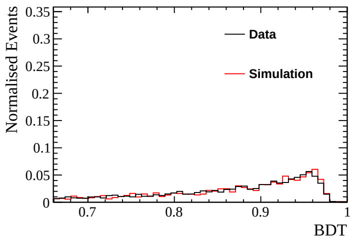
(g) 3.05< $\eta(\Lambda_b^0)$ <3.2



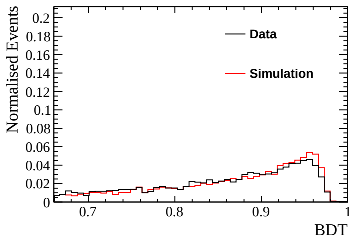
(h) 3.2< $\eta(\Lambda_b^0)$ <3.35



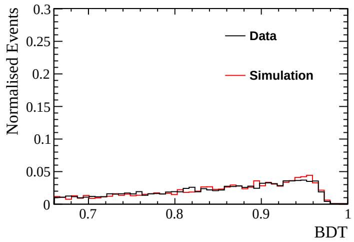
(i) 3.35< $\eta(\Lambda_b^0)$ <3.5



(j) 3.5< $\eta(\Lambda_b^0)$ <3.65

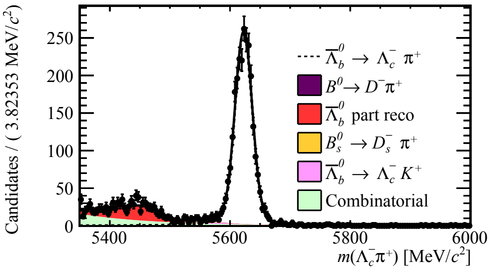


(k) 3.65< $\eta(\Lambda_b^0)$ <4.0

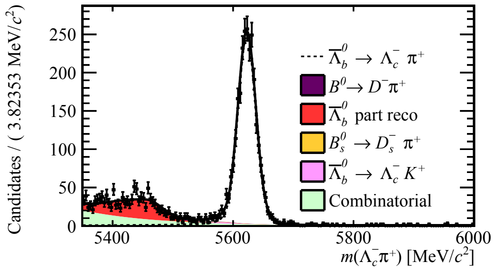


(l) 4.0< $\eta(\Lambda_b^0)$ <5.0

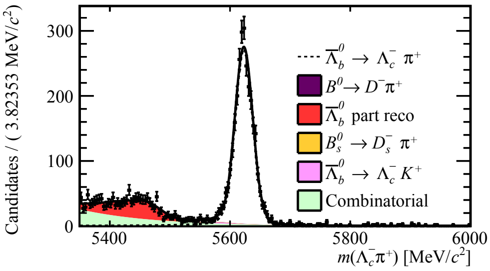
Figure D.4: The BDT output distribution in the binned Λ_b^0 signal samples. Events from simulation (red, grey in B&W) and data (black) are compared. The events in data are selected by requiring the Λ_b^0 invariant mass to be within $\pm 40\,\mathrm{MeV}/c^2$ of the world average Λ_b^0 mass.



