A study of $B \to DK$ and D^0 production using $D^0 \to K^+\pi^-\pi^+\pi^-$ decays at LHCb.

Philip Hunt

Keble College, Oxford

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term, 2012

Abstract

A precision measurement of the CKM angle γ from tree-level processes is one of the principal goals of the LHCb experiment. Two important channels for the extraction of γ are $B^{\pm} \to D^0 K^{\pm}$ and $B^{\pm} \to \overline{D}{}^0 K^{\pm}$; a first study of these modes is presented using an integrated luminosity of $\mathcal{L}_{\text{int}} = 35.6 \pm 3.6$ pb^{-1} at a pp centre-of-mass collision energy of $\sqrt{s} = 7\,\text{TeV}$. A measurement is made of the ratio of branching fractions $\mathcal{B}_{DK}/\mathcal{B}_{D\pi} = \Gamma(B^{\pm} \to DK^{\pm})/\Gamma(B^{\pm} \to D\pi^{\pm}) = 0.057 \pm 0.007 \pm 0.004$ using $B^{\pm} \to \left(K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}\right)_D K^{\pm}$ and $B^{\pm} \to \left(K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}\right)_D \pi^{\pm}$ four-body D decay modes. This has an uncertainty competitive with the previous world average of $(7.6 \pm 0.6)\%$. A measurement is also made of the rate asymmetry between $B^- \to \left(K^-\pi^+\pi^-\pi^+\right)_D K^-$ and $B^+ \to \left(K^+\pi^-\pi^+\pi^-\right)_D K^+$ decays $A^{K3\pi}_{fav} = \frac{\Gamma\left(B^- \to \left(K^-\pi^+\pi^-\pi^+\right)_D K^-\right) - \Gamma\left(B^+ \to \left(K^+\pi^-\pi^+\pi^-\right)_D K^+\right)}{\Gamma\left(B^- \to \left(K^-\pi^+\pi^-\pi^+\right)_D K^-\right) + \Gamma\left(B^+ \to \left(K^+\pi^-\pi^+\pi^-\right)_D K^+\right)} = 0.01 \pm 0.11 \pm 0.06$. As expected, this result is consistent with zero within statistical and systematic uncertainties.

The $D^0/\overline{D}{}^0$ production cross-section in bins of transverse momentum and rapidity from prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ decays (and its charge conjugate) is presented, using $\mathcal{L}_{\text{int}} = 14.96 \pm 0.52$ nb^{-1} of $\sqrt{s} = 7\,\text{TeV}$ LHCb pp collision data. This measurement is important for testing predictions of QCD theoretical models, and determining LHCb's sensitivity to measurements of *CP* violation, mixing measurements and rare decays of charmed hadrons. The results are compared to predictions from two different theoretical models and from the default LHCb tuning of the Pythia Monte Carlo event generator, and the results shown to be in good agreement. The cross-section results are also compared to an independent LHCb measurement in the $D^0 \to K^-\pi^+$ decay channel.

LHCb analyses rely on the ability to identify kaons and pions with a high efficiency and low mis-identification rate, achieved by two Ring Imaging Cherenkov (RICH) detectors. To ensure optimal performance of the RICH detectors, the time alignment of the Level-0 (L0) front-end electronics modules has been optimised using a combination of a pulsed laser system installed in the LHCb cavern and pp collision data. After the time-alignment procedure, the L0 modules have been time-aligned to within approximately $\pm 1\,$ns across both detectors.

Dedicated to the memory of Jean Armstrong.

Acknowledgements

Firstly, I would like to thank the Science and Technology Facilities Council (STFC) for funding my four-year studentship, including my two year long-term attachment at CERN. I would also like to thank Keble College, the Institute of Physics (IoP) and the Oxford physics sub-department for providing crucial funding that allowed me to attend the Kruger2010 conference.

I owe a great many thanks to Neville Harnew, my supervisor, for his continual support and encouragement throughout my DPhil. In particular, I have benefited greatly from his thorough proofreading of this thesis.

Thanks to all the members of the LHCb Oxford group who contributed to the $B \to Dh$ analysis: Malcolm John, Andrew Powell, Paolo Gandini and Chris Thomas. In particular, I would like to thank my contemporary, Chris Thomas, for making my time at both CERN and Oxford so enjoyable. I am also extremely grateful to Andrew Powell for all the help and advice he has given me during my DPhil, not least during the development of PID calibration method used by the prompt charm production analysis, and for helping proofread several sections of my thesis.

I owe many thanks to the Bristol group, in particular Jonas Rademacker, Paras Naik and Tom Hampson, for their work on the measurement of the prompt D^0 production cross-section from $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. I'd also like to thank the numerous other people who contributed to the prompt charm production analysis, particularly Patrick Spradlin for his great work as head of the open-charm production task force.

Thanks to Sean Brisbane and Stig Topp-Jorgensen for their excellent work on developing the RICH L0 PVSS projects and RICH L0 DIM servers respectively. I would also like to thank Sean Brisbane, Nick Styles and Conor Fitzpatrick for their work on the RICH time alignment.

Thanks to Carmelo d'Ambrosio, who was the RICH coordinator during a significant proportion of my stay at CERN, and who spend an inordinate amount of time in the LHCb control room in the run-up to the first pp collisions. Thanks also to Ken Wylie, Sean Brisbane, Stig Topp-Jorgensen, Phil Sullivan and Marco Adinolfi for imparting their knowledge and experience of the RICH L0 electronics and control systems and to Tom Blake for sharing his general knowledge of the RICH detectors.

Thanks to all the various hardware and software experts that worked so hard to ensure the proper working of the RICH detectors during my time at CERN: Tito Bellunato, Christophe Frei, Clara Gaspar, Thierry Gys, Chris Jones, Ulrich Kerzel, Antonis Papenestis, Alessandro Petrolini, Didier Piedigrossi, Gareth Rogers, Olav Ullaland, Steve Wotton, Philip Xing, Ross Young and many others.

Many thanks to all the members of the LHCb Oxford group for their hard work and dedication: Guy Wilkinson, Mat Charles, Andrei Nomerotski, Sneha Malde, Lauren Martin, Andrea Contu, Hamish Gordon, Sophie Redford, Dan Johnson, Ed Greening, Ed Smith, Nick Torr, Rhorry Gauld, Donal Hill and Nazim Hussain. I also own many thanks to the Oxford secretaries Sue Geddes, Kim Proudfoot and Laura Nevay for all the administrative aid they have given me during my DPhil.

Thanks to countless people who made my time at both CERN and Oxford so memorable, space forbids me from mentioning all your names.

Finally, thanks to my parents and brother, to whom I owe so much; your continual encouragement and support over the years has been invaluable.

Preface

This thesis describes the analysis of proton-proton collision data accumulated by the LHCb experiment during 2010, corresponding to a total integrated luminosity of $\mathcal{L}_{\text{int}} = 35.6 \pm 3.6\ \text{pb}^{-1}$. This analysis is a first step towards a tree-level measurement of γ from the study of $B \to DK$ decays, as while these decay modes have little or no sensitivity to γ, much of the analysis procedure presented here will also be applicable to the future measurements of more γ-sensitive observables. This thesis also describes the measurement of the $D^0 + \overline{D}^0$ production cross-section at LHCb.

Chapter 1 introduces the Standard Model (SM) and describes the concept of *CP* violation, how it can manifest itself in the decays of neutral mesons and how it enters the SM through the CKM quark mixing matrix. A summary of various strategies for performing a tree-level measurement of γ using $B \to DK$ decays is then given.

An account of the hardware components and performance of the LHCb detector follows in Chap. 2. The work of a large number of people is necessarily described; however, particular focus is given to the aspects of the detector that are most important to the author's work, presented in Chaps. 3 to 5. This includes the description of the hardware and performance of the RICH sub-detectors, and the LHCb data acquisition and experimental control systems.

Chapter 3 describes the internal time alignment of the RICH L0 front-end boards using a pulsed laser system installed in the LHCb cavern and a *global* time alignment of the RICH detectors to the rest of the LHCb experiment using *pp* collision data. This is primarily the author's own work; however, this work has been done in collaboration with Conor Fitzpatrick[1] and Stig Topp-Jorgensen[2]. The author's work has also been aided by Sean Brisbane[2] and

[1] University of Edinburgh.
[2] University of Oxford.

Nick Styles[1], who initially developed the internal time-alignment procedure [1, 2].

In Chap. 4, the analysis of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ decays using data collected at LHCb is described, where $h^{\pm}$ indicates a kaon or pion. A measurement of various independent observables was performed by the author in collaboration with Malcolm John[2], Andrew Powell[2] and Paolo Gandini[2] using a combination of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ and the more numerous $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, h^{\pm}$ and $B^{\pm} \to (K^{+}K^{-})_D\, h^{\pm}$ modes [3]. The author focused on obtaining a high purity sample of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ decays by minimising the contamination from the principal background: the random combination of tracks, using a series of selection requirements applied to the reconstructed $B^{\pm}$ meson and its decay products. The parametrisation of the signal and background components used in the fit to extract the observables, which mainly utilised the $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, h^{\pm}$ and $B^{\pm} \to (K^{+}K^{-})_D\, h^{\pm}$ modes, was largely performed by others. The author has adapted the fitting procedure used in Ref. [3] to measure two independent observables purely from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ decays: the ratio of branching fractions, $\mathcal{B}_{DK}/\mathcal{B}_{D\pi} = \Gamma(B^{\pm} \to DK^{\pm})/\Gamma(B^{\pm} \to D\pi^{\pm})$, and the decay rate asymmetry $A^{K3\pi}_{fav} = \frac{\Gamma(B^{-}\to DK^{-})-\Gamma(B^{+}\to DK^{+})}{\Gamma(B^{-}\to DK^{-})+\Gamma(B^{+}\to DK^{+})}$.

In Chap. 5, measurements of the $D^0 + \overline{D}^0$ production cross-section at $\sqrt{s} = 7\,\text{TeV}$ in bins of transverse momentum and rapidity of the D^0 meson from prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ (and charge conjugate) decays at LHCb is presented. This study has been principally performed by the author, in collaboration with Jonas Rademacker[3], Tom Hampson[3], Paras Naik[3]. This chapter focuses on the author's contribution to this study.

Finally, Chap. 6 summarises the major conclusions of the works described in this thesis.

[1] University of Edinburgh.
[2] University of Oxford.
[3] University of Bristol.

Contents

Chapter 1

CP violation, and the CKM angle γ

1.1 Introduction

The Standard Model (SM) is an attempt at a unified theory of the fundamental particles and forces. It has so far withstood all experimental tests. However, it is clearly not a complete theory, since it fails to explain several observed phenomena, including the force of gravity, dark matter, dark energy and the expansion of the universe.

The Standard Model contains eighteen fundamental particles. These can be split into two groups: the *bosons*, which have integer intrinsic spin and are mediators of the fundamental physical forces, and *fermions*, which have half integer spin. The fermions can be further sub-divided into *leptons*, with charge of 0 or -1[1], and *quarks*, which have fractional charge. Each fermion in the SM has an *anti-matter* equivalent. These have similar properties to the fermions, except that their internal *quantum numbers*, such as charge, have opposite sign.

There are three *generations* of negatively charged leptons that differ only in their rest mass: the *electron* (e^-), *muon* (μ^-) and *tau* (τ^-). Each of the six charged leptons and *anti-leptons* has an associated neutral lepton called a *neutrino* (ν). The quarks can also be split into three generations. Each generation contains two quarks, one of which has a charge

[1] Unless otherwise stated, charge refers to the electric charge, and units of elementary charge (e) are implied throughout this document when referring to electric charge.

of $-1/3$ (labelled d, s and b in order of increasing mass) and one which has charge $+2/3$ (u, c and t).

The EM force is mediated by the photon (γ), the strong force is mediated by the gluon (g) and the weak force is mediated by the W^+, W^- and Z^0 bosons. The EM and weak forces act on both fermions and quarks, whilst the strong force only acts on quarks (and the gluons themselves). Because the weak bosons are massive (unlike the photon and gluon, which are massless), the weak force acts over a relatively short range. In addition, the weak force is the only interaction which allows changing quark flavour. The final boson of the SM is the Higgs, which is the mediating particle of the Higgs field; the Higgs mechanism, in which the Higgs field acquires a non-zero vacuum expectation energy, is the process by which the weak bosons (and the fermions) acquire mass in the SM.

Unlike the EM and weak forces, the strong force gets stronger with distance. This means that quarks can only exist in bound states, known as *hadrons*. Quarks have an additional quantum number called *colour*, which can take one of three values, referred to as *red*, *green* and *blue*. Similarly, anti-quarks can have colour quantum number *anti-red*, *anti-green* and *anti-blue*. The hadronic states are required to be colourless. A colourless state can be achieved by combining a quark with an antiquark of the equivalent anti-colour (*e.g.* red and anti-red). These are referred to as *mesons*. It is also possible to have a colourless state with three quarks (or anti-quarks), each with different colour charge[1]. These are called *baryons*.

Studying weak-interaction processes is a crucial part of LHCb's physics programme, as this is the only fundamental interaction in the SM which can violate *CP* symmetry, as discussed in the following section.

1.2 C, P and T symmetries

Symmetries are an essential part of the SM. Noether's theorem [4] tells us that for every continuous symmetry, there exists a corresponding conservation law, *i.e.* an observable which

[1] More complex quark states such as penta-quarks ($qqqq\bar{q}$) are theoretically possible, but have so far not been observed.

is constant over time.

In the gauge theory of interactions, the Lagrangian is invariant under a continuous group of local transformations. These transformations form a *symmetry* or *gauge* group. The generators of the group correspond to the conserved quantities. For example, *Quantum Electrodynamics* (QED) is the gauge theory of electromagnetism and its Lagrangian is invariant under transformations involving a single complex phase that varies in space-time. This leads to the conservation of charge.

However, Noether's theorem does not necessary hold for *discrete* symmetries. As the name implies, a discrete symmetry means that the system only displays symmetry for a discrete group of transformations. Discrete symmetries of particular interest to particle physics are *inversion* transformations. These are transformations in which the application of the transformation twice results in the original state, such as a mirror reflection. The discrete symmetries discussed in this thesis are *parity transformation*, *charge conjugation* and *time reversal*.

The effect of the parity inversion operator P is to transform a right-handed reference system into a left-handed one. The parity-inverted wave function, $\phi_P(t, \mathbf{x})$, is related to the original wave function, $\phi(t, \mathbf{x})$, by

$$\phi_P(t, \mathbf{x}) = P\phi(t, \mathbf{x}) = \eta_P \phi(t, -\mathbf{x}) \ , \tag{1.1}$$

where P is the parity operator, and η_P is the eigenvalue of the parity operation, which satisfies $|\eta_P|^2 = 1$ by conservation of probability.

Charge conjugation involves the exchange of particles with their anti-particle equivalents. A particle with four-momentum $\mathbf{p}$, spin s and charge q transforms under charge the conjugation operator, C, as

$$C|\mathbf{p}, s, q\rangle = \eta_C |\mathbf{p}, s, -q\rangle \ , \tag{1.2}$$

where η_C is the eigenvalue of the C operator.

Finally, the transformation of the time co-ordinate of a system $t \rightarrow -t$ is referred to as *time reversal*[1]. In quantum mechanics, the wave-function of the time-reversed state $\psi_T(t, \mathbf{x})$ is defined by

$$\psi_T(t, \mathbf{x}) = T\psi(t, \mathbf{x}) = \psi^*(-t, \mathbf{x}) \ , \tag{1.3}$$

where T is the time reversal operator.

In 1957, the first evidence for P and C violation was found in the beta decay of cobolt-60 [5]. It now appears that parity and charge conjugation are conserved in all EM and strong interactions, but are maximally violated in weak interactions.

1.2.1 The CPT theorem

The CPT theorem states that the product of parity, charge conjugation and time reversal is an exact symmetry (for proofs, see for example [6, 7, 8]). The CPT transformation of the Lagrangian density satisfying this theorem is given by

$$(CPT)^\dagger \mathcal{L}\,(t, \mathbf{x})\,(CPT) = \mathcal{L}\,(-t, -\mathbf{x}) \ \ , \tag{1.4}$$

where $\mathcal{L}\,(t, \mathbf{x})$ is the (Hermitian) Lagrangian density for a local (*i.e.* no action at a distance) and Lorentz-invariant quantum field theory. Such a Lagrangian density exists in the SM, as well as each of the separate field theories governing interactions under the fundamental forces according to the SM. Thus, CPT invariance is a fundamental property of the SM.

CPT invariance has important consequences for particle physics. In particular, it implies that a particle and its anti-particle must have the same total decay width (lifetime) [9]. For stable particles, it also implies that the particle and anti-particle have equal rest mass, whilst for unstable particles, it means that the expectation value of the Hamiltonian for the particle and antiparticle eigenstates are equal.

Furthermore, if CPT symmetry holds and T symmetry is violated, then this implies that

[1] A more intuitive name would be *motion reversal*, since classically, time reversal symmetry means that for each possible motion in the system allowed by the dynamical equations there exists one in which the direction of motion is reversed.

the symmetry of the combined C and P transformations (CP) must be violated by an equal and opposite amount and vice versa.

1.2.2 CP violation

After the discovery that P (and C) was violated in the weak interaction, the V-A theory was proposed [10]. In this theory, the weak interaction acts only on left-handed particles and right-handed anti-particles[1]. Thus, the weak interaction must maximally violate P and C but the combination of parity inversion and charge conjugation, CP, must be conserved.

For some years after the V-A theory was proposed, CP symmetry appeared to hold in all physical systems. However, in 1964, *CP* violation was observed in the neutral kaon system [11]. Since then, CP violation has been observed in the decays of B mesons by the BABAR [12] and BELLE [13] experiments. In addition, LHCb has recently reported evidence for *CP* violation in neutral D mesons [14].

1.3 Formalism of flavour mixing

Before discussing *CP* violation in detail, it is informative to introduce the general formalism of the mixing of *flavoured* mesons[2]. Flavoured neutral mesons have the property that the particle and antiparticle are only distinguishable by their flavour quantum number and this number is *not* conserved in weak interactions. Four such families of mesons are known to exist: neutral kaons, neutral D mesons, and the B_s^0 and B^0 mesons.

The flavour eigenstates of a generic flavoured neutral meson, M^0, with a non-zero flavour eigenvalue and its antiparticle, $\bar{M}^0$, are defined by

$$F|M^0\rangle = +|M^0\rangle \quad \text{and} \quad F|\bar{M}^0\rangle = -|\bar{M}^0\rangle \ , \tag{1.5}$$

[1] *Chirality* (handedness) is an intrinsic property of a particle. For a massless particle it is proportional to the helicity, $h = \mathbf{S} \cdot \hat{\mathbf{p}}$, where $\mathbf{S}$ and $\hat{\mathbf{p}}$ are the particle's spin and momentum unit-vector respectively.

[2] A *flavoured* hadron is one which has a non-zero strangeness, charm or beauty quantum number (where the s and b quark have strangeness and beauty of -1 respectively and the c quark has charm +1, whilst the anti-quarks have the opposite sign).

where F signifies the flavour operator. Since the two states are CP-conjugates, they must satisfy

$$CP|M^0\rangle = e^{i\xi_{CP}}|\bar{M}^0\rangle \quad \text{and} \quad CP|\bar{M}^0\rangle = e^{-i\xi_{CP}}|M^0\rangle \ , \tag{1.6}$$

where ξ_{CP} is an arbitrary real number. It is convenient to chose the phase such that $\xi_{CP} = 0$. In this case, Eq. (1.6) becomes

$$CP|M^0\rangle = |\bar{M}^0\rangle \quad \text{and} \quad CP|\bar{M}^0\rangle = |M^0\rangle \ . \tag{1.7}$$

The weak Hamiltonian does not commute with flavour, since the weak interactions do not conserve the flavour quantum numbers. As such, the flavour eigenstates do not correspond to the physical states of the system. If CP symmetry holds and the weak Hamiltonian commutes with the CP operator, then then a good candidate for the physical states are the eigenstates of the CP operator, namely

$$|M_+\rangle = \frac{1}{\sqrt{2}}\left(|M^0\rangle + |\bar{M}^0\rangle\right) \quad \text{and} \quad |M_-\rangle = \frac{1}{\sqrt{2}}\left(|M^0\rangle - |\bar{M}^0\rangle\right) \ . \tag{1.8}$$

By definition, the CP eigenstates satisfy $CP|M_\pm\rangle = \pm|M_\pm\rangle$. Since the CP eigenstates are not CP conjugates of one another, they are allowed to have different invariant masses and lifetimes without violating CPT symmetry.