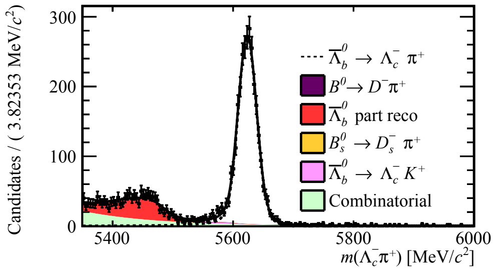
(a) 1.5< p_T (Λ_b^0)(GeV/c) <3.6



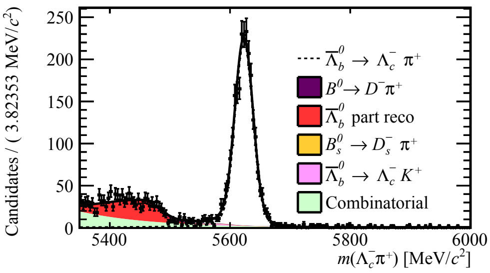
(b) 3.6< p_T (Λ_b^0)(GeV/c) <4.5



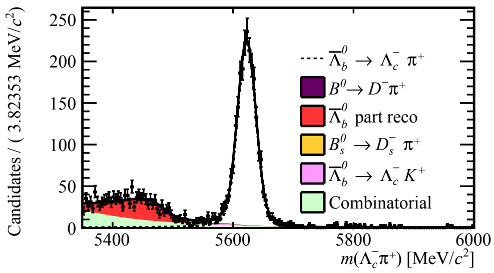
(c) 4.5< p_T (Λ_b^0)(GeV/c) <5.3



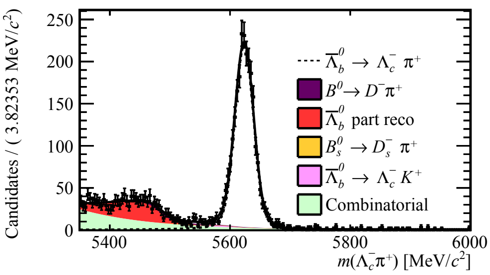
(d) 5.3< p_T (Λ_b^0)(GeV/c) <6.0



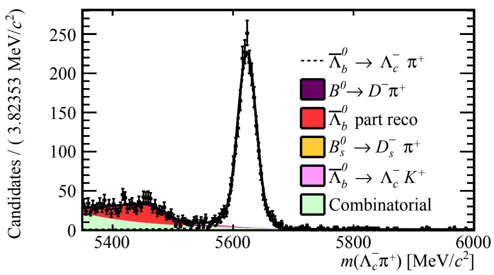
(e) 6.0< p_T (Λ_b^0)(GeV/c) <6.5



(f) 6.5< p_T (Λ_b^0)(GeV/c) <7.0

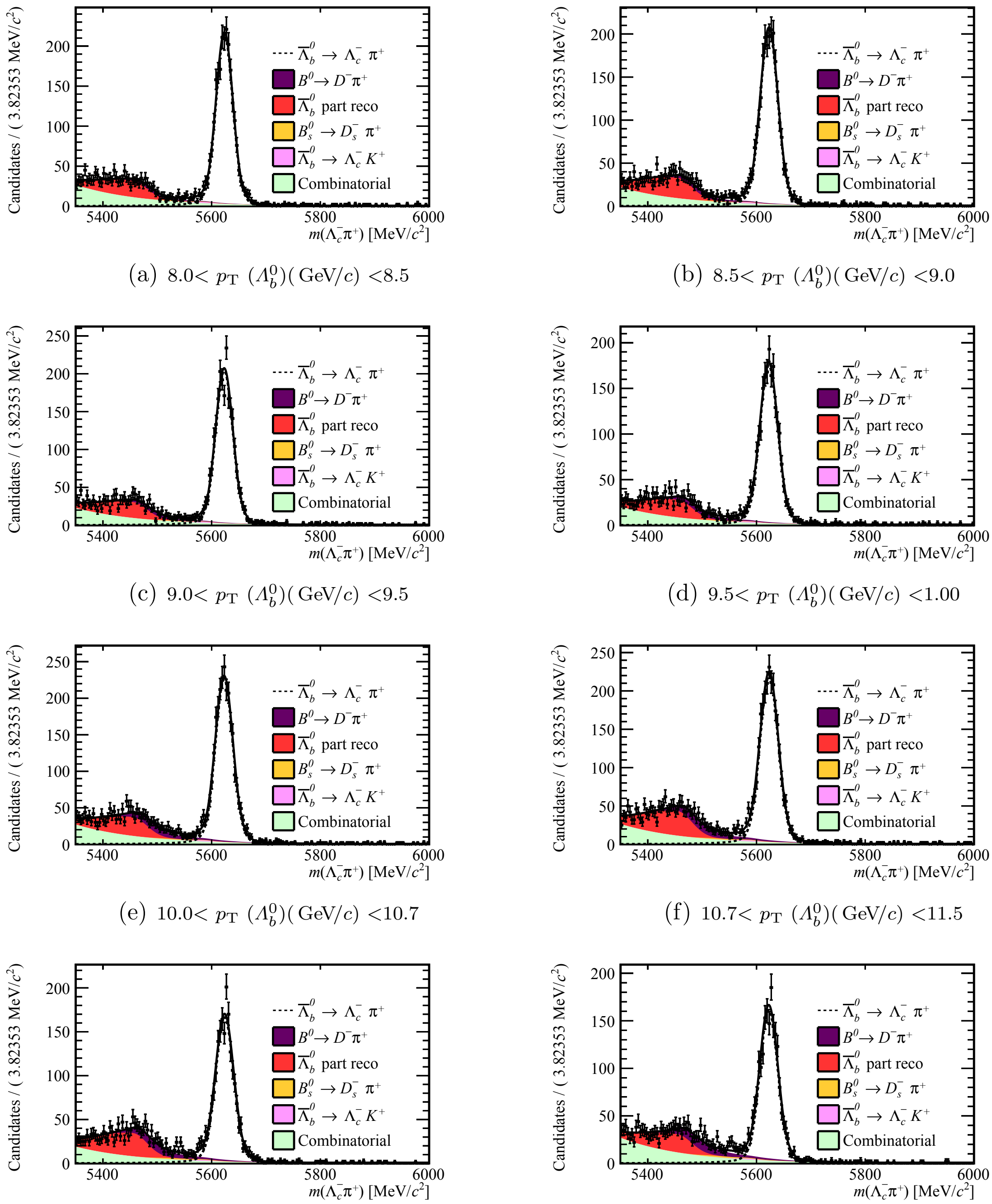


(g) 7.0< p_T (Λ_b^0)(GeV/c) <7.5



(h) 7.5< p_T (Λ_b^0)(GeV/c) <8.0

Figure D.5: The fitted invariant $\Lambda_c^+\ \pi^-$ mass distributions of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in bins of the Λ_b^0 p_T. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labelled in the legend. The 'Λ_b^0 part reco' component refers to the combination of the $\Lambda_b^0 \to \Sigma_c^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\rho^-$ and other partially reconstructed decays.



(a) 8.0< p_T (Λ_b^0)(GeV/c) <8.5

(b) 8.5< p_T (Λ_b^0)(GeV/c) <9.0

(c) 9.0< p_T (Λ_b^0)(GeV/c) <9.5

(d) 9.5< p_T (Λ_b^0)(GeV/c) <1.00

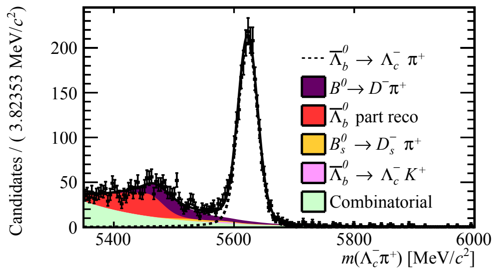
(e) 10.0< p_T (Λ_b^0)(GeV/c) <10.7

(f) 10.7< p_T (Λ_b^0)(GeV/c) <11.5

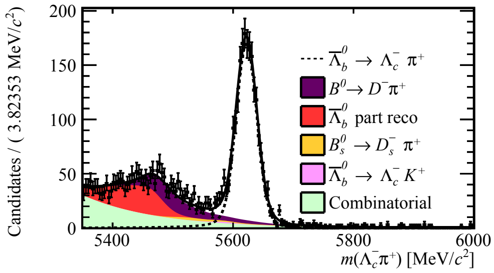
(g) 11.5< p_T (Λ_b^0)(GeV/c) <12.2

(h) 12.2< p_T (Λ_b^0)(GeV/c) <13.0

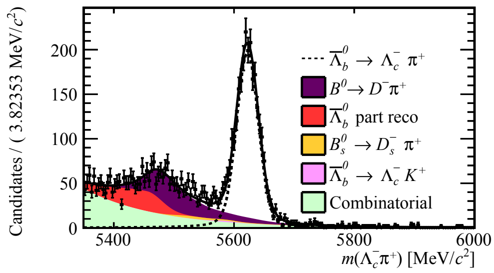
Figure D.6: The fitted invariant $\Lambda_c^+\ \pi^-$ mass distributions of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in bins of the Λ_b^0 p_T. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labelled in the legend. The 'Λ_b^0 part reco' component refers to the combination of the $\Lambda_b^0 \to \Sigma_c^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\rho^-$ and other partially reconstructed decays.



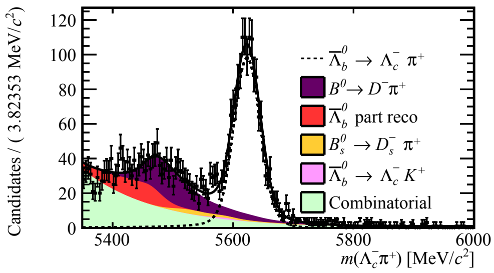
(a) 13.0< p_{T} (Λ_b^0)(GeV/c) <14.3



(b) 14.3< p_{T} (Λ_b^0)(GeV/c) <16.0

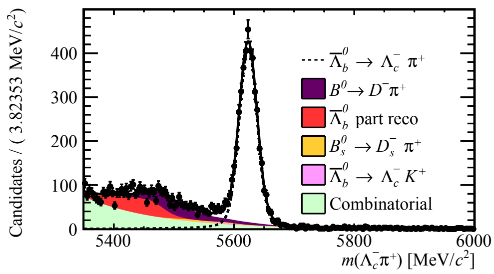


(c) 16.0< p_{T} (Λ_b^0)(GeV/c) <20.2

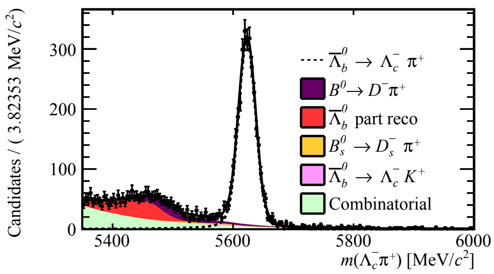


(d) 20.2< p_{T} (Λ_b^0)(GeV/c) <40.0

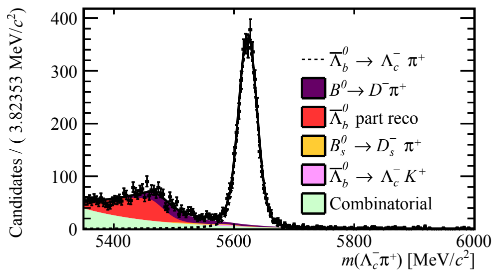
Figure D.7: The fitted invariant $\Lambda_c^+\ \pi^-$ mass distributions of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in bins of the Λ_b^0 p_{T}. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labelled in the legend. The 'Λ_b^0 part reco' component refers to the combination of the $\Lambda_b^0 \to \Sigma_c^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\rho^-$ and other partially reconstructed decays.