1.3.1 Time evolution of the flavour eigenstates

Determining the time evolution equations for neutral mesons is complicated by the mixing of the flavour eigenstates and the fact that they can decay into various other states. The situation is much simpler if the following approximations are made:

1. the initial states are only linear combinations of the flavour eigenstates;
2. the final states do not couple to the weak interaction;
3. the time evolution of the coefficients of the flavour eigenstates occurs over a much larger time-scale than the typical time-scale of the strong interactions.

This is referred to as the *Weisskopf-Wigner* approximation [15, 16]. Under this approximation, the weak interactions can be considered as a perturbation of the strong interactions, with an effective weak Hamiltonian $\boldsymbol{\mathcal{H}}$, which is defined by

$$\boldsymbol{\mathcal{H}} = \boldsymbol{H}_0 + \boldsymbol{H}_{\text{int}} + \sum_n \frac{\boldsymbol{H}_{\text{int}}|n\rangle\langle n|\boldsymbol{H}_{\text{int}}}{m_0 - E_n} , \tag{1.9}$$

where $\boldsymbol{H}_0$ and $\boldsymbol{H}_{\text{int}}$ are the unperturbed and perturbed parts of the Hamiltonian respectively, m_0 is the unperturbed mass of M^0 and $\bar{M}^0$, and E_n is the energy of the intermediate state n in the rest frame of the meson.

The generic state under the Weisskopf-Wigner approximation is given by:

$$|\psi\rangle = a(t)|M^0\rangle + b(t)|\bar{M}^0\rangle \equiv \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} . \tag{1.10}$$

The time evolution of $|\psi\rangle$ is therefore given by

$$i\hbar\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \equiv \boldsymbol{\mathcal{H}}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} \mathcal{H}_{11} & \mathcal{H}_{12} \\ \mathcal{H}_{21} & \mathcal{H}_{22} \end{pmatrix} , \tag{1.11}$$

where $\mathcal{H}_{12} = \langle M^0|\boldsymbol{\mathcal{H}}|\bar{M}^0\rangle$ and similar for the other matrix elements.

The effective weak Hamiltonian is not Hermitian, however it can be expressed as the linear combination of two Hermitian matrices,

$$\boldsymbol{\mathcal{H}} = \mathbf{M} - \frac{i}{2}\boldsymbol{\Gamma} , \tag{1.12}$$

where $\mathbf{M}$ and $\boldsymbol{\Gamma}$ are the *mass* and *decay* matrices respectively. The off-diagonal elements represent the dispersive and absorptive parts (virtual and real intermediate states) respectively of the transition amplitude between M^0 and $\bar{M}^0$.

Since the matrices on the right-hand side of Eq. (1.12) are Hermitian, the off-diagonal terms are complex conjugates of one another. If CPT invariance is also imposed, then the diagonal terms are equal (*i.e.* the two flavour eigenstates have the same invariant mass and total decay width). Thus, the matrix elements are given by

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{11} \end{pmatrix} \quad \text{and} \quad \mathbf{\Gamma} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{11} \end{pmatrix} . \tag{1.13}$$

The physical states $|M_L\rangle$ and $|M_H\rangle$ are the eigenstates of $\boldsymbol{\mathcal{H}}$, with (complex) eigenvalues

$$\boldsymbol{\mathcal{H}}|M_L\rangle = \lambda_L|M_L\rangle \quad \text{and} \quad \boldsymbol{\mathcal{H}}|M_H\rangle = \lambda_H|M_H\rangle \ , \tag{1.14}$$

where $\lambda_{(L,H)} \equiv m_{(L,H)} - \frac{i}{2}\Gamma_{(L,H)}$, with $m_{(L,H)}$ and $\Gamma_{(L,H)}$ signifying the masses and total widths of the two physical states[1]. Choosing a phase convention where the mixing parameters of the two states are equal, the physical states are written in terms of the flavour eigenstates as

$$|M_L\rangle = p|M^0\rangle + q|\bar{M}^0\rangle \quad \text{and} \quad |M_H\rangle = p|M^0\rangle - q|\bar{M}^0\rangle \ , \tag{1.15}$$

where p and q are complex numbers satisfying $|p|^2 + |q|^2 = 1$.

The physical states evolve in time as

$$|M_L(t)\rangle = e^{-i\lambda_L t}|M_L(0)\rangle = e^{-im_L t}e^{-\frac{1}{2}\Gamma_L t}|M_L(0)\rangle \ , \tag{1.16a}$$

$$|M_H(t)\rangle = e^{-i\lambda_H t}|M_H(0)\rangle = e^{-im_H t}e^{-\frac{1}{2}\Gamma_H t}|M_H(0)\rangle \ . \tag{1.16b}$$

Rearranging Eq. (1.15) in terms of the flavour eigenstates gives

$$|M^0\rangle = \frac{|M_L\rangle + |M_H\rangle}{2p} \ , \tag{1.17a}$$

$$|\bar{M}^0\rangle = \frac{|M_L\rangle - |M_H\rangle}{2q} \ . \tag{1.17b}$$

Finally, applying the time evolution operator to Eqs. (1.17a) and (1.17b), and substituting Eqs. (1.16a) and (1.16b), gives the time evolution of the meson and anti-meson states from a

[1] The labelling of the physical states is arbitrary. Since the B and D meson states can be distinguished by their masses we adopt the convention $|M_L\rangle$ and $|M_H\rangle$ to label the heavy and light states respectively.

time $t = 0$ to a time t:

$$|M^0(t)\rangle = f_+(t)|M^0(0)\rangle + \frac{q}{p} f_-(t)|\bar{M}^0(0)\rangle \ , \tag{1.18a}$$

$$|\bar{M}^0(t)\rangle = \frac{p}{q} f_-(t)|M^0(0)\rangle + f_+(t)|\bar{M}^0(0)\rangle \ . \tag{1.18b}$$

The time evolution parameters $f_+(t)$ and $f_-(t)$ are defined by

$$f_\pm(t) = \frac{1}{2} e^{-im_L t} e^{-\frac{1}{2}\Gamma_L t} \left(1 \pm e^{-i\Delta m\, t} e^{-\frac{1}{2}\Delta\Gamma\, t}\right) \ , \tag{1.19}$$

where $\Delta m = m_H - m_L$ and $\Delta\Gamma = \Gamma_H - \Gamma_L$.

From the above equations, the probability that the system remains in the original flavour eigenstate at a time t is given by

$$P\left(M^0(t) \to M^0\right) = P\left(\bar{M}^0(t) \to \bar{M}^0\right) = |f_+|^2 \ , \tag{1.20}$$

The fact that the two mixing probabilities are the same is a consequence of *CPT* invariance.

The probability of an initial state M^0 being measured in state $\bar{M}^0$ at time t and the equivalent probability for an initial state $\bar{M}^0$ and final state M^0 are given by

$$P\left(M^0(t) \to \bar{M}^0\right) = \left|\frac{q}{p}\right|^2 |f_-|^2 \quad \text{and} \quad P\left(\bar{M}^0(t) \to M^0\right) = \left|\frac{p}{q}\right|^2 |f_-|^2 \ . \tag{1.21}$$

If *CP* is conserved in the mixing (see Sec. 1.4.1), then these two probabilities are also equal.

The mass difference, Δm, is often referred to as the *oscillation frequency*. The origin of this term can be seen if the oscillation probability $|f_-|^2$ is determined using Eq. (1.19):

$$|f_-|^2 = \frac{e^{-\bar{\Gamma}\, t}}{2} \left(\cosh\left(-\frac{\Delta\Gamma}{2} t\right) - \cos\left(\Delta m\, t\right)\right) \tag{1.22}$$

where $\bar{\Gamma} = \frac{1}{2}\left(\Gamma_H + \Gamma_L\right)$ is the average decay width. Thus, Δm represents the frequency by which oscillation between the flavour eigenstates occurs.

1.4 Types of CP violation

A convenient way of classifying the different ways in which CP violation can occur in neutral meson decays is to subdivide into the following three mechanisms:

1. CP violation in mixing;
2. CP violation in decays;
3. CP violation in the interference between mixing and decay.

These mechanisms are described in the following sub-sections. At this point, it is useful to introduce the *decay amplitudes* A_f and $\bar{A}_f$, which are the transition amplitudes between initial states M^0 and $\bar{M}^0$ respectively and final state f, and are defined by

$$A_f = \langle f|\boldsymbol{H}_{\text{int}}|M^0\rangle \quad \text{and} \quad \bar{A}_f = \langle f|\boldsymbol{H}_{\text{int}}|\bar{M}^0\rangle \ , \tag{1.23}$$

where $\boldsymbol{H}_{\text{int}}$ is the perturbative part of the Hamiltonian[1].

1.4.1 CP violation in mixing

From Eqs. (1.8) and (1.15), the physical states can be written in terms of the CP eigenstates as

$$|M_L\rangle = \frac{p+q}{\sqrt{2}}|M_+\rangle + \frac{p-q}{\sqrt{2}}|M_-\rangle \quad \text{and} \quad |M_H\rangle = \frac{p+q}{\sqrt{2}}|M_-\rangle + \frac{p-q}{\sqrt{2}}|M_+\rangle \ . \tag{1.24}$$

If CP is conserved in the neutral flavour mixing, then $|M_L\rangle$ and $|M_H\rangle$ are also CP eigenstates. Since $CP|M_\pm\rangle = \pm|M_\pm\rangle$, it follows from Eq. (1.24) that the condition for CP violation in mixing is

$$\left|\frac{q}{p}\right| \neq 1. \tag{1.25}$$

[1] This definition of the transition amplitude is in fact an approximation (known as the *Born approximation*), which is valid providing that the matrix elements of $\boldsymbol{H}_{\text{int}}$ are much smaller than the elements of the unperturbed Hamiltonian. In this case, it means that the $\boldsymbol{H}_{\text{int}}$ is equivalent to the weak Hamiltonian, *i.e.* weak processes are the only interactions in the decay. This approximation breaks down if we consider final-state interactions.

This type of *CP* violation is also referred to as *indirect CP* violation.

1.4.2 *CP* violation in decay

In this case, *CP* violation occurs when the moduli of the physical decay amplitudes for *CP* conjugate processes into final states f and $\bar{f}$ are different, *i.e.* $|A_f| \neq |\bar{A}_{\bar{f}}|$. This is also referred to as *direct CP* violation. To understand how this arises, it is informative to consider the simplest case in which direct *CP* violation can occur: the interference between two decay amplitudes. Let A_1 and A_2 represent the two possible ways (intermediate resonances) in which initial state M^0 can decay to final state f. Since the final states are the same, the two decays will interfere. As the decay amplitudes are complex we can write the total amplitude in terms of the magnitudes and phases as

$$A_f = |A_1| \, e^{i\phi_1} + |A_2| \, e^{i\phi_2} \ . \tag{1.26}$$

CP violation would imply that $A_f \neq e^{i\phi} \bar{A}_{\bar{f}}$, *i.e.* the amplitude of the *CP* conjugate decay cannot be expressed in terms of the original amplitude and an arbitrary phase. This requires that $\phi_1 \neq \phi_2$ and that the phases must change sign under *CP* transformation, *i.e.*

$$\bar{A}_{\bar{f}} = |A_1| \, e^{-i\phi_1} + |A_2| \, e^{-i\phi_2} \ . \tag{1.27}$$

Since the weak interaction is the only known process in the SM by which *CP* violation can occur, these *CP* violating phases are referred to as *weak* phases.

Obviously, the decay amplitudes are not observable, only the transition probabilities, defined by

$$|A_f|^2 = |A_1|^2 + |A_2|^2 + 2 \, |A_1| \, |A_2| \cos\left(\phi_1 - \phi_2\right) \ , \tag{1.28a}$$

$$\left|\bar{A}_{\bar{f}}\right|^2 = |A_1|^2 + |A_2|^2 + 2 \, |A_1| \, |A_2| \cos\left(\phi_2 - \phi_1\right) \ . \tag{1.28b}$$

Since the cosine function is even, the two decay amplitudes are equal. Thus, if Eqs. (1.26) and (1.27) were the physical decay amplitudes, then it would not be possible to measure direct *CP* violation. This problem is solved if the phases in the decay amplitudes consist of both a *CP* violating and *CP* conserving part, *i.e.*

$$A_f = |A_1|\, e^{i\phi_1} e^{i\delta_1} + |A_2|\, e^{i\phi_2} e^{i\delta_2} \ , \tag{1.29a}$$

$$\bar{A}_{\bar{f}} = |A_1|\, e^{-i\phi_1} e^{i\delta_1} + |A_2|\, e^{-i\phi_2} e^{i\delta_2} \ . \tag{1.29b}$$

These *CP*-conserving phases can arise due to final-state interactions (FSI), even without the presence of *CP* violation. The main sources of FSI are processes in which the final state particles interact under the strong interaction. Thus, the *CP*-conserving phases are referred to as the *strong* phases.

After the introduction of the strong phases, the difference in transition rates is given by

$$\left|\bar{A}_{\bar{f}}\right|^2 - \left|A_f\right|^2 = 2\,|A_1|\,|A_2| \sin\left(\phi_1 - \phi_2\right) \sin\left(\delta_1 - \delta_2\right) \ . \tag{1.30}$$

This difference vanishes if the two interfering amplitudes have the same weak phase, but also if they have the same strong phase. Thus, the presence of FSI is essential for the observation of direct *CP* violation.

In general, the decays proceed not by two, but by numerous intermediate resonances. Thus, the general requirement for direct *CP* violation is given by

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| = \left|\frac{\sum_{j=1}^{n} |A_j|\, e^{i(\delta_j - \phi_j)}}{\sum_{j=1}^{n} |A_j|\, e^{i(\delta_j + \phi_j)}}\right| \neq 1 \ . \tag{1.31}$$

Direct *CP* violation is the only possible type of *CP* violation in the decay of charged particles, since mixing is forbidden due to charge conservation.

1.4.3 *CP* violation in the interference of mixing and decay

For neutral mesons, a third classification of *CP* violation is possible, which can occur if there exists at least one final state which can be reached by both flavour eigenstates (this includes, but is not limited to all *CP* eigenstates). This type of *CP* violation can occur even if there is no direct or indirect *CP* violation.

If a flavour eigenstate M^0 decays to final state f and $\bar{M}^0$ can also decay to f, then the final state can be reached either directly or via $M^0 - \bar{M}^0$ mixing followed by the decay of $\bar{M}^0$ to f. The overall decay amplitude for this process can be written as

$$A_f = A\left(M^0 \to f\right) + A\left(M^0 \to \bar{M}^0\right) A\left(\bar{M}^0 \to f\right) \tag{1.32}$$

and for the *CP* conjugate process as

$$\bar{A}_f = A\left(\bar{M}^0 \to f\right) + A\left(\bar{M}^0 \to M^0\right) A\left(M^0 \to f\right) \ . \tag{1.33}$$

Denoting θ_M and $\bar{\theta}_M$ as the phases of $M^0 \to \bar{M}^0$ and $\bar{M}^0 \to M^0$ mixing respectively, and θ_f and $\bar{\theta}_f$ as the decay amplitudes of $M^0 \to f$ and $\bar{M}^0 \to f$ respectively, then the relative phase between the two terms in Eq. (1.32), θ, and the relative phase between the two terms in Eq. (1.33), $\bar{\theta}$, are given by

$$\theta = \theta_f - \bar{\theta}_f - \theta_M \ , \tag{1.34}$$

$$\bar{\theta} = \bar{\theta}_f - \theta_f - \bar{\theta}_M \ . \tag{1.35}$$

If *CP* is conserved in both the mixing and decay, then the moduli of the decay amplitudes in Eqs. (1.32) and (1.33) are equal. However, this does not constrain the relative phases in Eqs. (1.34) and (1.35). For *CP* symmetry to hold in the interference between mixing and decay, then *both* mixing and decay phases must be equal (*i.e.* $\theta_f = \bar{\theta}_f$ and $\theta_M = \bar{\theta}_M$). Since the mixing phase is independent of the final state, any difference in the level of *CP* violation between different final states would be unambiguous evidence of direct *CP* violation.

It is convenient to introduce the complex quantity η_f, defined by

$$\eta_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \ . \tag{1.36}$$

This quantity is invariant to the choice of the arbitrary phase between $|M^0\rangle$ and $|\bar{M}^0\rangle$, so its phase is a physical quantity. The general condition for *CP* symmetry can then be expressed as

$$\eta_f = \eta_{\bar{f}}^{-1} \ . \tag{1.37}$$

If the final states are *CP* eigenstates, then $f = \bar{f}$, and the condition reduces to $\eta_f = \pm 1$.

1.5 *CP* violation in the Standard Model

At this point, it is useful to discuss how *CP* violation is incorporated into the SM. Cabibbo introduced the *Cabibbo angle*, θ_C, to explain why processes such as $K^+ \to \mu^+ \nu_\mu$ and $\mu^+ \to \mu^+ \nu_\mu$ have have different decay amplitudes. This difference in rates seemed to violate *universality*, the idea that the coupling strength should be the same for all interactions under a particular force. Cabibbo restored universality by suggesting that the weak current consists of an admixture of $\Delta S = 0$ and $\Delta S = 1$ axial and vector currents, where S represents the strangeness quantum number [17].

At the time of Cabibbo's work, the quark model had not been theorised. Subsequently, the GIM mechanism [18] was proposed to explain why Flavour Changing Neutral Currents (FCNCs), such as $K^+ \to \pi^+ e^+ e^-$ are heavily suppressed; this theory requires the existence of quarks. The hadrons that had been observed at the time were all consistent with a three-quark model comprising u, d, and s quarks. However, the three-quark model could not explain the level of FCNCs observed. To suppress FCNCs, the GIM mechanism also required the existence of a fourth type of quark called the charm (c) quark, the discovery of which corroborated both the GIM mechanism and the quark model.

With the knowledge that hadrons consist of quarks, the Cabibbo angle reflects the fact that the weak eigenstate d', which couples to the u quark, is a superposition of the d and s mass eigenstates of the strong (and EM) interactions. The existence of the c quark indicates that there is also a second weak eigenstate s'. The mixing of the mass eigenstates to form the weak eigenstates can be written in matrix form as:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{1.38}$$

where the quantities V_{ij} represent the weak coupling of quark i to quark j. This mixing matrix is required to be unitary, *i.e.* $V_{ji} = V_{ij}^*$. This ensures that the mixing matrix is the same regardless of whether we consider an up-type quark coupling to a mixture of down-type quarks, or vice-versa (*i.e.* we could consider the weak eigenstates as u' and c' rather than d' and s').

The GIM mechanism predicted only two generations of quarks. However, Kobayashi and Maskawa determined that *CP* violation could be introduced into the quark mixing matrix if there exists at least one additional quark generation [19]. This so-called *CKM* matrix is described in more detail in the following sub-section.

1.5.1 The CKM matrix

An $N \times N$ unitary matrix can be expressed in terms of N^2 parameters: $\frac{N(N-1)}{2}$ Euler angles, and $\frac{N(N+1)}{2}$ phases. The number of independent phases of the mixing matrix are reduced to $\frac{(N-1)(N-2)}{2}$ by the fact that both the mass and weak eigenstates must be orthogonal and normalised. Thus, there is a single Euler angle (the Cabibbo angle) for a 2×2 matrix, and no phases. However, if the number of quark generations is extended to three, then there are three angles and a single phase. As discussed in Sec. 1.4.2, direct *CP* violation requires an irreducible complex phase between two decay amplitudes which changes sign under *CP* conjugation. Thus, this phase is responsible for *CP* violation in the SM.