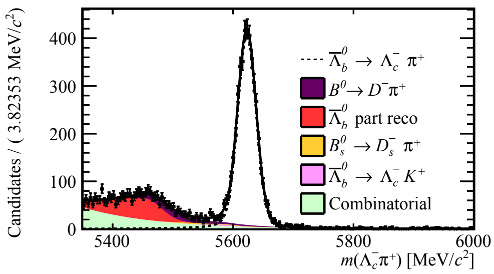
(a) 2.0< $\eta(\Lambda_b^0)$ <2.6



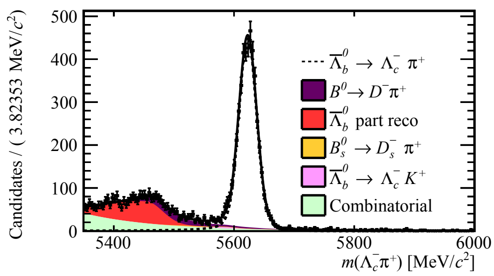
(b) 2.6< $\eta(\Lambda_b^0)$ <2.75



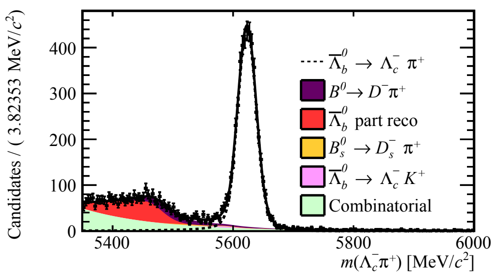
(c) 2.75< $\eta(\Lambda_b^0)$ <2.9



(d) 2.9< $\eta(\Lambda_b^0)$ <3.05

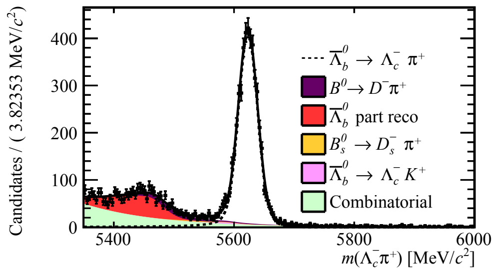


(e) 3.05< $\eta(\Lambda_b^0)$ <3.2

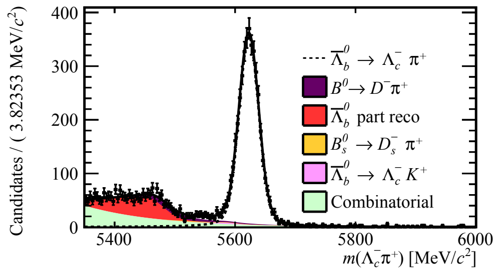


(f) 3.2< $\eta(\Lambda_b^0)$ <3.35

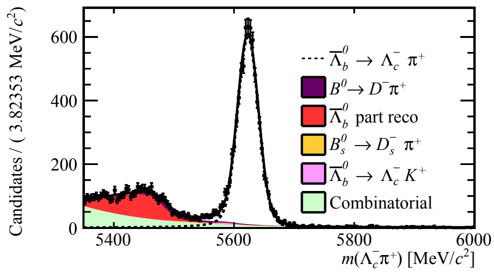
Figure D.8: The fitted invariant $\Lambda_c^+\ \pi^-$ mass distributions of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in bins of the Λ_b^0 η. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labelled in the legend. The 'Λ_b^0 part reco' component refers to the combination of the $\Lambda_b^0 \to \Sigma_c^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\rho^-$ and other partially reconstructed decays.



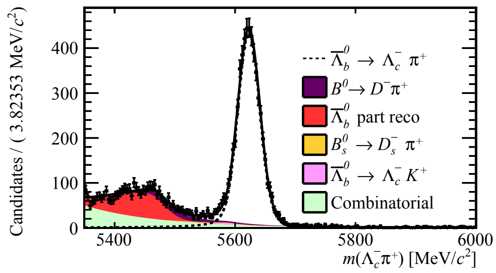
(a) $3.35 < \eta(\Lambda_b^0) < 3.5$



(b) $3.5 < \eta(\Lambda_b^0) < 3.65$



(c) $3.65 < \eta(\Lambda_b^0) < 4.0$



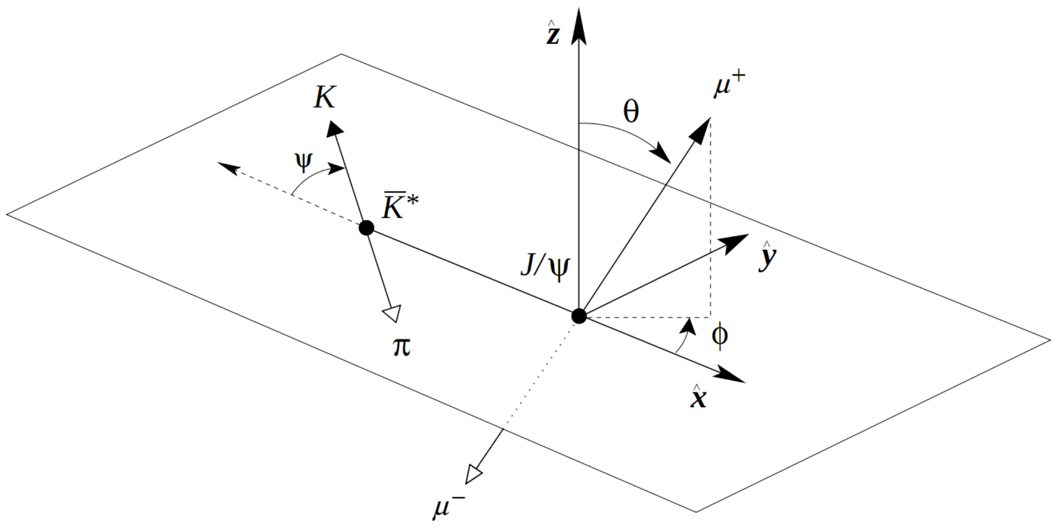
(d) $4.0 < \eta(\Lambda_b^0) < 5.0$

Figure D.9: The fitted invariant $\Lambda_c^+\ \pi^-$ mass distributions of selected $\Lambda_b^0 \to \Lambda_c^+\pi^-$ events in bins of the Λ_b^0 η. The dashed black line shows the signal component of the fit. The shaded regions show the different background fit components, each of which is labelled in the legend. The 'Λ_b^0 part reco' component refers to the combination of the $\Lambda_b^0 \to \Sigma_c^+\pi^-$, $\Lambda_b^0 \to \Lambda_c^+\rho^-$ and other partially reconstructed decays.

Appendix E

$B^0 \rightarrow J/\psi K^{*0}$ s-wave analysis

The analysis described in this chapter was not performed by the author. It is described for reference and completeness only.



(a)

Figure E.1: An illustration of the angles defined for the s-wave analysis.

The non-resonant s-wave component of the $B^0 \rightarrow J/\psi\ K^+\ \pi^-$ decay is measured by fitting the distribution of the three angles shown in Fig. E.1, where ψ is the angle between the K^+ in the rest frame of the $K^+\ \pi^-$ system and the $K^+\ \pi^-$ in the rest frame of the B^0. The angles are related to the differential decay rate by [70]:

$$\begin{aligned}
\frac{\mathrm{d}^3\Gamma}{\mathrm{d}\Omega} &\propto 2|A_0|^2\cos^2\psi(1-\sin^2\theta\cos^2\varphi) \\
&+ |A_\parallel|^2\sin^2\psi(1-\sin^2\theta\sin^2\varphi) \\
&+ |A_\perp|^2\sin^2\psi\sin^2\theta \\
&+ \frac{1}{\sqrt{2}}|A_0||A_\parallel|\cos(\delta_\parallel-\delta_0)\sin 2\psi\sin^2\theta\sin 2\varphi \qquad \text{(E.1)} \\
&+ \frac{2}{3}|A_\mathrm{S}|^2\left[1-\sin^2\theta\cos^2\varphi\right] \\
&+ \frac{4\sqrt{3}}{3}|A_0||A_\mathrm{S}|\cos(\delta_\mathrm{S}-\delta_0)\cos\psi\left[1-\sin^2\theta\cos^2\varphi\right] \\
&+ \frac{\sqrt{6}}{3}|A_\parallel||A_\mathrm{S}|\cos(\delta_\parallel-\delta_\mathrm{S})\sin\psi\sin^2\theta\sin 2\varphi, \\
&\equiv PDF(\Omega)_{phys} \qquad \text{(E.2)}
\end{aligned}$$

2769 where $|A_0|$,$|A_\parallel|$ and $|A_\perp|$ are moduli of the decay amplitudes of the longitudinal
2770 and transversely polarised J/ψ and K^{*0} vector mesons, $|A_S|$ is the amplitude of the
2771 s-wave, δ are the phases of the amplitudes, defined with respect to $\delta_0 = 0$ and $\mathrm{d}\Omega \equiv$
2772 $\mathrm{d}\cos\psi\,\mathrm{d}\cos\theta\,\mathrm{d}\varphi$.

2773 E.1 Angular acceptance

2774 The dependency of the angular distribution on the reconstruction, selection and trigger
2775 efficiencies is referred to as the 'angular acceptance'. It needs to be determined in order to
2776 extract the pure amplitudes in (E.1). The acceptance is measured by fitting the angular
2777 distributions of simulated $B^0 \to J/\psi K^{*0}$ events after application of the selection criteria
2778 detailed in Table 6.1, including a $\pm 25\,\mathrm{MeV}/c^2$ $K^+\,\pi^-\,\mu^+\,\mu^-$ invariant mass window around
2779 the nominal B^0 mass with the following PDF:

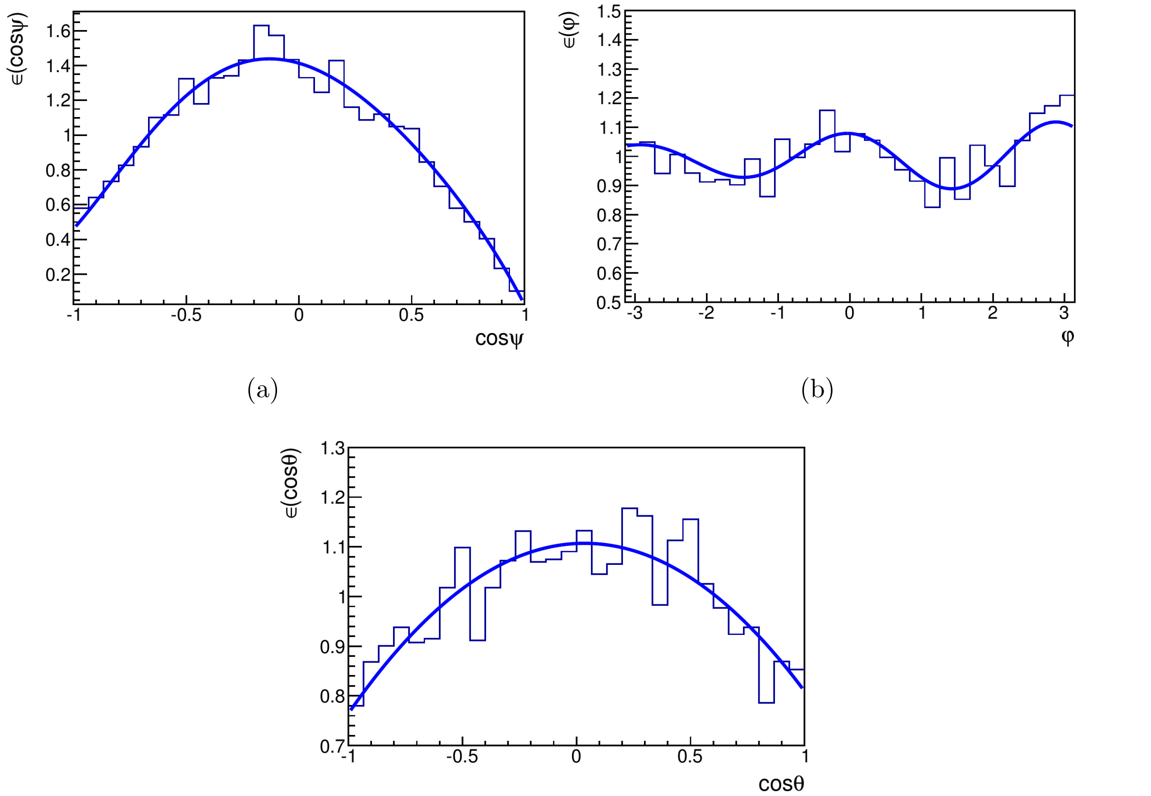
$$PDF(\Omega)_{tot} = PDF(\Omega)_{phys} \times PDF(\Omega)_{acc} \qquad \text{(E.3)}$$