Extending Eq. (1.38) to three down-type quarks, the SM quark mixing matrix $\mathbf{V_{CKM}}$, is defined by

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathbf{V_{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cd} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.39}$$

Several parametrisations exist for the elements of the CKM matrix, but "standard" notation, also known as the "Chau-Keung" parametrisation, uses three Euler angles $\theta_{12} \equiv \theta_C$, θ_{13} and θ_{23}, and the CP violating phase δ_{13} [20], which satisfy

$$\mathbf{V_{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -\left(s_{12}c_{23} + c_{12}s_{23}s_{13}e^{i\delta_{13}}\right) & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -\left(c_{12}s_{23} + s_{12}c_{23}s_{13}e^{i\delta_{13}}\right) & c_{23}c_{13} \end{pmatrix}, \quad (1.40)$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. Under this notation, the mixing between generations i and j vanishes if $\theta_{ij} = 0$.

An alternative approximation was developed by Wolfenstein [21], in which $\mathbf{V_{CKM}}$ is expanded in orders of $\lambda \equiv s_{12} \approx 0.22$. To $\mathcal{O}(\lambda^3)$, the Wolfenstein parametrisation gives

$$\mathbf{V_{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \quad (1.41)$$

where A, η and ρ are real parameters of order 1. This parametrisation is useful because it gives a simpler indication of the scale of the magnitudes of the CKM parameters than the standard parametrisation. It also indicates that CP is violated at order λ^3 for $b \to u$ and $t \to d$ transitions, but at higher orders of λ for other quark transitions. In addition, as will be shown later in this section, the ratio of some of the branching fractions can be directly represented in terms of the Wolfenstein parameters.

1.5.2 Unitary triangles

The unitary of the CKM matrix can be summarised by

$$\sum_{k=1}^{3} V_{ik}V_{kj}^{\dagger} = V_{ik}V_{jk}^{*} = \delta_{ij} , \quad (1.42)$$

where δ_{ij} is the Kronecker delta. The three cases where $i = j$ (*i.e.* $\delta_{ij} = 1$) indicate that the sum of the couplings of one of the up-type quarks to the down-type quarks is the same for all generations, which is the weak universality condition predicted by Cabibbo. The remaining six conditions can be represented as triangles in the complex plane. Unitarity then requires that the three angles sum to 180°.

The *unitary triangles* are independent of any parametrisation of the CKM matrix elements, since changing the parametrisation amounts to rotating the triangle in the complex plane without changing the lengths and internal angles.

1.5.3 The CKM angles α, β and γ

Of particular importance to *CP* violation is the following unitarity condition:

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \ . \tag{1.43}$$

The triangle formed by this unitary condition is commonly referred to as *The* Unitary Triangle (UT). The three terms in Eq. (1.43) are all $\mathcal{O}\left(\lambda^3\right)$ in the Wolfenstein parametrisation, and therefore the three sides of the UT have similar magnitudes, as do the angles. Thus, they are all are accessible to experimental determination. The angles of the UT, labelled α, β and γ are defined by

$$\alpha = \arg\left(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}}\right) \approx \arg\left(-\frac{1-\rho-i\eta}{\rho+i\eta}\right) \ , \tag{1.44a}$$

$$\beta = \arg\left(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}}\right) \approx \arg\left(\frac{1}{1-\rho-i\eta}\right) \ , \tag{1.44b}$$

$$\gamma = \arg\left(-\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}}\right) \approx \arg\left(\rho+i\eta\right) \ . \tag{1.44c}$$

It is convenient to normalise the sides of the UT by $V_{cb}^* V_{cd}$. This means that one side of the triangle is real and of unit length. The normalised UT is shown pictorially in Fig. 1.1. The real and complex components of the apex of the triangle, $\bar{\rho}$ and $\bar{\eta}$ respectively, are defined in terms of the Wolfenstein parameters (to $\mathcal{O}(\lambda^3)$) by

$$\bar{\rho} = \rho\left(1 - \lambda^2/2\right) \quad \text{and} \quad \bar{\eta} = \eta\left(1 - \lambda^2/2\right) \; . \tag{1.45}$$

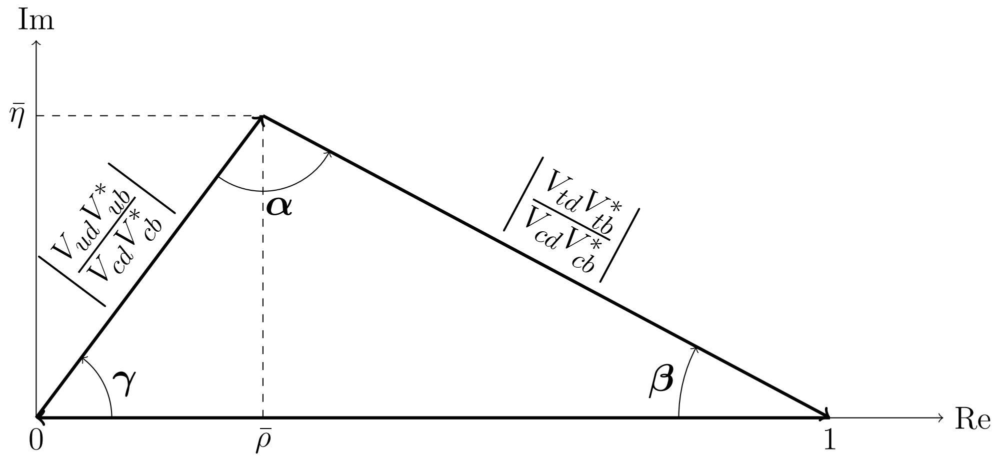


Figure 1.1: Diagram of the Unitary Triangle in the complex plane. The sides are normalised to $V_{cd}V_{cb}^*$, such that the base is real and of unit length.

1.5.4 Constraints on the Unitary Triangle parameters

In principle, the normalised UT can be fully constrained by the measurement of any two of the UT parameters: the sloping sides, two of the three internal angles, or $\bar{\rho}$ and $\bar{\eta}$. However, as previously mentioned, the geometry of the UT allows all the lengths and angles to be determined experimentally and thus over-constrain the UT, increasing the overall precision of the quark mixing parameters and introducing better limits on the level of SM *CP* violation. If the CKM matrix is unitary, then all of the various UT parameters should converge on a single point in the complex $(\bar{\rho}, \bar{\eta})$ plane. If any one parameter deviates from this, then this would be an indicates a violation of unitarity and constitute indirect evidence of New Physics (NP), *i.e.* either extensions to the SM or an entirely new model of particle physics.

Figure 1.2 shows the constraints on the UT as of summer 2011. The red and yellow hatched regions represent the 68% and 95% confidence levels of the apex. The quantities Δm_d and Δm_s represent the mass difference between the heavy and light B^0 and B_s^0 mass eigenstates respectively. These quantities provide constraints on $|V_{td}|$ and $|V_{ts}|$ through neutral B mixing mediated by Feynman diagrams involving the exchange of a virtual $t\bar{t}$ quark pair. The parameter ϵ_K is a measure of indirect CP violation in the K^0-$\overline{K}^0$ system (related to the ratio of mixing parameters p/q), and provides a hyperbolic constraint on the apex. The results indicate that the various measurements are consistent with the global fit result at the 95% confidence level. It should be noted that the CKMFitter results assume the unitarity of the CKM matrix. The experimental uncertainty of γ in particular, is significantly larger when considering only direct measurements. The current world averages of the UT internal angles from direct measurements [22] are:

$$\alpha = \left(89.0^{+4.4}_{-4.2}\right)^\circ \ , \qquad \beta = \left(21.38^{+0.79}_{-0.77}\right)^\circ \quad \text{and} \quad \gamma = \left(68^{+10}_{-11}\right)^\circ \ ,$$

where the uncertainties quoted represent the 68% confidence level. The sum of the UT angles are therefore consistent with 180°, albeit with large uncertainties. Precision measurements of the UT angles, particularly α and γ, is thus one of the key goals of flavour physics.

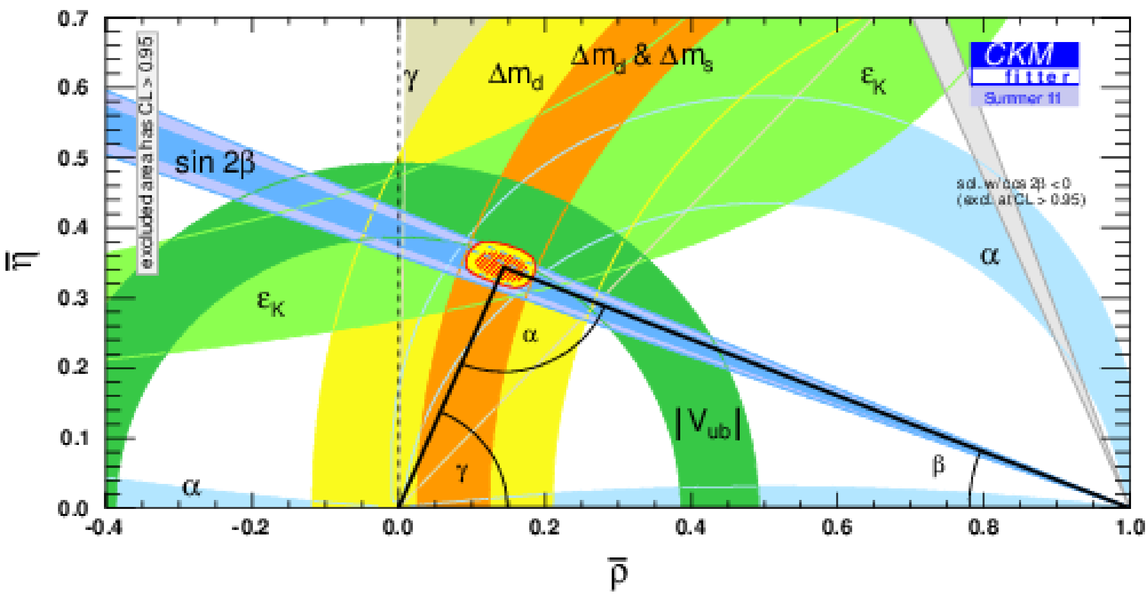


Figure 1.2: Constraints of the Unitary Triangle as of summer 2011, from the CKMFitter group [22]. The shaded areas represent the 95% confidence levels of each parameter.

1.5.5 Motivation for precision measurements of γ

As already indicated, γ is the internal angle of the UT with the largest experimental uncertainty, and so a precision measurement is vital for testing unitarity. Moreover, as can be seen by Eqs. (1.44a) to (1.44c), it is the only angle which does not depend on couplings to the top quark. This means it can be extracted from tree-level processes (*i.e.* the only Feynman diagrams which contribute to the decay amplitude at leading order are tree diagrams). It is assumed that there is no significant contribution to the decay amplitudes of tree-level SM processes from NP. Any difference in γ measured from tree-level processes, and processes with contributions from both tree and loop diagrams would therefore be an indication of NP. Thus, a precision tree-level measurement of γ provides an important SM benchmark, against which processes sensitive to NP can be compared.

1.6 Extracting γ from tree-level B decays

Sensitivity to γ from tree-level processes comes from the interference between $b \to c$ and $b \to u$ transitions (and their charge conjugates). One such process of particular importance to γ extraction comes from the decays $B^\pm \to D^0 K^\pm$ and $B^\pm \to \overline{D}^0 K^\pm$. As indicated by Figs. 1.3a and 1.3b, these are both tree-level processes, and therefore they can be used to determine a SM benchmark for γ. The ratio of the decay amplitudes are given by

$$\frac{A\left(B^- \to \overline{D}^0 K^-\right)}{A\left(B^- \to D^0 K^-\right)} = r_B e^{i(\delta_B - \gamma)} \quad \text{and} \quad \frac{A\left(B^+ \to D^0 K^+\right)}{A\left(B^+ \to \overline{D}^0 K^+\right)} = r_B e^{i(\delta_B + \gamma)} , \tag{1.46}$$

where r_B is the magnitude of the ratio of the decay amplitudes, hereby referred to as the *amplitude ratio* and δ_B is the CP-conserving strong phase between the two amplitudes.If D^0 and $\overline{D}^0$ decay to the same final state, f_D[1], then the D^0 and $\overline{D}^0$ are indistinguishable and so the two decays interfere. The CKM angle γ can then be extracted by exploiting the fact that the phase involving γ changes sign under CP conjugation, as indicated by Eq. (1.46). Thus, a difference in the rates of the B^+ and B^- decays would indicate CP

[1] Here and subsequently, D refers to either a D^0 or $\overline{D}^0$.

violation and the level of asymmetry can be used to determine γ. Since the $B^- \to \overline{D}^0 K^-$ and $B^+ \to D^0 K^+$ diagrams are colour suppressed[1], r_B is comparably small; the current world average is $r_B = 0.113^{+0.024}_{-0.021}$ [23]. As a result, interference effects tend to be small. There are a variety of methods for extracting γ from these decays. They tend to be grouped by the final state considered.

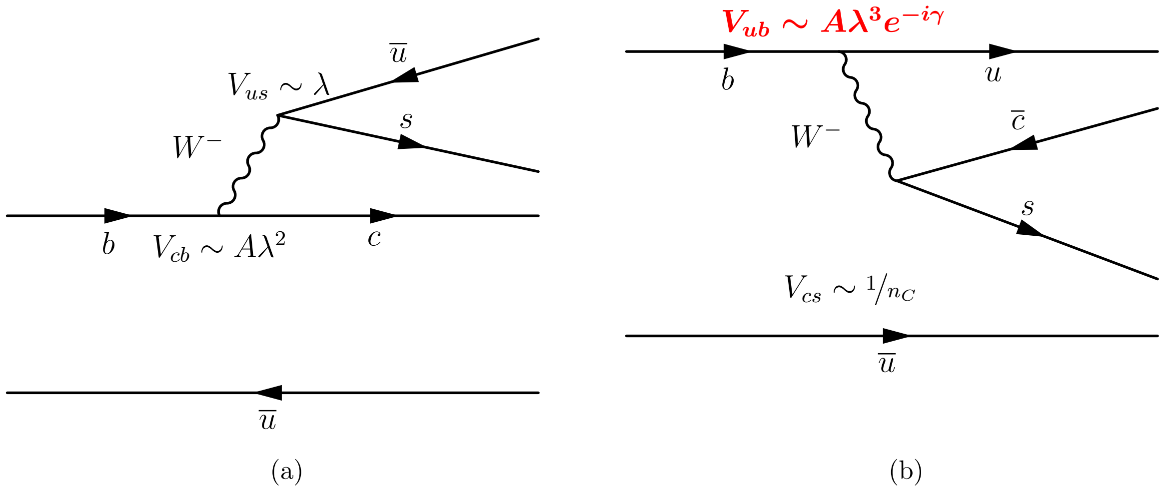


Figure 1.3: Feynman diagrams of the decays (a) $B^- \to D^0 K^-$ and (b) $B^- \to \overline{D}^0 K^-$. The ratio of the decay amplitudes has a magnitude, r_B, and phase, $\delta_B - \gamma$. The approximate couplings (neglecting the weak coupling constant, which is the same for all vertices) are indicated in terms of the Wolfenstein parameters. The coupling ${}^1/_{n_C}$ in (b) indicates that this vertex is colour suppressed, where $n_C = 3$ is the number of quark colours (or anticolours).

1.6.1 GLW and ADS methods

In the method for extracting γ proposed by Gronau, London and Wyler (the *GLW* method), the final state f_D is a *CP* eigenstate, such as K^+K^- or $\pi^+\pi^-$ [24, 25]. The extraction of γ is then simply a matter of comparing the numbers of reconstructed $B^+ \to DK^+$ and $B^- \to DK^-$ decays to the *CP* eigenstate, which, as already mentioned, will be different if *CP* is violated. However, this method suffers from the fact that r_B is small and so any difference in event rates will also be small.

An alternative method proposed by Atwood, Dunietz and Soni (the *ADS* method)

[1] Colour suppression refers to the fact that the two quarks in the charm meson originate from different (colour-conserving) vertices in the Feynman diagram. This means that the colour of one quark determines the colour of the other (since the final-state meson must be colourless). Conversely, in the colour-favoured decays the quarks originate from the same vertex and so the quarks can be any colour.

considers decays which are not *CP* eigenstates, such as $K^{\pm}\pi^{\mp}$ [26]. In this method, as indicated by the interference diagrams in Fig. 1.4, there is an additional suppression factor, r_D[1]. This is because $\overline{D}^0 \to K^-\pi^+$ and its charge conjugate are *Doubly Cabibbo Suppressed* (DCS)[2] as indicated by the Feynman diagram Fig. 1.5b, whilst $D^0 \to K^-\pi^+$, shown in Fig. 1.5a is Cabibbo Favoured (CF). Measurements indicate that the amplitude ratio, r_D, has a value 0.0575 ± 0.007 [27]. Since r_B and r_D are comparable, when the two kaons in the final state have opposite sign, as in Fig. 1.4b, the interference is maximal. The disadvantage of this method is that the combination of colour suppression and Cabibbo suppression makes the optimal decays for γ extraction relatively rare compared to the same-sign and *CP* eigenstate decays. The decay rates of the four $K^{\pm}\pi^{\mp}$ modes are given by

$$\Gamma\left(B^- \to \left(K^-\pi^+\right)_D K^-\right) = N_{K\pi}\left(1 + \left(r_B r_D\right)^2 + 2 r_B r_D \cos\left(\delta_B - \delta_D - \gamma\right)\right) \ , \tag{1.47a}$$

$$\Gamma\left(B^+ \to \left(K^+\pi^-\right)_D K^+\right) = N_{K\pi}\left(1 + \left(r_B r_D\right)^2 + 2 r_B r_D \cos\left(\delta_B - \delta_D + \gamma\right)\right) \ , \tag{1.47b}$$

$$\Gamma\left(B^- \to \left(K^+\pi^-\right)_D K^-\right) = N_{K\pi}\left(r_B^2 + r_D^2 + 2 r_B r_D \cos\left(\delta_B + \delta_D - \gamma\right)\right) \ , \tag{1.47c}$$

$$\Gamma\left(B^+ \to \left(K^-\pi^+\right)_D K^+\right) = N_{K\pi}\left(r_B^2 + r_D^2 + 2 r_B r_D \cos\left(\delta_B + \delta_D + \gamma\right)\right) \ . \tag{1.47d}$$

Equations (1.47a) to (1.47d) indicate that there are four observables and six parameters, including the constant of proportionality, $N_{K\pi}$. Since r_D is already well-measured, the number of parameters can be reduced to five by fixing r_D. The strong phases are currently measured to be $\delta_B = (125 \pm 16)^\circ$ and $\cos\delta_D = \left(1.03^{+0.32}_{-0.18}\right)$ [23]. Thus, the remaining parameters are not as well constrained as r_D, and so it is preferable not to have to fix them when extracting γ. One possible solution is to include additional final states, such as the *CP* eigenstate K^+K^-. Then, there are two additional observables, the B^+ and B^- decay rates, and one additional unknown parameter, the constant of proportionality between the two decay rates. Thus, if

[1] Here, and throughout this thesis, the levels of *CP* violation and mixing in neutral D mesons are assumed to be negligible.

[2] Doubly Cabibbo Suppressed refers to the fact that the coupling strength of both flavour-changing vertices are proportional to the off-diagonal CKM elements of the first and second generations, whose magnitudes are approximately equal to the Cabibbo angle. Decays with no CKM suppression factors are referred to as *Cabibbo Favoured*.

r_D is fixed, then there are six parameters and six observable rates, enough to extract all the relevant parameters, including γ.