2780 where the parameters of the physical PDF shown in (E.1) are fixed to the values in [1],
2781 which are used to generate the simulated events:

$$\begin{aligned}
|A_\perp|^2 &= 0.1601 \\
|A_\parallel|^2 &= 0.2397 \\
\delta_\parallel - \delta_0 &= 2.501 \\
|A_S| &= 0 \qquad \text{(E.4)}
\end{aligned}$$

The acceptance function $PDF(\Omega)_{acc}$ is parameterised as:

$$\begin{aligned} PDF(\Omega)_{acc} &= PDF(\psi)_{acc} \times PDF(\theta)_{acc} \times PDF(\varphi)_{acc}, \\ PDF(\psi)_{acc} &= 1 + \sum_{n=1}^{5} c_n^{\psi} \cos^n \psi \qquad \text{(E.5)} \\ PDF(\theta)_{acc} &= 1 + c_\theta \cos^2 \theta \qquad \text{(E.6)} \\ PDF(\varphi)_{acc} &= 1 + (c_1^{\varphi} + c_4^{\varphi}\varphi) \cos(c_2^{\varphi}\varphi + c_3^{\varphi}) \qquad \text{(E.7)} \end{aligned}$$



(a) (b)

(c)

Figure E.2: The angular distributions of simulated $B^0 \to J/\psi K^{*0}$ events. Projections of the fitted PDF in (E.3) with the physical parameters fixed to those in (E.4) are shown.

The simulated angular distributions and the resultant fit are shown in Fig. E.2. The fitted acceptance variables are shown in Table E.1.

E.2 Background angular distribution

Background events under the B^0 mass peak can distort the angular distributions. A sample of pure background events with invariant $K^+ \pi^- \mu^+ \mu^-$ mass 5400-5756 MeV/c^2 is taken from selected data $B^0 \to J/\psi K^{*0}$ events. The angular distributions of these events

Table E.1: Parameters of the angular acceptance obtained from the acceptance fit to simulated $B^0 \to J/\psi K^{*0}$ events.

c_1^ψ	-0.265 ± 0.059
c_2^ψ	-0.952 ± 0.067
c_3^ψ	0.33 ± 0.19
c_4^ψ	0.0135 ± 0.0083
c_5^ψ	-0.115 ± 0.029
c_1^θ	0.020 ± 0.016
c_2^θ	-0.289 ± 0.028
c_1^ϕ	0.086 ± 0.016
c_2^ϕ	2.15 ± 0.13
c_3^ϕ	0.14 ± 0.21
c_4^ϕ	$(1.35 \pm 0.83) \times 10^{-2}$

Table E.2: Values of the variables of the background PDF (E.8), obtained from a fit to background $B^0 \to J/\psi K^{*0}$ events in data.

Parameter	Value
k_1	-0.086 ± 0.077
k_2	$(0.4 \pm 2.3) \times 10^{-2}$
k_3	0.019 ± 0.012

are shown in Fig. E.3. The ψ and θ distributions are well described by $PDF(\psi)_{acc}$ and $PDF(\theta)_{acc}$, using the acceptance PDFs obtained from the simulation. The background φ distribution does not well reproduce the $PDF(\varphi)_{acc}$ so it is fitted a third order polynomial PDF. The total background PDF is:

$$PDF(\Omega)_{bkg} = PDF(\psi)_{acc} \times PDF(\theta)_{acc} \times \left(1 + \sum_{n=1}^{3} k_n \varphi^n\right) \tag{E.8}$$

The values of the fitted φ coefficients k are shown in Table E.2.