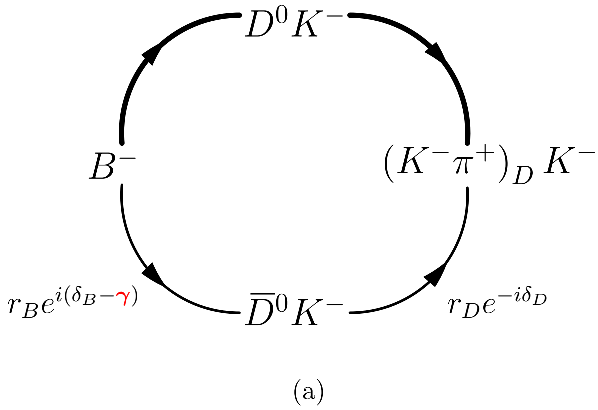


Figure 1.4: Interference diagrams of the decays (a) $B^\pm \to (K^\pm\pi^\mp)_D K^\pm$ and (b) $B^\pm \to (K^\mp\pi^\pm)_D K^\pm$. The thick lines represent the favoured transition. The relative magnitudes and phases are indicated for the relevant suppressed transitions.

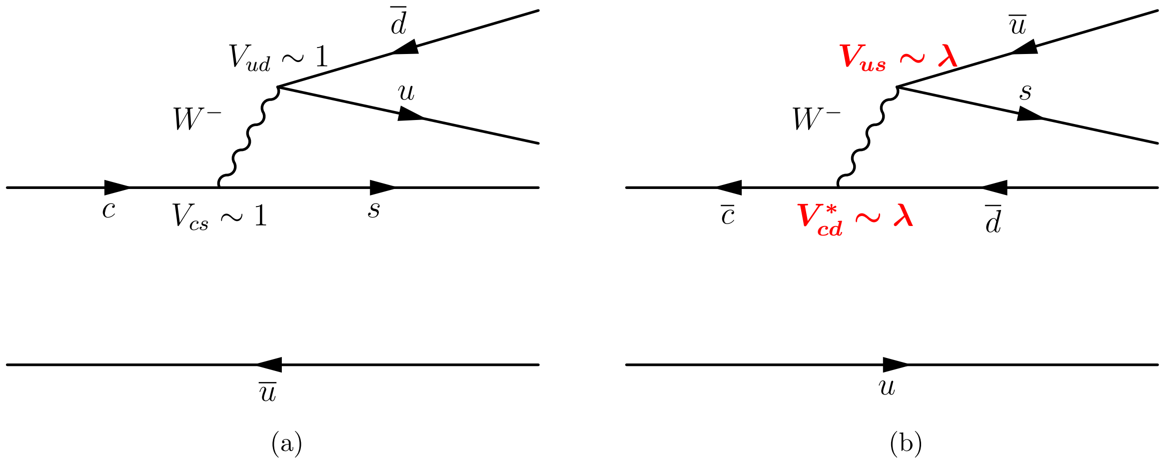


Figure 1.5: Feynman diagrams of the decays (a) $D^0 \to K^-\pi^+$ and (b) $\overline{D}^0 \to K^-\pi^+$. The ratio of the decay amplitudes has a magnitude, r_D, and phase, δ_D. The latter diagram is referred to as Doubly Cabibbo Suppressed because both of the vertices are suppressed by a factor $\approx \lambda$, where λ is the Wolfenstein parameter.

Experimentally, it is generally advantageous to consider the ratio of decay rates rather than the rates themselves, as many sources of experimental uncertainty will cancel. For example, for the $K^\pm\pi^\mp$ decay modes, the observables R_{ADS}, A_{ADS} and $A^{K\pi}_{fav}$ are useful alternatives to the raw decay rates, defined by

$$R_{ADS} = \frac{\Gamma\left(B^- \to (K^+\pi^-)_D K^-\right) + \Gamma\left(B^+ \to (K^-\pi^+)_D K^+\right)}{\Gamma\left(B^- \to (K^-\pi^+)_D K^-\right) + \Gamma\left(B^+ \to (K^+\pi^-)_D K^+\right)}\,, \tag{1.48a}$$

$$A_{ADS} = \frac{\Gamma\left(B^- \to \left(K^+\pi^-\right)_D K^-\right) - \Gamma\left(B^+ \to \left(K^-\pi^+\right)_D K^+\right)}{\Gamma\left(B^- \to \left(K^+\pi^-\right)_D K^-\right) + \Gamma\left(B^+ \to \left(K^-\pi^+\right)_D K^+\right)} \;, \tag{1.48b}$$

$$A_{fav}^{K\pi} = \frac{\Gamma\left(B^- \to \left(K^-\pi^+\right)_D K^-\right) - \Gamma\left(B^+ \to \left(K^+\pi^-\right)_D K^+\right)}{\Gamma\left(B^- \to \left(K^-\pi^+\right)_D K^-\right) + \Gamma\left(B^+ \to \left(K^+\pi^-\right)_D K^+\right)} \;. \tag{1.48c}$$

Substituting Eqs. (1.47a) to (1.47d) into Eqs. (1.48a) to (1.48c) gives these observables in terms of the unknown parameters:

$$R_{ADS} = \frac{r_B^2 + r_D^2 + 2r_B r_D \cos\left(\delta_B + \delta_D\right)\cos\gamma}{1 + r_B^2 r_D^2 + 2r_B r_D \cos\left(\delta_B - \delta_D\right)\cos\gamma} \approx r_B^2 + r_D^2 + 2r_B r_D \cos\left(\delta_B + \delta_D\right)\cos\gamma \;, \tag{1.49a}$$

$$A_{ADS} = \frac{2r_B r_D \sin\left(\delta_B + \delta_D\right)\sin\gamma}{r_B^2 + r_D^2 + 2\cos\left(\delta_B + \delta_D\right)\cos\gamma} \approx \frac{2r_B r_D \sin\left(\delta_B + \delta_D\right)\sin\gamma}{R_{ADS}} \;, \tag{1.49b}$$

$$A_{fav}^{K\pi} = \frac{2r_B r_D \sin\left(\delta_B - \delta_D\right)\sin\gamma}{1 + r_B^2 r_D^2 + 2r_B r_D \cos\left(\delta_B - \delta_D\right)\cos\gamma} \approx 2r_B r_D \sin\left(\delta_B - \delta_D\right)\sin\gamma \;. \tag{1.49c}$$

It is preferable to consider all possible final states when extracting γ in order to over-constrain the parameters, and so increase sensitivity to γ. This includes the multi-body final states such as $K^\pm\pi^\mp\pi^\pm\pi^\mp$, the treatment of which will now be discussed.

1.6.2 Multi-body extension to ADS method

The ADS method can be extended to consider multi-body decays, such as $K^\pm\pi^\mp\pi^\pm\pi^\mp$. However, the situation is complicated by the fact that the D^0 and $\overline{D}^0$ can decay to these final states via several intermediate resonances, introducing a multitude of strong phases and amplitude ratios. The implementation of this method therefore requires an extensive modelling of the resonance structure of the decays. However, Atwood and Soni indicated that an alternative is to modify the formulae for the decay rates in Eqs. (1.47a) to (1.47d) by a *coherence factor* C_f, reflecting the interference structure of the decay modes [28].

First, the decay amplitude at a single point $\mathbf{x}$ in the D decay phase space is considered. For the B^- decaying to a multi-body "ADS-like" final state, f_D, the decay rate is given by

$$\Gamma\left(B^- \to (f_D)_D K^-\right)(\mathbf{x}) \propto \left|A_f(\mathbf{x})\right|^2 + r_B^2\left|\bar{A}_f(\mathbf{x})\right|^2 + 2r_B\left|A_f(\mathbf{x})\right|\left|\bar{A}_f(\mathbf{x})\right|\cos\left(\delta_B + \xi(\mathbf{x}) - \gamma\right) \;, \tag{1.50}$$

where $A_f(\mathbf{x})$ and $\bar{A}_f(\mathbf{x})$ are the decay amplitudes of the D^0 and $\overline{D}^0$ decaying to f_D via the resonance defined by point $\mathbf{x}$, and $\xi(\mathbf{x})$ is the relative phase difference between the two decays, defined by

$$\xi(\mathbf{x}) = \arg\left(A_f^*(\mathbf{x})\bar{A}_f(\mathbf{x})\right) \ . \tag{1.51}$$

Integrating Eq. (1.50) over the full D decay phase space gives

$$\Gamma\left(B^- \to (f_D)_D K^-\right) = A_f^2 + r_B^2 \bar{A}_f^2 + 2 r_B A_f \bar{A}_f C_f \cos\left(\delta_B - \delta_D^f - \gamma\right) \ , \tag{1.52}$$

where A_f and $\bar{A}_f$ are the total D^0 and $\overline{D}^0$ decay amplitudes respectively, C_f is the aforementioned coherence factor, and δ_D^f is the average strong phase difference. These four quantities are defined by

$$A_f^2 = \int |A_f(\mathbf{x})|^2 \, d\mathbf{x} \ ; \qquad \bar{A}_f^2 = \int \left|\bar{A}_f(\mathbf{x})\right|^2 d\mathbf{x} \ , \tag{1.53a}$$

$$C_f e^{i\delta_D^f} = \frac{\int |A_f(\mathbf{x})| \left|\bar{A}_f(\mathbf{x})\right| e^{i\xi(\mathbf{x})} d\mathbf{x}}{A_f \bar{A}_f} \ . \tag{1.53b}$$

From the definition in Eq. (1.53b), it is clear that C_f can take values between zero and unity. If the decay amplitudes are dominated by a single resonance, then C_f will be close to 1, whilst if C_f is small, then this indicates that a multitude of decay amplitudes with differing strong phases all contribute to the overall decay amplitude.

For $K^\pm\pi^\mp\pi^\pm\pi^\mp$ decays, the amplitudes are given in terms of the coherence factor $C_{K3\pi}$, and average strong phase difference $\delta_D^{K3\pi}$ by

$$\Gamma\left(B^\pm \to \left(K^\pm\pi^\mp\pi^\pm\pi^\mp\right)_D K^\pm\right) \propto \left(1 + \left(r_B r_D^{K3\pi}\right)^2 + 2 r_B r_D^{K3\pi} C_{K3\pi} \cos\left(\delta_B - \delta_D^{K3\pi} \mp \gamma\right)\right) \ , \tag{1.54a}$$

$$\Gamma\left(B^\pm \to \left(K^\mp\pi^\pm\pi^\mp\pi^\pm\right)_D K^\pm\right) \propto \left(r_B^2 + (r_D^{K3\pi})^2 + 2 r_B r_D^{K3\pi} C_{K3\pi} \cos\left(\delta_B + \delta_D^{K3\pi} \mp \gamma\right)\right) \ , \tag{1.54b}$$

where $r_D^{K3\pi}$ is the ratio of the magnitude of the suppressed and favoured decays. As with the $K^\pm\pi^\mp$ modes, it is useful to consider the ratio of decay rates. The equivalent observables,

$R_{ADS}^{K3\pi}$, $A_{ADS}^{K3\pi}$ and $A_{fav}^{K3\pi}$ are defined by

$$R_{ADS}^{K3\pi} \approx r_B^2 + (r_D^{K3\pi})^2 + 2r_B r_D^{K3\pi} C_{K3\pi} \cos\left(\delta_B + \delta_D^{K3\pi}\right) \cos\gamma \ , \tag{1.55a}$$

$$A_{ADS}^{K3\pi} \approx \frac{2r_B r_D^{K3\pi} C_{K3\pi} \sin\left(\delta_B + \delta_D^{K3\pi}\right) \sin\gamma}{R_{ADS}^{K3\pi}} \ , \tag{1.55b}$$

$$A_{fav}^{K3\pi} \approx 2r_B r_D C_{K3\pi} \sin\left(\delta_B - \delta_D^{K3\pi}\right) \sin\gamma \ . \tag{1.55c}$$

The CLEO-c experiment has determined the coherence factor and average strong phase difference to be $C_{K3\pi} = 0.33^{+0.20}_{-0.23}$ and $\delta_D^{K3\pi} = \left(114^{+26}_{-23}\right)^\circ$ [29]. This indicates that the coherence factor in these modes is relatively small, although the uncertainties are very large. The amplitude ratio $r_D^{K3\pi}$ is comparable to that of the $K^\pm\pi^\mp$ decays [30] ($(r_D^{K3\pi})^2 = (1.18^{+0.10}_{-0.09}) \tan^4\theta_C$), so even with a reasonably low coherence factor, $B^\pm \to (K^\mp\pi^\pm\pi^\pm\pi^\pm)_D K^\pm$ decays should provide good sensitivity to γ. Furthermore, even if the decay-rate asymmetry, $A_{fav}^{K3\pi}$, is small due to an unfavourable coherence factor, $B^\pm \to (K^\mp\pi^\pm\pi^\mp\pi^\pm)_D K^\pm$ decays can still reduce the uncertainty on γ indirectly, since $R_{ADS}^{K3\pi}$ will have heightened sensitivity to the poorly-measured r_B. Thus, $B^\pm \to (K^\mp\pi^\pm\pi^\mp\pi^\pm)_D K^\pm$ decays will be an important part of the tree-level determination of γ at LHCb.

1.6.3 Sensitivity to γ from $B \to D\pi$ processes

The strange quark can be exchanged for a down quark in Fig. 1.3 and the interference of the two processes will still contain information about γ. This suggests that analysing $B \to D\pi$ decays is also a potential method for extracting γ. Figure 1.6 shows the Feynman diagrams for B^- decays to $D^0\pi^-$ and $\overline{D}^0\pi^-$. Comparing Figs. 1.3a and 1.6a, it can be seen that the magnitude of the latter decay amplitude is around $^1/_\lambda \approx 4.5$ (to $\mathcal{O}(\lambda^3)$) times larger than the former. However, there is an additional λ^2 suppression between the two $B \to D\pi$ diagrams, so r_{B_π} is smaller than r_B by around a factor $\lambda^2 \approx 0.05$. This large suppression of the colour-suppressed mode means that any sensitivity to γ from either the CP eigenstates or the "ADS-like" final states such as $K^\pm\pi^\mp$ and $K^\pm\pi^\mp\pi^\pm\pi^\mp$ is expected to be small compared to $B \to DK$ decays.

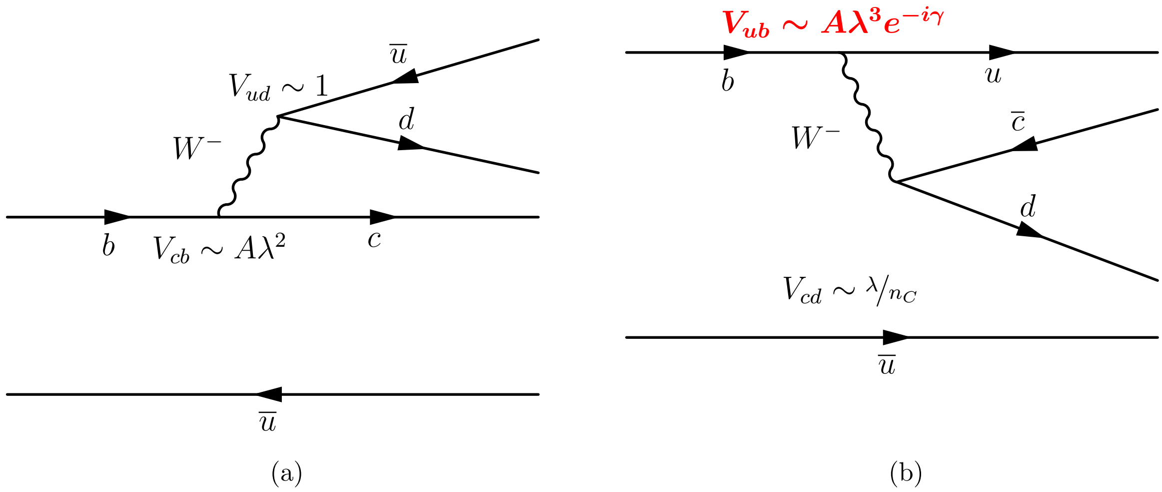


Figure 1.6: Feynman diagrams of the decays (a) $B^- \to D^0\pi^-$ and (b) $B^- \to \overline{D}^0\pi^-$. The ratio of the decay amplitudes has a magnitude, r_{B_π}, and phase, $\delta_{B_\pi} - \gamma$.

When considering $B \to D\pi$ and $B \to DK$ decays where the D decays to the same final state f_D, there is an additional observable that can be measured which contains information on γ, namely the ratio of branching fractions $R^{f_D}_{K/\pi}$, defined by

$$R^{f_D}_{K/\pi} = \frac{\Gamma\left(B^- \to (f_D)_D K^-\right) + \Gamma\left(B^+ \to (f_D)_D K^+\right)}{\Gamma\left(B^- \to (f_D)_D \pi^-\right) + \Gamma\left(B^+ \to (f_D)_D \pi^+\right)} \ . \tag{1.56}$$

For $B^\pm \to (K^\pm\pi^\mp)_D\, K^\pm$ and $B^\pm \to (K^\pm\pi^\mp)_D\, \pi^\pm$, and the equivalent multi-body modes such as $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, K^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, \pi^\pm$, Eq. (1.56) can be written in terms of γ as

$$R^{f_D}_{K/\pi} = \mathcal{B}_{DK}/\mathcal{B}_{D\pi} \frac{1 + \left(r_B r^f_D\right)^2 + 2 r_B r^f_D C_f \cos\left(\delta_B - \delta^f_D\right)\cos\gamma}{1 + \left(r_{B_\pi} r^f_D\right)^2 + 2 r_{B_\pi} r^f_D C_f \cos\left(\delta_{B_\pi} - \delta^f_D\right)\cos\gamma} \ , \tag{1.57}$$

where r^f_D and δ^f_D are the amplitude ratio and strong phase difference for D final state f_D, C_f is the coherence factor for state f_D, defined in Eq. (1.53b), which is equal to unity by definition for $f_D = K^\pm\pi^\mp$, and $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ is defined by

$$\mathcal{B}_{DK}/\mathcal{B}_{D\pi} \equiv \Gamma(B^- \to D^0K^-)/\Gamma(B^- \to D^0\pi^-) = \Gamma(B^+ \to \overline{D}^0K^+)/\Gamma(B^+ \to \overline{D}^0\pi^+) \ . \tag{1.58}$$

Since the second and third terms in the numerator and denominator on the right-hand side of Eq. (1.57) are heavily suppressed compared to the first terms, $R^{f_D}_{K/\pi} \approx \mathcal{B}_{DK}/\mathcal{B}_{D\pi}$[1] for the favoured modes $B^{\pm} \to (K^{\pm}\pi^{\mp})_D h^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ (where $h^{\pm}$ denotes a kaon or pion). This means that whilst $R^{f_D}_{K/\pi}$ has negligible sensitivity to γ for the favoured modes, it is a useful way of determining $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$, can be measured using multiple decay modes and does not rely on obtaining large samples of suppressed decays. The measurement of this observable using data collected at LHCb from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ decays, as well as the decay-rate asymmetry, $A^{K3\pi}_{fav}$, is discussed in Chap. 4.

[1] Hereby, this approximation is assumed throughout this thesis.

Chapter 2

The LHCb detector

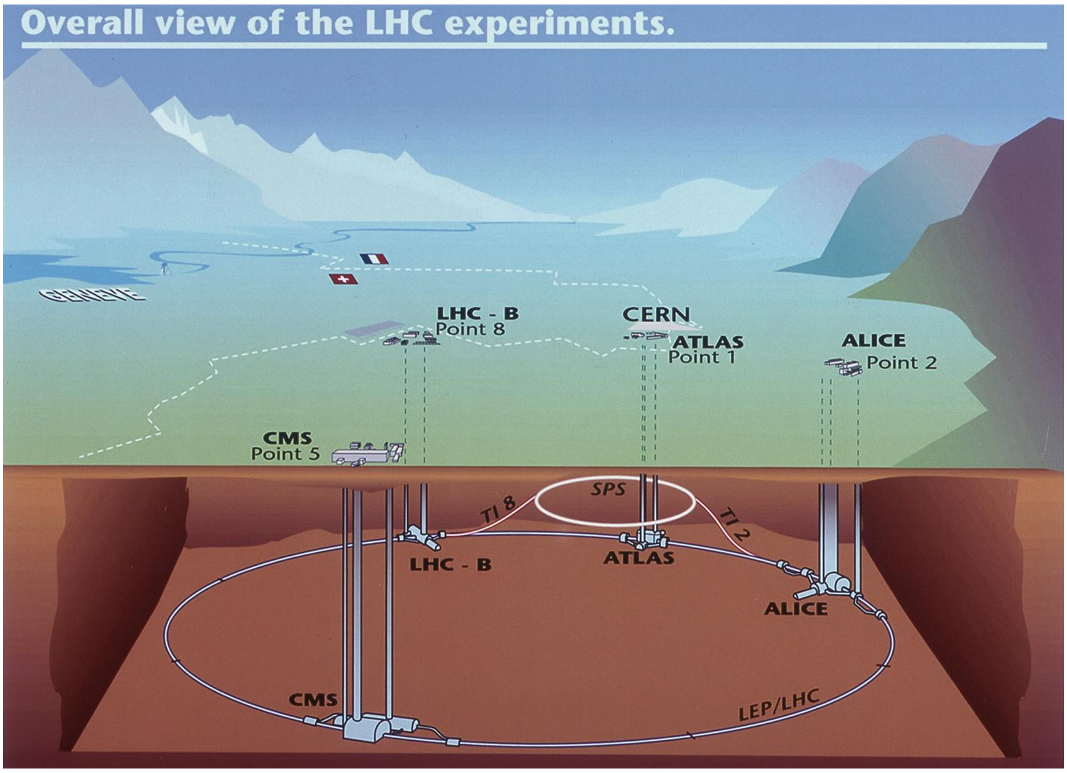
This chapter begins with a description of the Large Hadron Collider and its various detectors. An explanation of the LHCb detector and its various hardware components and data acquisition system is then given, focusing on those parts of the detector which are most relevant to the physics analyses discussed in subsequent chapters. The LHCb Ring Imaging Cherenkov detectors in particular are described in detail. An explanation of how charged and neutral particle identification is performed in LHCb is presented. The chapter concludes with a summary of the LHCb trigger and online systems, and the detector software used to produce simulated data, perform the event reconstruction and analyse the reconstructed data.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC), located at the European Organisation for Nuclear Research (CERN)[1] is the world's largest and most-powerful particle accelerator, designed to collide two beams of protons at a centre-of-mass energy ($\sqrt{s}$) up to 14 TeV, with a maximum instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [31]. The LHC comprises two rings of counter-rotating beams, constructed in the same 27 km tunnel that originally housed the Large Electron Positron (LEP) collider. The protons are injected from the Super Proton Synchrotron (SPS) into the main rings with an energy of 450 MeV. A series of radio-frequency (RF) cavities and superconducting dipole magnets, located around the collider, accelerate and bend the beams respectively. To accelerate the beams to the design energy whilst keeping

[1] CERN stands for *Consiel Europèene pour la Recherche Nucléaire*, the original French title.

them in a stable orbit, a magnetic field strength of 8.3 T is required. Each beam consists of multiple high particle-density regions known as *bunches*, with each bunch nominally spaced by 25 ns. This allows multiple collisions to be produced in the same beam cycle. There are four regions along the LHC tunnel in which the two beams share a common beam pipe and here the beams collide. These correspond to the *interaction points* (IPs) of the four main experiments: the two general-purpose experiments ATLAS and CMS, the heavy-ion experiment ALICE, and the heavy flavour experiment LHCb which is outlined in the following sections. A pictorial representation of the LHC is shown in Fig. 2.1.



© CERN (1999)

Figure 2.1: Diagram of the LHC, showing the geographical location of the four main experiments. The smaller ring close to the ATLAS detector represents the SPS.

2.2 Overview of LHCb

The LHC is the most copious source of B hadrons in the world, with an expected $b\bar{b}$ cross-section of $\sim 500\,\mu\text{b}$ at $\sqrt{s} = 14\,\text{TeV}$ [32]. Even at half the design centre-of-mass energy, $b\bar{b}$ pairs are produced with a cross-section of $284 \pm 20 \pm 49\,\mu\text{b}$ [33]. LHCb utilises the large production cross-section of both $b\bar{b}$ and $c\bar{c}$ pairs to make precision measurements of the decays of B and D hadrons, in particular, the search for and study of CP-violating processes and rare decays.

LHCb is a single-arm spectrometer, orientated in the forward direction, with an approximate angular coverage of $10 - 300(250)$ mrad in the magnetic bending (non-bending) plane. This design utilises the fact that at LHC energies, the B hadrons produced by both the b and $\bar{b}$ tend to be heavily boosted in the forward or backward direction (along the beam axis) in the laboratory frame.

LHCb is a "precision" experiment. Its physics requirements make it necessary to accurately reconstruct the full decay chain of any forward-going B or D hadron. At the LHC's optimal luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, there is on average significantly more than one collision per bunch crossing. The resulting track multiplicity is far too high to be able to cleanly reconstruct B and D decays. As such, LHCb is designed to operate with an average peak luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which ensures that the data collected are dominated by events with one pp interaction per bunch crossing. Since the ATLAS and CMS experiments require the maximum luminosity available, LHCb reduces its luminosity by *defocusing* the beams at the IP.

The physics programmes at LHCb require the reconstruction of exclusive leptonic, semi-leptonic and fully hadronic decays, and the ability to separate these signals from various background processes. As such, LHCb comprises a system of sub-detectors, responsible for a number of crucial tasks:

- **Proper-time resolution**: Time-dependent CP violation analyses require a good knowledge of the B or D hadron's proper decay time. This requires that the production (primary) and decay (secondary) vertices be well measured. This measurement is achieved by the *Vertex Locator* (VELO).

- **Energy and momentum resolution**: To distinguish signal from background processes, it is important that LHCb has a good invariant mass resolution, which means that the momentum and energy of the tracks must be well measured. The LHCb dipole magnet, the tracking system and the calorimeters are all necessary to perform this task.

- **Particle identification (PID)**: For many B and D decays, there exist decays that are topologically similar to the signal, often with significantly larger branching fractions, whose invariant-mass distributions overlap that of the signal. It is therefore vital that LHCb has excellent particle identification, particularly between kaons and pions but also between other species of particles. Discrimination between different types of charged hadrons is achieved primarily by the two Ring Imaging Cherenkov (RICH) detectors, RICH1 and RICH2, whilst the calorimeters and muon chambers are used to identify electrons and photons and muons respectively.

- **Fast event filtering**: Hadron colliders produce primary interactions via quark/gluon fragmentation and this generally results in high multiplicity events. At design luminosity, LHCb is expected to have a *visible interaction*[1] rate of 10 MHz. Of these, only about 100 kHz comprises $b\bar{b}$ pairs and only $\approx 15\%$ of these have at least one B meson with all decay products within the LHCb acceptance. It is essential to have rapid response event filtering without incurring dead time for the next bunch crossing. The accepted data need to be further reduced to about 2 kHz before the events are written to offline storage. This requires a combination of hardware and software triggers, retaining only events likely to contain decays of interest.

The various sub-detectors are indicated in Fig. 2.2. A description of these sub-detectors is given in the subsequent sections of this chapter.

2.3 Tracking and vertexing

This section describes the LHCb detector components which are used to provide information of the production and decay vertices of particles and to track them as they traverse the LHCb detector. An explanation of how tracks are reconstructed in the LHCb software is also given.

[1] Events which contain at least two *reconstructible* charged particles (*i.e.* sufficient hits in the VELO and inner tracker).

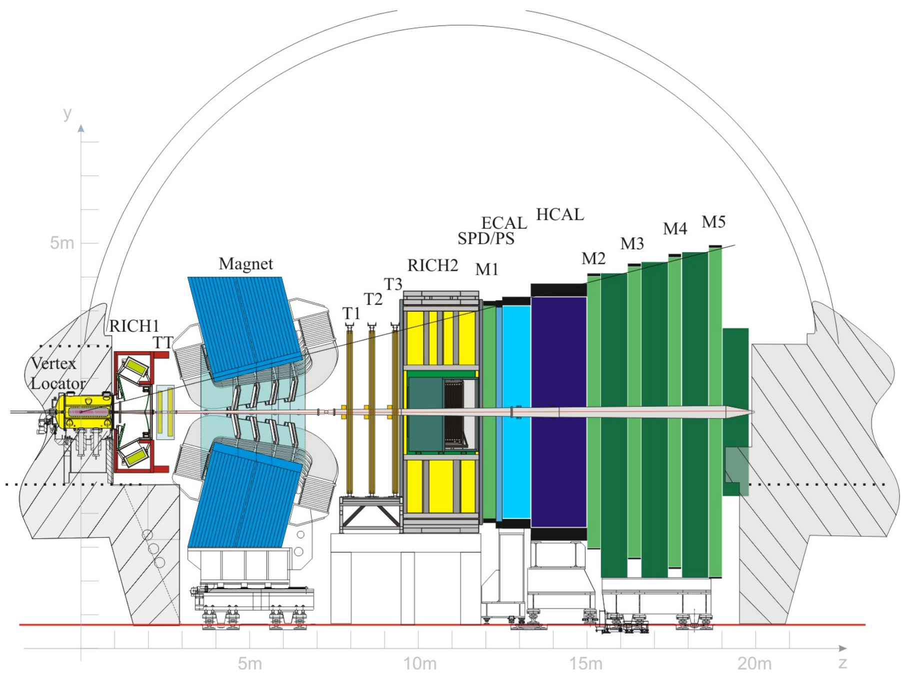


Figure 2.2: Schematic of the LHCb detector viewed in the non-bending plane. The y and z axes shown define the LHCb global co-ordinate system convention, along with the x axis, which points into the page.

2.3.1 VELO

One of the key challenges of LHCb was to design a system which would allow measurement of the position of tracks very close to the interaction region, and therefore accurately reconstruct the primary and secondary vertices. This is essential for precision lifetime measurements and to identify D and B hadrons, which tend to have secondary vertices displaced from the PV. To reach an impact parameter resolution of 20 μm with respect to the PV, the vertex detector silicon sensor modules must be located very close to the beam. To avoid radiation damage at injection, when the beam spread is large and the orbit is unstable, the VELO modules are able to be retracted away from the beam pipe.

The top diagram of Fig. 2.3 shows a cross-section of the VELO system in the fully closed position. As can be seen, the sub-detector comprises two sets of 21 modules oriented perpendicular to the beam pipe. The VELO modules are contained in a vacuum vessel to minimise the material traversed by charged particles before hitting the VELO sensors. The two long lines on the diagram indicate the minimum and maximum angle with respect to the z-axis in which a track can be produced and register hits in all the VELO sensors in the forward direction ($15 - 60\,\mathrm{mrad}$) whilst the shorter line shows an example of a track which is *reconstructible* by the VELO (has hits in three VELO stations) but is outside the angular acceptance of the tracking stations. The different track types considered by the LHCb reconstruction software can be found in Sec. 2.3.4.

During data-taking, the beams are well constrained in the x and y directions at the IP (the RMS spread of the beams should be less than $100\,\mu\mathrm{m}$ [32]). However, the resolution of the IP in the beam (z) direction is significantly larger. The precise uncertainty depends on the crossing-angle of the beams, the bunch size and β^{*}[1], but under nominal run conditions, the IP resolution in the z direction is expected to be $\approx 5.3\,\mathrm{cm}$, as indicated by the shaded area of Fig. 2.3 (top). Due to the lateral uncertainty, several VELO modules are placed behind the "nominal" interaction point (defined as $z = 0$ in the LHCb co-ordinate system) to encompass the entire luminous region. The four modules on the far left of the diagram are known as the *pile-up veto* stations, and are used by the *L0 trigger*, discussed in Sec. 2.8.1, to reject events with multiple pp interactions.

The bottom diagrams of Fig. 2.3 show the front face of two of the VELO half-modules. In the closed position (left diagram), the two modules are displaced from the beam pipe by as little as 8 mm. In this position, the modules can be seen to overlap. This is to ensure that the sensors cover the full angular acceptance. During injection, the modules are in the fully open position (right diagram) to minimise radiation damage. The modules can retract up to 6 cm in the x direction, and 8.4 cm in the y direction.

Each module comprises two silicon strip detectors, referred to as the R- and ϕ-sensors[2]. As indicated by the names, the R-sensors consist of a series of silicon strips orientated in concentric semi-circles about the centre of the beam pipe, allowing detection of the radial component of the tracks, whilst the strips in the ϕ-sensors are distributed radially around the

[1] The amplitude function, β, represents how strongly focused the beam is, which is determined by the field strength of the quadrupole magnets. β is defined as the distance (along the beam direction) from a focal point at which the beam size is twice as large as its size at the focal point. An important parameter of the beams is β^{*}, which is the value of β at the IP.

[2] With the exception of the pile-up veto modules, which contain only R-sensors.

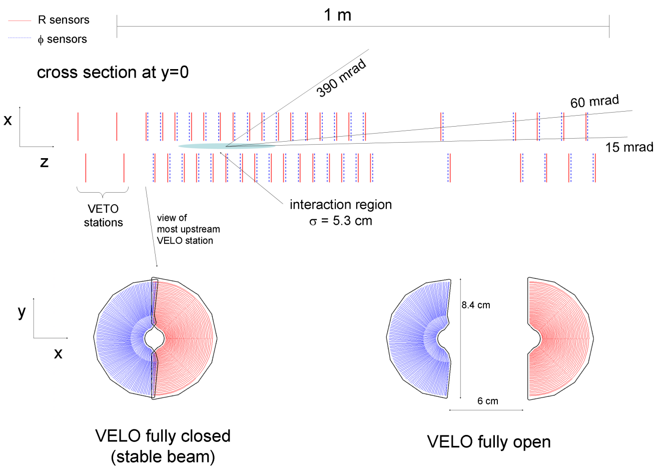


Figure 2.3: Above, cross-section of the VELO in the magnet bending plane at $y = 0$. The modules are shown in the fully closed position. The two pile-up veto stations are indicated on the far left of the diagram. Below, the front face of the first modules is shown in the fully closed and fully open positions.

module in order to measure the position of the track in the ϕ direction[1]. The final component needed to fully reconstruct the vertex and track locations is the z position. This is calculated by knowledge of the position of the module along the beam axis direction. The PV has a typical resolution of around 0.01 mm in the transverse directions, and 0.05 mm in the z direction [34].

2.3.2 The LHCb dipole magnet

As can be seen in Fig. 2.2, the dipole magnet is located between the TT and *silicon tracking stations* (discussed in Sec. 2.3.3). It comprises two saddle shaped iron yokes, one placed either side of the beam pipe in the vertical (y) direction, as indicated in Fig. 2.4. The magnetic coils are constructed from a series of "pancakes" wound from aluminium conducting cable. The LHCb dipole magnet has a maximum integrated magnetic field of 4 T$\cdot$ m and its orientation means that the Lorentz force is concentrated in the (x, z) (horizontal) plane.

An important feature of the magnet is the ability to revert the direction of the magnetic field between fills, which means that positive and negative charges will bend in the opposite

[1] The VELO uses a cylindrical co-ordinate system (R, ϕ, z) rather than the Cartesian co-ordinate system used by the rest of LHCb to allow fast reconstruction of vertices and tracks in the LHCb trigger.

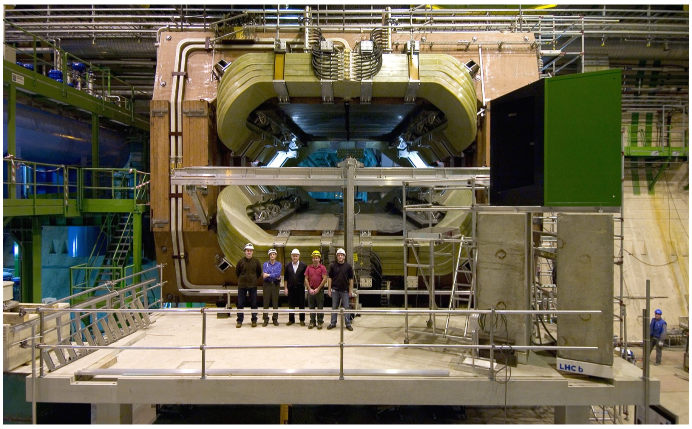


Figure 2.4: Image of the LHCb dipole magnet shortly after installation in the LHCb cavern. The interaction point is behind the magnet.

direction upon field reversal. This can be used to measure the level of detector left-right asymmetry, where the detection efficiency in the left side of the detector is different from the right (*e.g.* due to material differences). This is crucial for many *CP* asymmetry measurements, since the expected level of asymmetry due to *CP* violation can be very small.

2.3.3 Tracking stations

Aside from the VELO, tracking in LHCb is provided by four tracking stations: the TT located upstream of the magnet, and three further stations downstream of the magnet labelled T1, T2 and T3 in Fig. 2.2. The T1-T3 tracking stations (commonly referred to as the T-stations) comprise two separate sub-detectors: the Inner Tracker (IT), which is situated close to the beam and uses the same tracking technology as the TT, and the Outer Tracker (OT), which covers the remaining geometrical acceptance. The TT and IT sub-detectors form the Silicon Tracker (ST) system.

Silicon Tracker

Both the TT and IT comprise a series of silicon microstrip sensors, constructed in four separate detection layers. In the first and fourth layers, the strips are orientated vertically, whilst the second and third layers are rotated by a -5% and $+5\%$ stereo angle respectively. Each layer has a strip pitch of $\sim 200\,\mu\text{m}$, which corresponds to a spatial resolution of approximately 50 μm for a single hit. The layers of the TT are split into pairs separated by ≈ 27 cm in the z direction in order to aid the reconstruction of tracks.

The TT, shown in Fig. 2.5a, covers the full LHCb geometric acceptance and has two key roles. The primary purpose of the TT is to reconstruct long-lived neutral particles such as K^0_S and Λ, since they tend to decay outside of the VELO. K^0_S and Λ decays play an important role in evaluating the performance of the PID algorithms, as discussed in Chap. 5. The TT is also used to track particles with very low momentum that get swept out of the detector acceptance by the magnet before reaching the T-stations. This is important for some exclusive decays such as $D^{*+} \to \pi^+ D^0(K^-\pi^+)$, since most of the momentum from the D^* goes to the D^0, so the pion is generally very soft.

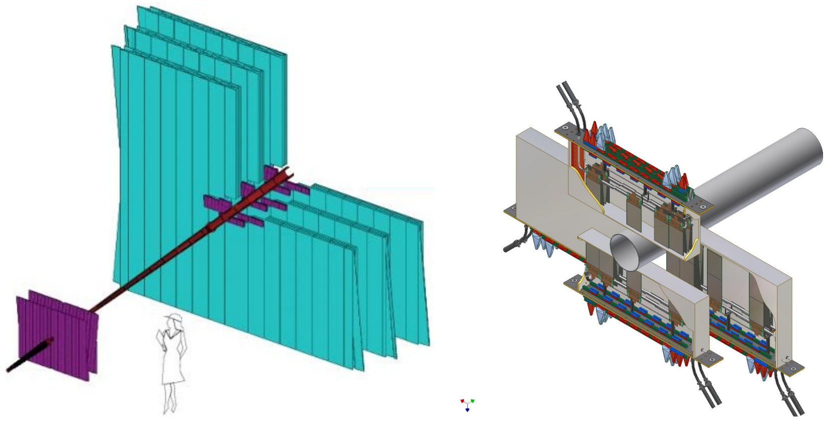
The IT covers the highest occupancy region close to the beam pipe, where particle fluxes are expected to be up to $5 \times 10^5\,\mathrm{cm^{-2}\,s^{-1}}$ [35]. Although the IT covers only 1.3% of the region covered by the tracking stations, approximately 20% of charged particles produced close to the nominal IP that pass through the T-stations do so in the area in which the IT stations are located.