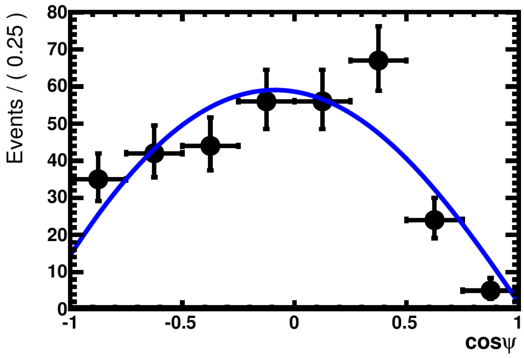
E.3 Fit to the data and extraction of κ

The amplitudes A and phases δ in (E.1) are extracted from the data by fitting the ψ, θ and φ variables of selected $B^0 \to J/\psi K^{*0}$ events within a $M_B^0 \pm 25\,\mathrm{MeV}/c^2$ K^+ π^- μ^+ μ^- invariant mass window. The fit PDF is:

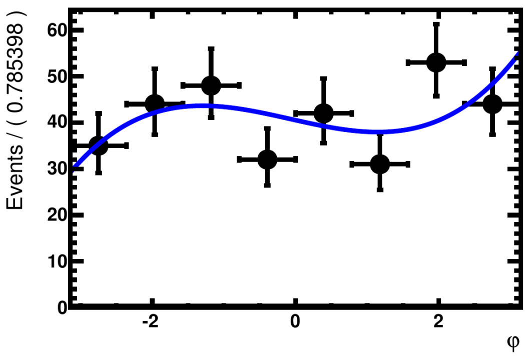
$$PDF(\Omega)_{data} = PDF(\Omega)_{phys} \times PDF(\Omega)_{acc} + PDF(\Omega)_{bkg} \tag{E.9}$$

where the parameters of $PDF(\Omega)_{phys}$ are left free in the fit, those of $PDF(\Omega)_{acc}$ and $PDF(\Omega)_{bkg}$ are fixed to the values obtained in Table E.1 and Table E.2, respectively.

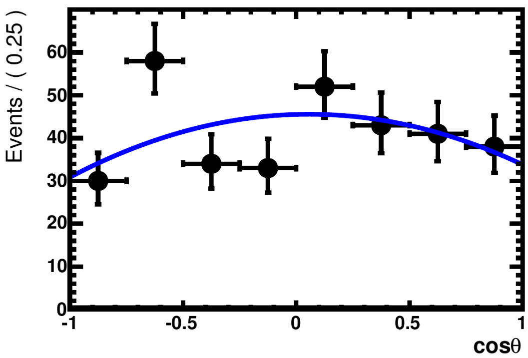
The signal data angular distributions are shown in Fig. E.4, along with the fitted



(a)



(b)



(c)

Figure E.3: The angular distributions of background data $B^0 \to J/\psi K^{*0}$ events, projections of the fitted PDF in (E.3) are shown.

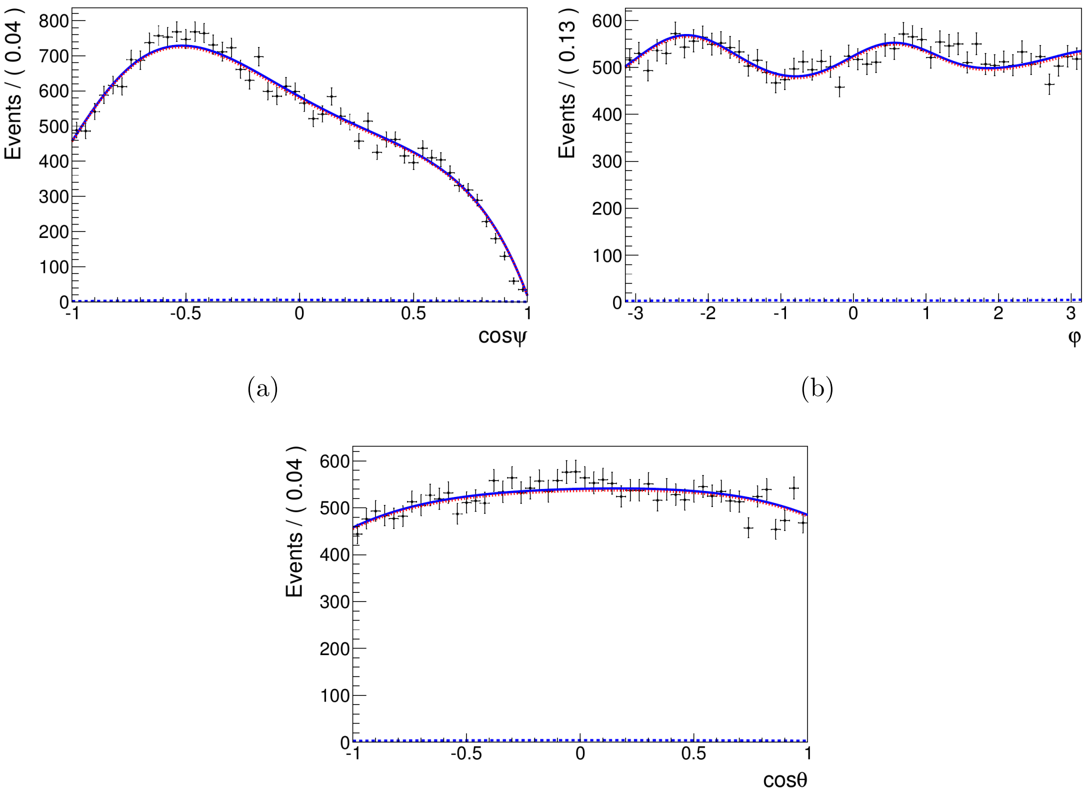
PDF described in (E.9). The fitted values for the amplitudes and phases in (E.1) are shown in Table E.3. The s-wave amplitude squared is found to be $|A_s|^2 = 0.054 \pm 0.005$, compared to 0.037 ± 0.010 in [70]. The two values are consistent within 2σ.