Each of the three IT stations comprises four detector boxes, as indicated in Fig. 2.5b, and each box contains four detection layers. Adjacent modules are shifted by 4 mm along the beam direction, and overlap by 3 mm in the horizontal direction. The overlap ensures there are no gaps in the detector acceptance, and the staggered configuration means that the modules can be aligned relative to one another.

Outer Tracker

The Outer Tracker (OT) system, shown in Fig. 2.5a, consists of a series of multi-wire proportional chambers (MWPCs). Each OT station comprises four detection layers, rotated by the same stereo angles as the ST. Each layer consists of a number of gas-tight modules comprising two layers of staggered drift-tubes. In order to ensure a drift time of less than 50 ns (two bunch crossings) and a drift-coordinate resolution of 200 μm or better, the drift-tubes have an internal diameter of 4.9 mm and contain an admixture of Argon (70%) and CO_2

(30%).



(a) (b)

Figure 2.5: Pictorial representation of (a) the four tracking stations and (b) a close-up view of one of the IT stations showing the beam pipe surrounded by the four IT detector boxes. In (a), the ST sub-detectors are shown in purple (TT on the left of the figure and the IT stations on the right) whilst the OT stations are indicated in light blue.

2.3.4 Track reconstruction

The way in which track reconstruction in LHCb is handled depends on the type of track considered. The LHCb tracks can be classified into one of five categories:

- **VELO tracks** have hits only in the VELO sensors. They are typically backward tracks or tracks produced at large angles, as indicated by Fig. 2.3. VELO tracks are not used directly in physics analyses, but are used as part of the PV reconstruction.
- **T tracks** have hits only in the T-stations. Fully reconstructed T tracks are principally produced by secondary interactions, and are used in the RICH2 global pattern recognition.
- **Downstream tracks** traverse the TT and T-stations, but not the VELO. They are primarily decay products of long-lived particles such as K^0_S and Λ. Since downstream tracks pass through all tracking stations, they have a reasonable momentum resolution.
- **Upstream tracks** have hits only in the VELO and TT. They are typically low momentum tracks such as the pion from $D^{*+} \to D^0\pi^+$, which are swept out of

acceptance by the dipole magnet before reaching the T-stations. Since tracks of this type pass through RICH1, they can produce Cherenkov photons if the momentum is above threshold. This means that in principle upstream tracks can be used in physics analyses; however, the momentum resolution of the tracks is obviously poor. Upstream tracks can also be used to understand backgrounds in the RICH PID algorithms.

- **Long tracks** traverse the full tracking system: the VELO, TT and T-stations, and are therefore the most important tracks for B and D hadron reconstruction.

The five track categories are illustrated in Fig. 2.6.

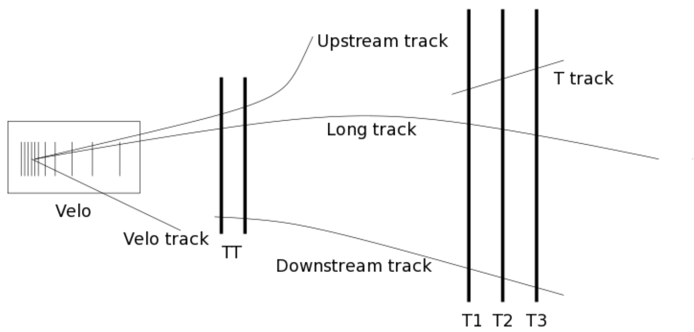


Figure 2.6: Illustration of the track types considered by the LHCb reconstruction software. The tracks are shown in the magnet bending plane.

Tracks with hits in multiple sub-detectors (*i.e.* downstream, upstream and long tracks) are constructed from VELO and/or T-track *seeds*[1]. The reconstruction algorithms then tries to find matching hits or track seeds in the other sub-detectors. In the case of long tracks, the track reconstruction can be performed either by the *forward-tracking* algorithm, which starts from VELO seeds, and searches for corresponding hits in the T-stations (and TT) or the *track-matching* algorithm, which takes reconstructed VELO and T-station track segments and tries to match them to one another. After tracks have been identified, their trajectories are refitted using a *Kalman fitter* [36]. This corrects for multiple scattering and radiative energy losses ($\frac{dE}{dx}$). The quality of the reconstructed tracks is monitored by the χ^2 of the track fit.

[1] Seeds are segments of candidate tracks from regions of the detector where the magnetic field is low.

Background in track pattern-recognition

There are two types of background which can occur in the track pattern-recognition algorithms: *ghosts* and *clones*.

Ghosts are tracks (or track segments) comprising hits which are not associated with a "true" particle in the event. These fake tracks can be the result of either noise or material interactions in the sub-detector(s), or *spill-over* (signal sub-detector hits from a previous event).

Clones are two or more tracks (or track segments) constructed from the same particle. This is generally identified by tracks sharing a large proportion of hits.

Clones are generally dealt with by applying a *clone-killer* algorithm, in which a quality requirement is applied to determine the "best" track amongst the clone tracks. The best track is usually defined as the track constructed from the most number of hits (or the lowest fit χ^2 if two or more tracks have the same number of hits) [37]. Such a clone-killer algorithm is applied by default in the LHCb reconstruction. Tracks which are likely to be clones can either be removed, or flagged as "clone-like" and stored for later analysis.

In LHCb, it is possible to have two tracks which are clones but do not share common hits, since the VELO reconstruction may split a particle trajectory into two tracks: one comprising hits from the forward stations and the other from the remaining stations. This type of clone is not picked up by the standard clone killer algorithm. Instead, these clones can be dealt with by selecting tracks based on the *Kullback-Leibner* (KL) distance [38]; if the KL distance is small, then the two tracks are likely to be clones of one another. As with the standard clone killer, the track with the most hits (or smallest fit χ^2) is generally considered the best track. The KL distance to be used as an additional clone killer by removing all tracks with values of KL distance in the range $[0, 5000]$.

Track reconstruction efficiency

The reconstruction efficiency is not a constant, but varies as function of track momentum. The efficiency of tracks from $B^0 \to J/\psi\, K_S^0\,(\pi^+\pi^-)$ MC decays (where the J/ψ decays to either

$\mu^+\mu^-$ or e^-e^-) is $\sim 94\%$ on average for long tracks with $|\vec{p}| > 10\,\mathrm{GeV}/c$ [32] and $\sim 83\%$ for downstream tracks with $|\vec{p}| > 5\,\mathrm{GeV}/c$ [39]. However, these efficiencies drop off rapidly as the track momentum decreases due to the increased level of multiple scattering. The momentum resolution of long tracks from $B^0 \to J/\psi\, K^0_\mathrm{S}$ MC decays normalised to the track momentum ($\sigma_{|\vec{p}|}/\,|\vec{p}|$) as a function of momentum and the impact parameter (IP) resolution (σ_{IP}) as a function of $1/p_\mathrm{T}$, where p_T is the *transverse momentum* of the track[1] are shown in Fig. 2.7. The linear dependence of the IP resolution is parametrised by $\sigma_{IP} = 14\,\mu\mathrm{m} + 35\,\mu\mathrm{m}/\mathrm{p_T}$, where p_T is expressed in units of GeV/c.

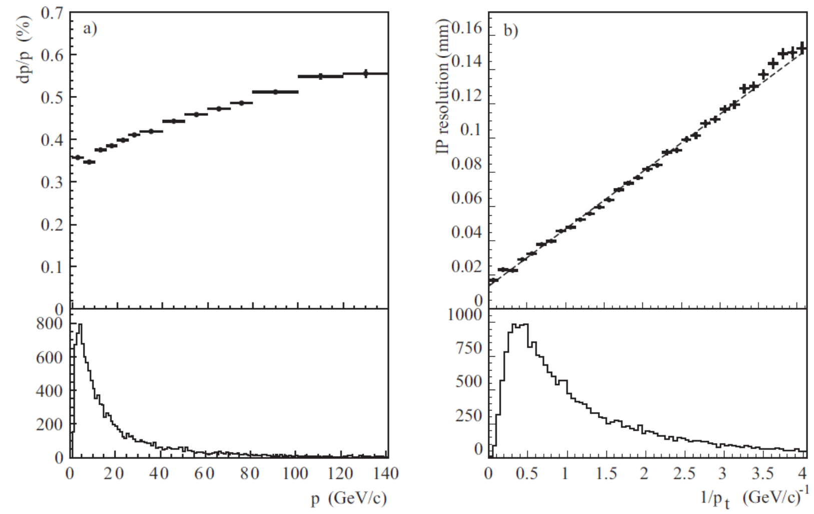


Figure 2.7: Resolution of track parameters from $B^0 \to J/\psi\, K^0_\mathrm{S}$ MC decays: (a) momentum resolution versus momentum, and (b) IP resolution versus $1/p_\mathrm{T}$. The IP resolution is determined by the quadrature sum of the uncertainties of the three projections of the IP (x, y and z in the LHCb global co-ordinate frame). The bottom plots show the (a) momentum spectrum and (b) $1/p_\mathrm{T}$ spectrum of the B^0 decay products for comparison. Results are shown for long tracks only.

2.4 Calorimetry system

The LHCb calorimetry system, as indicated by Fig. 2.2, comprises an electronic calorimeter (ECAL) and a hadronic calorimeter (HCAL), situated downstream of the detector, after RICH2. The calorimeters have several important roles. Firstly, they are used to calculate the transverse energies (E_T) of the hadrons, electrons and photons in the event, which are used

[1] The transverse momentum of a particle is defined by $p_\mathrm{T} = \sqrt{p_x^2 + p_y^2}$, where p_x and p_y are the two components of the momentum transverse to the beam direction.

in the hardware trigger system to provide a fast decision as to whether to keep or reject an event (see Sec. 2.8.1 for details). The system also provides identification of electrons, photons and hadrons, and measures their energies and positions. To identify electrons from the large pion backgrounds, two additional detectors are placed in front of the ECAL: the preshower (PS) and the scintillator pad detector (SPD).

The PS is located directly upstream of the ECAL, with the SPD in front of this. Sandwiched between them is a 15 mm lead converter, corresponding to roughly 2.5 radiation lengths ($2.5X_0$). The SPD and PS comprise a series of high granularity scintillator pads, with the photons from each pad transmitted to a multi-anode photomultiplier tube (MaPMT) via wavelength-shifting (WLS) fibres for read-out.

The ECAL is a sampling calorimeter. Like the PS/SPD it uses lead as the absorber and scintillating tiles as the active material, connected to WLS fibres. The ECAL uses a standard PMT for each fibre bunch. The absorber and scintillating plates are arranged in alternating layers with a total thickness of 42 cm ($25X_0$). Using the standard parametrisation of the energy resolution, $\sigma_E/E = a/\sqrt{E} \oplus b$ (with E in GeV/c^2), the energy resolution terms have been determined from test beam measurements to be $8.5\% < a < 9.5\%$ and $b \sim 0.8\%$ [40].

As with the ECAL, the HCAL uses scintillating tiles as the active material, with PMTs to collect the scintillating light. Iron is used as the absorber, with the absorber and active layers placed perpendicular to the beam axis. Owing to space requirements, the HCAL is only $5.6\lambda_I$ in depth, where λ_I is the hadron interaction length in steel, with a further $1.2\lambda_I$ coming from the ECAL. This is not enough to fully contain the shower at all energies, leading to some inefficiency in high E_{T} signal decays, though this does not affect the rejection of low-E_{T} minimum-bias events. In HCAL prototype test beam measurements [41], the energy resolution was determined to be $\sigma_E/E = (67 \pm 5)\,\%/\sqrt{E} \oplus (9 \pm 2)\,\%$.

2.5 RICH sub-detectors

LHCb has two RICH detectors which perform the vital role of distinguishing between charged hadron types. This combination of detectors provides discrimination between kaons, pions and protons over a momentum range of $2 - 100\,\mathrm{GeV}/c$. Cherenkov radiation is emitted when a charged particle travels through a medium faster than the phase velocity of light in the medium. The radiated photons are emitted in a cone of half-angle θ_C, referred to as the *Cherenkov angle*, according to

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

where n is the refractive index of the medium, and $\beta = v/c$ is the speed of the particle expressed as a fraction of the speed of light in vacuum.

Using the formula for the relativistic energy $E = \gamma mc^2$ and relativistic momentum $p = \gamma mc\beta$ of the particle, where m is the particle's rest mass and $\gamma = (1-\beta^2)^{-\frac{1}{2}}$, Eq. (2.1) can be written in terms of the rest mass and momentum of the charged particle:

$$\cos\theta_C = \frac{E}{npc} = \frac{1}{n}\sqrt{1 + \left(\frac{mc}{p}\right)^2} \;. \tag{2.2}$$

It can be seen that a measurement of the particle's momentum and the Cherenkov angle of the emitted light leads directly to a measurement of the particle's rest mass, from which the particle type can be inferred.

In order to minimise the degradation of the tracking systems from material interactions, the RICH photodetectors are located outside of the LHCb acceptance. As can be seen in Fig. 2.8, the Cherenkov light is focused using spherical mirrors, then reflected onto the photodetectors using planar (flat) mirrors. As shown in Fig. 2.2, RICH1 is located directly after the VELO, before the TT and the dipole magnet. Furthermore, for upstream tracks, RICH1 provides the only means of particle identification. RICH2 is located directly after the T-stations.

2.5.1 RICH1

As indicated in Fig. 2.8a, RICH1 contains two radiators. The first is a 5 cm thick layer of silica aerogel, located at the front of the detector. Aerogel has a very low density but a large refractive index (1.01-1.10). This makes it ideal for low-momentum particle identification

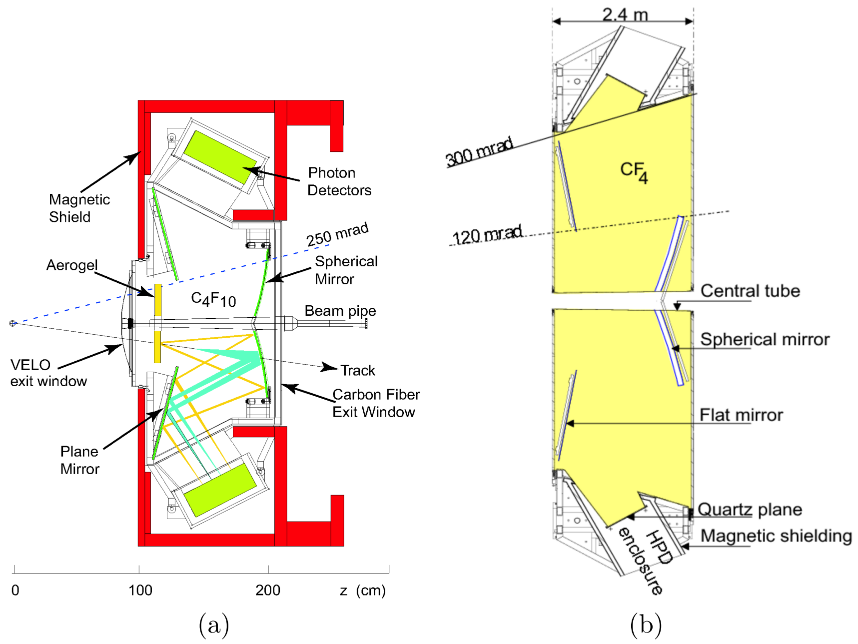


Figure 2.8: (a) side view of RICH1 and (b) top view of RICH2. In both cases the radiators, flat and spherical mirrors, beam pipe (labelled *central tube* in (b)), HPD enclosures (labelled *photon detectors* in (a)), and magnetic shielding are indicated. In addition, the length of the detectors along the beam axis is shown.

(around 2 GeV/c to a few GeV/c). The second radiator is C_4F_{10} gas, which has a refractive index of 1.0014 (at photon wavelength 400 nm) [42]. C_4F_{10} allows discrimination between charged hadrons in the momentum range $10 - 60$ GeV/c. The RICH1 detector covers the full angular acceptance of ± 250 mrad in the vertical (non-bending) plane, and ± 300 mrad in the horizontal (bending) plane. The low angular acceptance is limited to ± 25 mrad by the beam pipe, which passes through the centre of the detector.

As previously mentioned, minimising the material budget in the LHCb acceptance is a key requirement for the RICH detectors. Since the spherical mirrors need to be within the acceptance, they are constructed using a light-weight carbon-fibre reinforced substrate. This, and the supported structure, corresponds to a material budget around $1.5\% X_0$.

2.5.2 RICH2

Figure 2.8b shows that RICH2 contains a single gas radiator, CF_4, which covers a wide momentum range of $16 - 100$ GeV/c. RICH2 has a limited upper angular acceptance of

$\pm120(100)$ mrad in the bending(non-bending) plane. As with RICH1, its lower angular acceptance of ±15 mrad is limited by the beam pipe. The flat and spherical mirrors are manufactured from 6 mm thick glass. Including the gas radiator, the total material budget of RICH2 corresponds to a radiation length of $0.15X_0$.

2.5.3 Detection thresholds

It is clear from Eq. (2.1) that there is a threshold velocity, $\beta_{\min} = 1/n$ below which there will be no Cherenkov light. The threshold momentum, $p_{\min}$, can be written as

$$p_{\min} = \frac{mc\beta_{\min}}{\sqrt{1-\beta_{\min}^2}} = \frac{mc}{\sqrt{n^2-1}} \,. \tag{2.3}$$

The threshold momentum therefore depends on the particle type, with heavier particles having a larger $p_{\min}$. Conversely, as $\beta \to 1$, the Cherenkov angle converges to a maximum value $\theta_C^{\max} = 1/n$. Figure 2.9 shows how Cherenkov angle varies with particle momentum for the different charged particle types and the three radiator media.

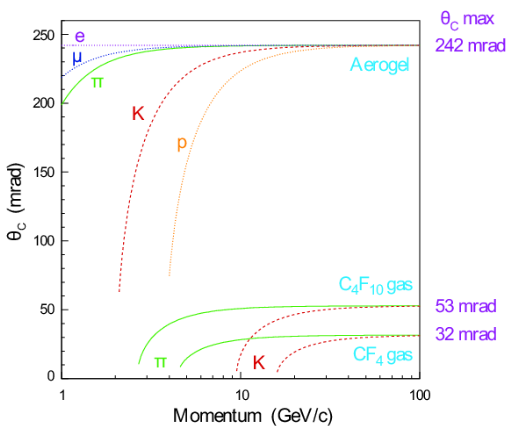


Figure 2.9: Cherenkov angle versus particle momentum for the three RICH radiators. Curves for the five charged particle types ($e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, $p/\overline{p}$) are shown. The maximum Cherenkov angles for the three radiators are indicated on the right-hand side of the plot.

2.5.4 Photon detectors

The RICH detectors use pixel *Hybrid Photon Detectors* (HPDs) as photodetectors for the incoming photons. In an HPD, the photons are converted into photoelectrons by a photocathode deposited on the inner surface of the quartz entry window. The photoelectrons are accelerated by a high voltage, generally between 16 and 20 kV, then focused (with a nominal factor 5 demagnification) onto a reverse-biased silicon detector. A schematic of an HPD is shown in Fig. 2.10.

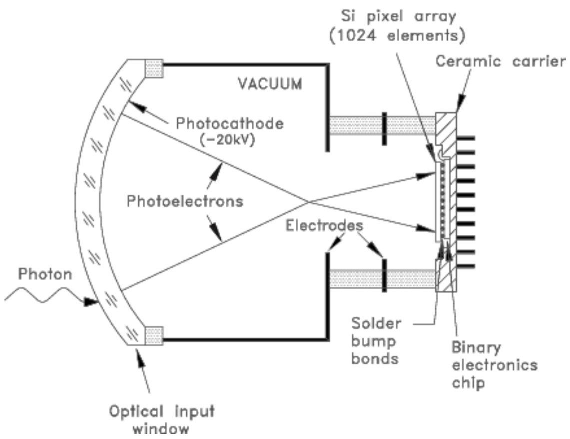


Figure 2.10: Schematic of an HPD. The trajectories of photoelectrons within the HPD are indicated.

The silicon pixel array of an HPD comprises 1,024 $500 \times 500\,\mu\mathrm{m}^2$ pixels arranged as a matrix of 32 rows and 32 columns. Each pixel row is further divided into eight sub-pixels. These sub-pixels can either be read out individually, referred to as *ALICE* mode[1], or combined into a single "super-pixel", referred to as *LHCb* or *physics* mode. Data-taking is always performed with the HPDs read out in LHCb mode, leading to a granularity at the HPD entrance window of 2.5×2.5 mm^2. ALICE mode is used for certain calibration operations which need higher granularity (at the expense of a significantly larger event size), such as the time-alignment of the HPDs, detailed in Chap. 3. The silicon pixel detector is bump bonded onto a binary readout chip. The flip chip assembly is then wire bonded onto a ceramic carrier, which forms the anode of the HPD. A total of 484 HPDs (196 in RICH1 and 288 in RICH2) are used in the detector boxes of the two RICH detectors.

[1] This name is a consequence of the fact that the ALICE experiment uses similar silicon pixel detectors in its inner tracker, which are read out in this way [43].

2.5.5 Magnetic shielding

The fringe field of the LHCb dipole magnet surrounding the RICH1 HPDs is around 60 mT. This field is too strong for the HPDs to operate effectively, so the field must be reduced by around a factor 20. The combination of the Ni-Fe shielding that surrounds each HPD and the thick iron magnetic shielding boxes shown in Fig. 2.8a results in a residual magnetic field of 3 mT around the HPDs, whilst maintaining a field integral of approximately 0.12 T$\cdot$ m between the IP and the TT.

There is also a strong fringe field of more than 15 mT around the RICH2 detector. This is reduced to below 1 mT by the iron HPD enclosures on each side of the detector, shown in Fig. 2.8a.

2.5.6 Photoelectron resolution

The photodetectors image the Cherenkov photons from a track as a ring on the HPD plane. The radius of the ring indicates the Cherenkov angle, which when combined with the momentum measurement from the tracking system can be used to infer the particle type. The distributions of Cherenkov angles, reconstructed under a particular mass hypothesis, are Gaussian distributed. The Gaussian width, σ_θ, corresponds to the angular resolution, which depends on the radiator. For effective discrimination between two particle hypotheses (*e.g.* a kaon and pion), the difference between the peak Cherenkov angles, $\Delta\theta_C$, must be large compared to the resolution (typically $\Delta\theta_C \geq 3\sigma_\theta$). The Cherenkov angle resolution comprises four main sources of uncertainty:

- **emission point error** ($\sigma_\theta^{\mathbf{emission}}$): As the spherical mirrors are tilted in order to focus the photons out of the detector acceptance, the point at which each photon is focused onto the HPD plane depends on the unknown point in the radiator at which it was emitted. The emission point error is a consequence of the fact that the reconstruction software has to assume that all photons are emitted from the mid-point of the trajectory of the track through the radiator.

- **chromatic dispersion** ($\sigma_\theta^{\textbf{chromatic}}$): This is a consequence of the fact that the refractive index of the radiator depends on the unknown wavelength of the emitted photon. The same wavelength is assumed for all photons in the reconstruction, which leads to an irreducible uncertainty.

- **tracking resolution** ($\sigma_\theta^{\textbf{tracking}}$): To calculate θ_C from the photon hit position, the track trajectory (which defines the centre of the ring) needs to be known. The uncertainty from the track reconstruction therefore results in an uncertainty on θ_C.

- **HPD spatial resolution** ($\sigma_\theta^{\textbf{HPD}}$): The blurring of the image of a point source by a focused optical system is known as the *point spread function* (PSF). The spatial resolution of the HPDs is a convolution of the PSF and the pixel granularity at the HPD entrance window, and is dominated by the latter.

Table 2.1 shows the magnitude of the above resolutions for the three RICH radiators according to MC. In all cases, the irreducible chromatic dispersion is by far the largest individual contribution to the total resolution. The Cherenkov resolution for a fully reconstructed ring also depends on the number of photoelectrons, $N_{\text{p.e}}$. The tracking resolution is obviously independent of the number of photoelectrons, but the remaining uncertainties approximately scale as $1/\sqrt{N_{\text{p.e}}}$.

	Aerogel	C_4F_{10}	CF_4
$\sigma_\theta^{\text{emission}}$	0.4	0.8	0.2
$\sigma_\theta^{\text{chromatic}}$	2.1	0.9	0.5
$\sigma_\theta^{\text{tracking}}$	0.4	0.4	0.4
$\sigma_\theta^{\text{HPD}}$	0.5	0.6	0.2
$\sigma_\theta^{\text{total}}$	2.6	1.5	0.7
$N_{\text{p.e}}$	6.7	30.3	21.9

Table 2.1: Cherenkov angle resolutions (in mrad) for an individual photon emitted from the three RICH radiators according to MC. Individual contributions are shown, as well as the total. The mean number of photoelectrons according to the MC are also indicated.

2.6 Muon chambers

Muons are an integral part of the LHCb physics programme since there are many *CP*-sensitive B decays which involve muonic decays, as well as several rare decays such as $B_s^0 \to \mu^+\mu^-$ [44] which can provide constraints or evidence for NP. The muon system also provides information for the high-p_{T} hardware muon trigger and muon identification for the HLT and physics analysis.

The muon system comprises five stations, labelled M1-M5 in Fig. 2.2. The lower and upper acceptances of the muon stations are 20 (16) mrad and 306 (258) mrad in the bending (non-bending) plane respectively, which roughly corresponds to a 20% acceptance of muons from semileptonic inclusive b decays.

The muon chambers M2-M5 are placed downstream of the calorimeters and are interleaved with 80 cm thick iron absorbers in order to remove any stray hadrons which fail to interact in the HCAL. The fifth station, M1, is placed in front of the PS. Its main job is to improve the muon p_{T} measurement for the L0 trigger.

Stations M1-M3 have a high spatial resolution in the bending plane and are used to quantify the track direction and calculate the muon p_{T} with a resolution of around 20% for the hardware trigger. The remaining stations M4 and M5 have a much poorer spatial resolution and are designed to confirm the existence of penetrating muons. The muon hardware trigger relies on a stand-alone track reconstruction algorithm to ensure a fast response and requires hits in all five muon stations. The minimum momentum required for a muon to traverse all five muon chambers is around 6 GeV/c.

Multi-wire proportional chambers (MWPCs) are used in all muon stations. Due to the higher radiation tolerance, the innermost region of the M1 station comprises a series of three *gas electron multiplier* (GEM) foils sandwiched between anode and cathode planes. The MWPCs have an efficiency of greater than 95% at a gas gain of $\approx 10^5$, whilst the inner region of M1 has an efficiency of greater than 96% for a gain of 6×10^3.

2.7 Particle Identification (PID)

The LHCb experiment uses a combination of calorimeters (Sec. 2.4), two RICH detectors (Sec. 2.5), and several muon chambers (Sec. 2.6) to distinguish between charged particles ($e^{\pm}$, $\mu^{\pm}$, $p/\overline{p}$, $\pi^{\pm}$ and $K^{\pm}$), whilst photons and neutral pions are identified using only the calorimetry system.

2.7.1 Calorimeter PID

Electrons and charged hadrons can be distinguished from the π^0 and photonic backgrounds by the lack of hits in the SPD in the latter case. Hadrons generally interact in the HCAL, though a non-negligible amount of charged pions are expected to deposit energy in the ECAL. These pions are distinguished from electrons by the lack of hits in the PS in the former case.

2.7.2 Muon PID

Muons are identified by extrapolating well constructed tracks with momentum above $3\,\mathrm{GeV}/c$ (tracks below this will not reach M2). Hits are searched for in the muon chambers around the extrapolation point of the track, with a *field of interest* (FOI) dependent on the momentum and acceptance angle of the track. Depending on its momentum, a track is considered likely to be a muon if it has hits in the FOI in 2-4 of the outer muon stations.

2.7.3 RICH PID

The RICH detectors are primarily designed to distinguish between different hadron types, but the detectors also have some discriminating powers for leptons. The task of pattern recognition involves many factors, including the distortion of the rings (they are slightly elliptical) due to the tilting of the spherical mirrors, the fact that some rings are disjoint (hits in multiple HPDs, which may be in both sides of the detector), partial rings caused by incomplete coverage, background caused by detector and electronic noise, radiation from secondary particles and so on. Many of the associated complications can be minimised by exploiting the tracking information. It is also possible to find rings without track information [45], and such *trackless ring finding* algorithms are of particular importance in LHCb for monitoring the data quality of the RICH.

The RICH pattern recognition/PID algorithms in LHCb can be divided into *local* and *global* algorithms.

Local pattern recognition

In the local approach, outlined in [46], each track is considered independently of the rest of the event. For each HPD pixel hit, the reconstructed emission angle, θ_i, is calculated with respect to the direction of the associated track. For each track, a log-likelihood, $\log \mathcal{L}(\theta_h)$, is then calculated assuming a mass hypothesis m_h, corresponding to an expected Cherenkov angle θ_h, where h represents the particle type. Assuming an angular resolution σ_θ, the log-likelihood for particle type h is then given by

$$\log \mathcal{L}(\theta_h) = N \sum_i \log \left(1 + \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp \left(-\frac{(\theta_i - \theta_h)^2}{2\sigma_\theta^2} \right) \right) . \tag{2.4}$$

The normalisation parameter, N, is chosen such that terms in which θ_h equals θ_i are unity. The log-likelihood then corresponds to the number of hits contributing to a certain θ_i. The sum is performed over all hits for a given track, though hits which have $\theta_i \gg \theta_{\max}$, where $\theta_{\max}$ is the saturation Cherenkov angle (see Fig. 2.9) are rejected prior to calculation to reduce CPU time. Equation (2.4) is used to determine the number of hits, n_h, which can be assigned to a particular mass hypothesis. A Poisson probability, $P(n_h, n_{\text{exp}})$ can then be constructed given the number of expected photoelectrons. n_{exp}, which can be used to determine the most likely particle type.

Global pattern recognition

The local method of RICH pattern recognition requires relatively modest computation, but does not account for background in the rest of the event, which can be a major hindrance for events with high track multiplicity. A more advanced, albeit much slower algorithm, considers all tracks in the event (from all radiators in both RICH detectors) simultaneously.

In this global approach, the log-likelihood is determined by comparing the number of photoelectrons detected in each pixel with the number expected from both signal and background contributions [47]. If $\mathbf{h} = h_1, h_2, \ldots, h_{n_{\text{tracks}}}$ is a vector of mass hypotheses spanning all n_{tracks} tracks in the event, then the log-likelihood of the event for these mass hypotheses is given by

$$\log \mathcal{L}(\mathbf{h}) = \sum_{i=1}^{n_{\text{pixels}}} \left(- \sum_{j=1}^{n_{\text{tracks}}} a_{ij}(h_j) + n_i \ln \left(\sum_{j=1}^{n_{\text{tracks}}} a_{ij}(h_j) + b_i \right) \right) , \tag{2.5}$$

where n_{pixels} is the total number of pixels with hits in the event, $a_{ij}(h_j)$ is the expected number of photoelectrons from track j in pixel i assuming mass hypothesis h_j, n_i is the total number of hits detected in pixel i, and b_i is the expected number of background hits (hits without an associated track).

The background quantities b_i are estimated per HPD, and Eq. (2.5) is iterated on for different vectors $\mathbf{h}$ until the log-likelihood converges on a maximum. The background estimate is then adjusted, and the procedure repeated until the log likelihood converges. This procedure is obviously CPU intensive, though it can be sped up by first seeding the particle hypotheses using the local approach.

2.7.4 Global PID

Since the individual detector components all provide information on the particle hypothesis, it makes sense to consider a *global log-likelihood* by taking the product of the sum of the individual log-likelihoods. The global log-likelihoods for an electron ($e^\pm$), muon ($\mu^\pm$) and charged pion ($\pi^\pm$) are given by

$$\log \mathcal{L}(e^\pm) = \log \mathcal{L}_{\text{RICH}}(e^\pm) + \log \mathcal{L}_{\text{CALO}}(e^\pm) + \log \mathcal{L}_{\text{MUON}}(!\,\mu^\pm) , \tag{2.6a}$$

$$\log \mathcal{L}(\mu^\pm) = \log \mathcal{L}_{\text{RICH}}(\mu^\pm) + \log \mathcal{L}_{\text{CALO}}(!\,e^\pm \;\&\; !\,h^\pm \;\&\; !\,\gamma) + \log \mathcal{L}_{\text{MUON}}(\mu^\pm) , \tag{2.6b}$$

$$\log \mathcal{L}(\pi^\pm) = \log \mathcal{L}_{\text{RICH}}(\pi^\pm) + \log \mathcal{L}_{\text{CALO}}(h^\pm) + \log \mathcal{L}_{\text{MUON}}(!\,\mu^\pm) , \tag{2.6c}$$

where $h^{\pm}$ refers to a charged hadron and γ refers to a photon or π^0. The ! symbol implies that the likelihood is not consistent with the specified hypothesis. The global log-likelihood for charged kaons and protons/anti-protons are equivalent to Eq. (2.6c) with $\pi^{\pm}$ replaced with $K^{\pm}$ or $p/\overline{p}$ respectively. These quantities are not useful by themselves because there are constant terms in the likelihood which will differ on an event-by-event basis. Hence, the difference between the log-likelihoods from two mass hypotheses, referred to as the *DLL* is considered. For example, the DLL between the kaon and pion mass hypotheses, $\log(\mathcal{L}_K/\mathcal{L}_\pi)$, is defined by

$$\log(\mathcal{L}_K/\mathcal{L}_\pi) \equiv \log\mathcal{L}(K^{\pm}) - \log\mathcal{L}(\pi^{\pm}) \ . \tag{2.7}$$

Thus, if Eq. (2.7) evaluates to greater than zero, then the particle is more likely to be a kaon than a pion. In this way, the DLL between two particle hypotheses is used as a means of particle identification (or more accurately of rejecting one mass hypothesis in favour of another).

2.8 Trigger system

The LHCb trigger system needs to reduce the (visible) event rate of $\approx 10\,\text{MHz}$ to around $2-3\,\text{kHz}$ for writing to storage tapes. To achieve this, a combination of hardware and software trigger systems are employed. The trigger system is illustrated in Fig. 2.11.

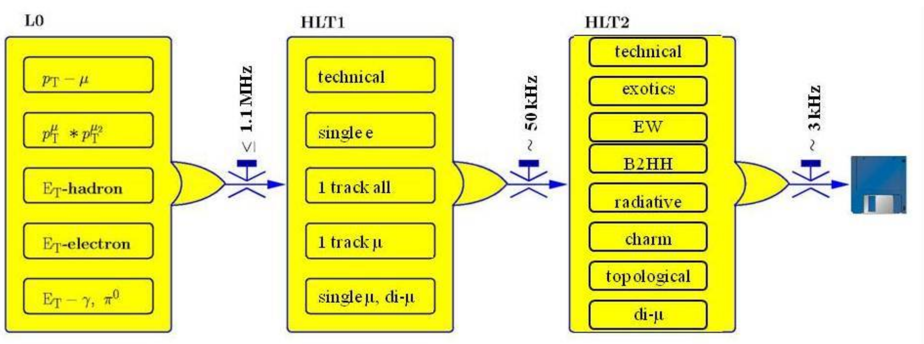


Figure 2.11: Schematic of the LHCb trigger system. The various stages and associated triggers are indicated, as well as the approximate event rate at each stage. At the HLT2 stage, each label indicates a group of triggers, rather than an individual trigger.

2.8.1 L0 trigger

The *Level-0* (L0) hardware trigger system is designed to reduce the 10 MHz event rate to around 1 MHz. Since the B and D mesons have a large invariant mass compared to their decay products (*daughters*) they tend to have high transverse energy (E_{T}) and p_{T} with respect to the minimum bias background, and this is exploited in the L0 trigger.

The calorimeter triggers (hadron, electron and photon) exclusively use information from the calorimetry system, and select events with at least one hadron, electron or photon with a minimum E_{T} threshold. This threshold is varied depending on the run conditions to avoid exceeding the 1 MHz total output rate of the L0 trigger. In addition, events otherwise passing these triggers are rejected if the SPD multiplicity or total E_{T} of the event are above a certain threshold, in order to avoid further processing of very high multiplicity events. Such events which can inhibit the trigger system due to extensive CPU processing at the software level.

The muon trigger utilises information from the muon system to select either singly or pair-produced muons above a certain p_{T} threshold. As with the calorimeter triggers, this threshold is configurable, and is varied based on the run conditions.

Finally, the pile-up trigger uses the pile-up veto stations to reject events with multiple pp interactions per bunch crossing.

The *level-0 decision unit* (L0DU) collates the aforementioned triggers and forms a global trigger decision, which is sent at 40 MHz to the Readout Supervisor (see Sec. 2.9) where it is relayed to the front-end electronics. The time between the pp interaction and the trigger decision reaching the front-end electronics, known as the *L0 latency*, is fixed to 4 μs.

2.8.2 High Level Trigger

The High Level Trigger (HLT) is a two-stage software trigger. The HLT algorithms, called *trigger lines*, run on an *event filter farm* (EFF) comprising several thousand CPU nodes, each having around 20 ms on average to run the track reconstruction and the two HLT trigger stages. The first stage, HLT1, comprises a series of trigger lines similar to the L0 triggers, but using track information from the VELO and T stations. The second stage, HLT2, exploits full event reconstruction to apply more sophisticated selection criteria, which are used to further reduce the event rate.

2.8.3 HLT1

The HLT1 trigger lines involve a partial reconstruction of the event, and are aimed at "confirming" the L0 decision. The calorimeter hits or muon track segments are matched with track segments in the VELO and T stations, but the full event is not reconstructed, since the size of the computing farm prevents this at 1 MHz. The trigger must reduce the event rate by around a factor 20 before the full reconstruction can be performed and has a CPU budget of around 10 ms to achieve this.

The *1track all* trigger line, hereby referred to simply as the 1track trigger line, is the principal HLT1 trigger line for B and D mesons decaying to fully-hadronic final-states. The 1track trigger lines rely on the fact that in B decays, there tends to be at least one good quality, high momentum track. Selection criteria are applied to the VELO track seeds to reduce the number of tracks which need to be extrapolated to the T-stations. The VELO seeds are then extrapolated to the T-stations using the forward-track reconstruction algorithm, which is configured to only look for tracks with momentum and p_{T} above a certain threshold. Finally, the Kalman fitter is run, and selection requirements are applied to the minimum $IP\,\chi^2$ and maximum χ^2 of the tracks, as well as more stringent requirements on momentum and p_{T} in order to to achieve the required event rate. The 1track trigger is described in more detail in [48].

2.8.4 HLT2

The reduced event rate from HLT1 allows the HLT2 stage to exploit the full event reconstruction. This allows the trigger lines to be more in union with the *stripping* (see Sec. 2.10.3) and *offline selections* (the selection criteria applied by analysts to stripped events to further reduce background). Some of these triggers are inclusive, encompassing a variety of

B or D hadron decays, whilst others are exclusive, aiming to provide the maximum possible efficiency for specific B or D decays, or optimised for a specific physics analysis. The main trigger lines for fully-hadronic B decays are known collectively as the *topological* trigger lines. For most multi-body fully-hadronic B decays (with three or more final state particles), these represent the only HLT2 lines which are efficiently triggered on.