The impact of the s-wave on both the $B^0 \to J/\psi K^{*0}$ yield and the efficiencies calculated in (6.1) is corrected for using the factor κ:

$$\kappa = \frac{\epsilon_{data}}{\epsilon_{sim}} \times \frac{1}{f_{K^{*0}}} = \frac{\int_\Omega PDF(\Omega, |A_S| = 0)_{phys,data} PDF(\Omega)_{acc} d\Omega}{\int_\Omega PDF(\Omega)_{phys,sim} PDF(\Omega)_{acc} d\Omega} \times \frac{1}{1 - |A_s|^2} = 1.09 \pm 0.01 \tag{E.10}$$

Table E.3: Results of the angular fit to selected signal $B^0 \to J/\psi K^{*0}$ events in data.

Parameter	Fitted value
$\|A_0\|^2$	0.504 ± 0.007
$\|A_\parallel\|^2$	0.248 ± 0.006
$\|A_\perp\|^2$	0.194 ± 0.005
$\|A_s\|^2$	0.054 ± 0.005
$\delta_\parallel$	2.934 ± 0.039
δ_s	$2.08^{+0.05}_{-0.04}$



(a) (b)

(c)

Figure E.4: The angular distributions of selected signal $B^0 \to J/\psi K^{*0}$ events in data. The fitted background (dashed blue, dark grey in B&W), signal (dotted red, light grey in B&W) and combined (solid blue, dark grey in B&W) PDFs, defined in (E.9), are shown.

where $f_{K^{*0}} = 1 - |A_s|^2$ is the fraction of p-wave $B^0 \to J/\psi K^+\pi^-$ events in the data, $\frac{\epsilon_{data}}{\epsilon_{sim}}$ corrects the data-simulation differences of the amplitudes and phases. The inputs to the physical angular PDFs $PDF(\Omega, |A_S| = 0)_{phys,data}$ and $PDF(\Omega)_{phys,sim}$ are taken from the fitted data and simulation, respectively. The s-wave amplitude is set to zero for both as the s-wave is removed in the $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ analysis in Sect. 6.6.

Appendix F

$B_s^0 \to SP$ efficiency scan

Table F.1 shows the variation in ϵ_{tot} when one $\mu^+\mu^-$ pair is required to have an invariant mass of $< 950\,\text{MeV}/c^2$ and the mass of the other is varied across 8 bins which avoid the ϕ and J/ψ mass vetoes. This gives an approximate measure of how ϵ_{tot} varies in the MSSM model where the mass of S is varied across its allowed phase-space. The efficiency in the $2363 - 2682$ MeV/c^2 dimuon mass bin is 0.375 ± 0.019 %, which is consistent with 0.366 ± 0.003 % for the $B_s^0 \to SP$ simulation sample with $m_S = 2.5\,\text{GeV}/c^2$. When these efficiencies are input to (6.1) the 95% confidence level branching fraction limits on $B_s^0 \to SP$ are seen to vary by $^{+6}_{-23}$% with respect to the simulated sample with $m_S = 2.5\,\text{GeV}/c^2$.

Table F.1: The variation of the $B_s^0 \to \mu^+\mu^-\mu^+\mu^-$ phase-space simulation sample efficiency ϵ_{tot} when one $\mu^+\mu^-$ pair is required to have an invariant mass of $< 950\,\text{MeV}/c^2$ and the mass of the other pair is varied to the values shown. The binning scheme avoids the ϕ and J/ψ mass vetoes.

$M_{\mu^+\mu^-}$ (MeV/c^2)	ϵ_{tot} (%)
<950	0.434 ± 0.034
1090-1408	0.389 ± 0.024
1408-1727	0.442 ± 0.023
1727-2045	0.383 ± 0.020
2045-2363	0.374 ± 0.019
2363-2682	0.375 ± 0.019
2682-3000	0.325 ± 0.017
>3200	0.381 ± 0.011

Appendix G

List of Acronyms

ATLAS A Toroidal LHC ApparatuS

BDT Boosted Decision Tree

BSM Beyond the Standard Model

CDF Collider Detector at Fermilab

CERN European Organization for Nuclear Research

CL confidence level

CMS Compact Muon Solenoid

CB Crystal Ball Function

CKM Cabibbo-Kobayashi-Maskawa

DCB Double Crystal Ball Function

ECAL Electromagnetic Calorimeter

EM Electromagnetic

EW Electroweak

FCNC Flavour Changing Neutral Current

GEM Gas Electron Multiplier

HCAL Hadron Calorimeter

HEP High Energy Physics

HFAG Heavy Flavour Averaging Group

HPD Hybrid Photon Detector

IP Impact Parameter

LEP Large Electron-Positron Collider

LHC Large Hadron Collider

LHCb Large Hadron Collider beauty Experiment

MC Monte Carlo

MSSM Minimal Supersymmetric Standard Model

MWPCs Multi-Wire Proportional Chambers

PDF Probability Density Function

PID Particle Identification

PS Proton Synchrotron

PSD Pre-Shower Detector

PSB Proton Synchrotron Booster

PV Primary Vertex

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QED Quantum Field Theory

RICH Ring Imaging Cherenkov Detector

RF Radio Frequency

SPD Scintillating Pad Detector

SPS Super Proton Synchrotron

SM Standard Model

SUSY Supersymmetry

SV Secondary Vertex

TIS Trigger Independent of Signal

TOS Trigger on Signal

TT Tracker Turicensis

VEV Vacuum Expectation Value

VELO Vertex Locator