A flow diagram of the topological trigger lines is shown in Fig. 2.12. The simplest trigger line is the two-body line. First, two of the Kalman-fitted tracks are combined to form a displaced vertex. Since the difference between the mass of the tracks under different track hypotheses is small compared to the masses of the B hadrons, all tracks are treated as kaons to reduce the number of required track combinations. If two constituent tracks have a *distance of closest approach* (*DoCA*) below a certain threshold, then they are treated as input to the two-body topological line. Additional selection requirements are then applied to the two-track vertex to determine whether or not it passes the two-body trigger line. Since the topological trigger aims to trigger inclusively on any B decay with at least two charged hadrons, the vertex formed by the two-body topological trigger line (and those formed by the other topological lines described below) does not necessarily correspond to the true, fully-reconstructed B decay vertex, since the trigger may not reconstruct all tracks in the decay, nor does the topological trigger reconstruct any further displaced vertices in the decay, such as a D decay vertex. Consequently, the cuts applied in the trigger line do not necessarily apply to all B daughters, allowing tight cuts to be applied to the tracks considered, and certain selection requirements, particularly those related to the quality of any secondary decay vertices, are avoided, since they cannot be properly calculated in the case of missing B daughters. Further details of the HLT2 topological trigger, including a summary of which cut variables are considered by the trigger, can be found in [49].

The track combinations that fail the *DoCA* requirement of the two-body topological trigger line are used as seeds for the three-body trigger line. Another *DoCA* requirement is applied between the two-track vertex created in the former stage and an additional track. If the track combination passes the *DoCA* requirement, the two-track vertex is combined with the third track to form another vertex, which is then treated as the vertex of a three-body decay (by having the three-track candidate instantaneously decay to the two-body vertex and the additional track). This seeding technique reduces the number of three-track combinations that need to be made by the topological trigger, significantly reducing processing time. As with the two-body topological trigger line, additional selection requirements are applied to the three-body vertex to determine whether the event passes or fails the trigger line. Three-track combinations that fail the *DoCA* reqirement are used as input to the four-body trigger line. The same procedure is then applied to create four-body topological trigger-line candidates.

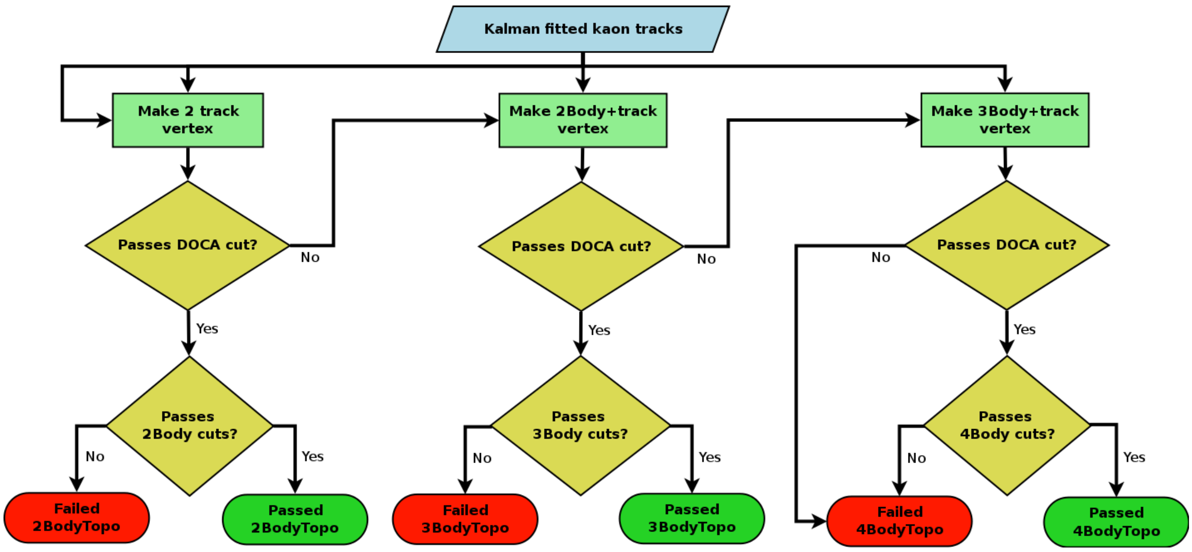


Figure 2.12: Simplified flow diagram of the HLT2 topological trigger lines.

2.9 Online system

The *online* system handles the transfer of data from the front-end electronics to permanent storage. It comprises the *Data Acquisition* (DAQ) system, *Timing and Fast Control* (TFC) system and the *Experiment Control System* (ECS).

In addition to the data transfer, the DAQ system handles the event building, filtering and monitoring of the data, and includes the *readout network*, which handles the event building, the EFF, defined in Sec. 2.8.2, which runs the HLT software, and the *monitoring farm*, which is used to monitor the data quality.

The front-end electronics of the LHCb experiment captures, processes, amplifies and buffers the detector signals, and then filters the events based on the decision of the L0DU, described in Sec. 2.8.1. Whilst the front-end system differs between sub-detectors, the front-ends all conform to a common specification [50]. The front-end and readout architecture is illustrated in Fig. 2.13.

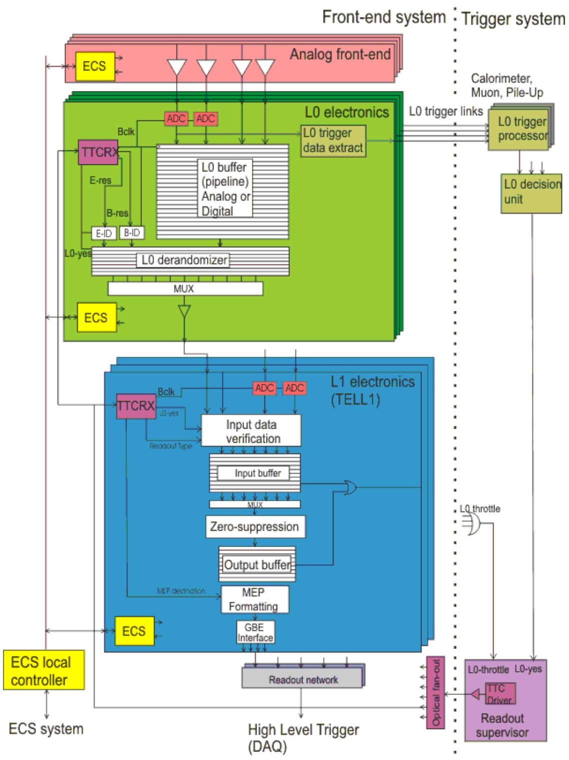


Figure 2.13: Diagram of the architecture of the front-end and readout electronics common to all LHCb sub-detectors.

2.9.1 L0 front-end electronics

Firstly, for all sub-detectors, an analogue front-end sub-system amplifies and shapes the detector signals. To minimise spill-over, the shaping time must be as small as possible (< 25 ns). The shaped signals are then further processed by the *L0 electronics*. For some systems such as the RICH, the L0 electronics are required to be on the detector itself. The L0 electronics contain two buffers. The first is the *L0 pipeline buffer*, which receives the signal from the analogue front-end and stores it temporarily during the L0 latency (as defined in Sec. 2.8.1). For some sub-detectors, analogue-to-digital converters (ADCs) are used to digitise the signal prior to storage in this buffer. If the L0 electronics receive an L0 trigger accept, then the relevant data are transferred to the *L0 derandomizer* buffer. The L0 accept rate is essentially random, so it is possible for multiple events to be accepted in quick succession. The derandomizer buffer alleviates bottlenecks as a result of the variable accept rate by storing 16 L0 accepted events and transferring these data to the L1 electronics using a constant rate multiplexed data link. With the exception of the VELO, the data transfer is performed using optical links based on the radiation-hard Gigabit Optical Link (GOL) serialiser chip [51].

2.9.2 L1 readout electronics

The L1 electronics[1] comprises a series of readout boards which are located in the *barracks*, situated some distance from the detector behind a thick concrete wall that acts as a radiation shield.

All sub-detectors, with the exception of the RICH detectors, use programmable *TELL1* modules for their readout [52]. The RICH detector has opted instead for *UKL1* boards [53], although from the point of view of the data-readout, the two types of module are equivalent. The L1 readout modules receive data from the L0 electronics and perform a basic verification of the data. In the case of the analogue signals from the VELO, the data are digitised prior to verification. Accepted events are then *zero suppressed* (redundant zeros are removed from the digitised data) to reduce the event size. The zero-suppressed data are then stored in another derandomizer buffer before being combined into *multi-event packets* (MEPs) to reduce the protocol overhead during transfer via a Gigabit Ethernet (GBE) interface to the readout network. Finally, the data are transferred from the readout network to nodes on the EFF, to be processed by the HLT. A subset of the data are also sent to the monitoring farm to perform additional data quality checks.

2.9.3 TFC

The TFC system comprises three components: the *optical distribution network*, the *trigger throttle network* and the *Readout Supervisor*.

[1] This name is a historical artifact. There was initially envisioned an additional L1 trigger at this stage of data processing.

Optical distribution network

The optical distribution network is based on the *Timing and Trigger Control* (TTC) system used throughout the LHC [54]. The TTC system distributes the 40 MHz LHC bunch crossing clock (Clock40) to receivers on the front-end and readout modules for each the various sub-detectors. The trigger accept/reject signal is generated on the output pins of the TTC receivers (TTCrx). For each accepted signal, an internal event ID counter referred to as the *Level0 ID* (L0ID) is incremented. The L0ID is reset to zero by an event ID reset command. Similarly, the bunch crossing is identified in the TTC system by a local bunch crossing ID (BXID) counter. The BXID reset command is issued at the beginning of the bunch structure to ensure that the counter is synchronised with the LHC counter. To ensure that the front-end and readout modules are correctly synchronised, the L1 verification process checks that the BXID and L0ID of the incoming L0 triggered event agree with numerical expectation.

The clock phase of the TTC receivers (TTCrx) is determined by delays in the distribution. These delays are controlled by two sets of delay generators with a resolution of $\sim 100\,\mathrm{ps}$, which are used to align the local clock to the arrival of the detector signals, accounting for delays in the system such as time of flight, detector response, delays in the analogue front-end and so on. These delays often vary the sub-detector modules, and these delay generators can be used to synchronise the sub-detector signals. This type of relative alignment within the clock cycle is known as *fine time alignment*.

The two deskewed (delayed) clocks are known as Clock40Des1 and Clock40Des2. The former clock is used by the majority of TTC signals. The latter is generally only used for broadcast commands which need to be offset relative to Clock40Des1, since the two fine-time delay generators can be set independently of one another. The L0 trigger is phase-aligned with the Clock40Des1 reference clock. It is also possible to align the trigger in 25 ns clock cycles using a programmable pipeline delay in the TTCrx, with a maximum delay of 15 clock cycles. This is referred to as *coarse time alignment*. The coarse and fine time alignment of the RICH detectors is discussed in Chap. 3.

Trigger throttle network

The design of the front-end and readout electronics aims to avoid buffer overflow. However, due to the large fluctuations in event size after zero suppression, it is possible for the L1 output buffer to become congested. In this case, the zero suppression is stopped and the output buffer is forced to store the event data. At this point, the trigger throttle network issues a L0 *trigger inhibit* (throttle) preventing the broadcast of any further L0 triggers until the output buffers are emptied.

Readout Supervisor

The final part of the TFC system is the *Readout Supervisor*, also known as the *ODIN*. The ODIN is the master of the TFC system, and handles the interface of the trigger systems and the readout system. It synchronises the trigger decisions and beam synchronous commands to the LHC clock. It is also capable of producing a variety of "auto-triggers", which can be used to perform calibration or test runs, or for monitoring purposes. The ODIN also has the vital roles of collecting the L0 trigger decisions, ensuring that trigger accepts are only broadcast if there is no risk of buffer overflow in either the front-end or readout systems and requesting L0 throttles. Dynamic load balancing of the nodes on the DAQ CPU farm is also handled by the ODIN. For each trigger accept, the ODIN transmits a data bank, which is appended to the event data in the L1 system. This data bank contains the event identifier (event number *etc.*), the (GPS) time of the event, and the trigger type (physics, or the auto-trigger type).

There are in fact multiple Readout Supervisors, which allow the detector to be *partitioned*. This means that all sub-detectors can be controlled by a single ODIN, as is the case for normal data-taking, or each sub-detector can be controlled by its own ODIN, each utilising a subset of the farm nodes. This allows several independent activities, such as the time alignment of the sub-detectors, to take place concurrently.

2.9.4 ECS

The ECS controls and monitors the state of the LHCb detector. The ECS includes the *Detector Control System* (DCS), which supervises the high voltages (HVs) and low voltages (LVs), temperature and pressure controls and the gas flow systems. The ECS is also responsible for controlling and monitoring the trigger, DAQ and TFC systems. To simplify the ECS system, the number of different types of interfaces and buses to the various detector systems has been minimised, so there are only three types of *field buses* in use by LHCb:

- **SPECS**, a *Serial Protocol for ECS*, which is a high speed (10 Mbytes/s) serial bus used to provide access control to the front-end electronics [55];
- **CAN** (Controller Area Network), which is a standard system bus[1] that allows devices within a system to communicate without a host computer;
- **Ethernet**.

The Ethernet components utilised by LHCb are not radiation-hard, so Ethernet is only used in radiation-free zones, such as the barracks or on the surface. SPECS and CAN are also not fully radiation-hard, but they are able to withstand modest levels of radiation, so they are mainly used in the peripheral areas surrounding the detector. Apart from the control of individual PCs, Ethernet is used to communicate with the readout boards through Credit-card sized PCs (CCPCs) mounted on the TELL1 and UKL1 boards.

The ECS software uses the *Joint Controls Project* (JCOP) framework [56] developed for the LHC experiments, which is based on *PVSS II* (usually referred to simply as PVSS), a commercial SCADA (Supervisory Control And Data Acquisition) system. It is a *distributed system*, meaning that PVSS managers are distributed across different machines, each responsible for a different part of the DAQ and DCS systems. This is a requirement for LHCb partitioning (see Sec. 2.9.3).

[1] ISO 11898. See for *e.g.*, `http://www.iso.org`.

2.10 LHCb software framework

The majority of the LHCb software is based on the C++ software framework known as GAUDI [57, 58]. The GAUDI framework provides the necessary interfaces and services for constructing the various LHCb software packages.

The *offline* software packages (as opposed to the software running on the online DAQ systems) can be classified into two groups: *simulation* software, and *reconstruction* software. These software packages are described below.

2.10.1 Simulation

The LHCb software package GAUSS is used to produce simulated *Monte Carlo* (MC) event data. The GAUSS software consists of a *generator phase*, consisting of the generation of the pp collisions and the decay of the particles, and a *simulation phase*, in which the traversal of the particles through the LHCb detector is simulated. The generator phase employs several external packages, primarily PYTHIA [59], used to generate the pp collisions and *EvtGen* [60], which is a library that specialises in the decay of b hadrons. The simulation phase utilises the GEANT4 toolkit [61].

The final stage of the LHCb detector simulation is the *digitisation*, performed by the BOOLE software package. BOOLE takes the sub-detector hits generated by GEANT4 and simulates the detector response, producing digitised data which mimics the raw data output by the online systems. The association of the MC hits to the digitised hits is also incorporated by BOOLE.

2.10.2 Reconstruction and data analysis

The offline reconstruction of both the simulated data and the “real” pp collision data output from the DAQ system is performed by components of the BRUNEL software package.

The reconstruction software decodes the raw data from the various sub-detectors and performs the full event reconstruction discussed in previous sections. For MC data, the reconstruction software also determines the association between the reconstructed tracks and the MC particles produced by the simulation software. The reconstruction creates *ProtoParticle* objects, which contain the full reconstruction information for a charged or neutral particle.

The final stage of reconstruction is performed by the physics analysis software package DaVinci. DaVinci is the only part of the LHCb software which is used by the majority of LHCb analysts. The DaVinci algorithms take the ProtoParticles from the output files produced by Brunel, known as *DST* files and reconstructs the full decay, including assigning mass hypotheses to charged particles and fitting vertices. The kinematic and vertex information from the reconstructed decays can then be used to select events which are likely to be of interest to a particular analysis. These events of interest can either be stored as a DST file, or the parameters of interest to the analyst (the particle mass, the momentum of the daughters, the vertex and track quality *etc.*) can be output to *nTuples*, which can be analysed using the *ROOT* data analysis software framework [62].

2.10.3 Stripping

Even with the substantial reduction in the event rate from the L0 and HLT triggers, the size of the reconstructed events means that allowing every analyst to run over all the reconstructed data produced by LHCb would be unfeasible.In addition, many events sent to permanent storage will be of no interest to the analyst, since LHCb is designed to collect data for a wide variety of physics analyses.

To deal with the above issues, LHCb performs an additional filtering of events after reconstruction, a process known as *stripping*. Stripping is effectively an additional trigger, except that the raw data are still available in permanent storage, though the analysts generally are only able to analyse the stripped data. This philosophy permits the possibility of changing the selection criteria used to strip the raw data at a later date, allowing the LHCb analysis groups to make changes to their selection criteria and permitting the stripping of data relevant to new physics analyses.

A particular set of stripping cuts is known as a stripping *round*. Each round of stripping is given a unique name, such as "Stripping12", so the analysts can easily identify and process data for a specific set of evolving conditions. Any new data collected by LHCb is processed using the cuts from the latest stripping round. At the end of the year, a special stripping round is produced, which is used to reprocess the entire year's worth of accumulated data.

The stripping is organised into *streams*, which contain a number of *stripping lines*. Stripping lines are similar to the HLT2 trigger lines; they comprise a number of inclusive or exclusive selection criteria. The content of each stripping stream is generally handled by a specific working group in the LHCb collaboration, which is responsible for ensuring that the contents of the stream keep to a reasonable retention rate and processing time.

Chapter 3

Time alignment of the RICH detectors

This chapter summaries the series of analyses which were performed to time align the RICH detectors shortly before the first *pp* collisions and the early stages of operation. It also describes the hardware and software components of the RICH DAQ system.

3.1 Introduction

As discussed in Sec. 2.9.3, the time alignment of the detectors is a crucial part of the detector operation. For the RICH detectors, the TTCrx delay generators are located on the *L0 interface modules*, described in Sec. 3.2. Each HPD has a different response time. This is mainly due to the *silicon-bias leakage current*, which varies between HPDs (at a given temperature and silicon bias voltage). The leakage current is reduced by increasing the *discriminator threshold* (as defined in Sec. 3.2.1), which in turn increases the response time. Clearly, the ideal situation would be to time align each individual HPD. Unfortunately, this is not possible since two HPDs are attached to each L0 interface module. Thus, to maximise photon efficiency, each L0 board was connected to two HPDs with similar leakage currents during the detector assembly. However, subsequent replacements of HPDs has meant that matching the timing profiles of the two HPDs on a board has not always been possible, and this can prove to be challenging for the time alignment.

This chapter describes how the time alignment of the RICH detectors was performed shortly before, and during the early stages of detector operation, using a both a pulsed laser system, and $\sqrt{s} = 900\,\mathrm{GeV}$ and $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data.

Before the time alignment analysis is discussed, a brief description of the RICH L0 electronics and L0 DAQ ECS systems is given.

3.2 RICH L0 electronics

The RICH on-detector electronics is split into two separate parts:

- the **LHCBPIX1 binary front-end pixel chip**, which is located inside the vacuum tube of the HPD, and performs the amplification and shaping of the HPD signal;
- the **L0 interface module**, which interfaces with the LHCb TFC and ECS systems (and indirectly, the L0 trigger system), the LV power supplies, the L1 electronics and the LHCbPIX1 pixel chip.

3.2.1 LHCBPIX1 pixel chip

Figure 3.1 shows the circuit diagram for an individual pixel of the LHCbPIX1 chip. Each pixel contains its own analogue front-end, which comprises a pre-amplifier, a two-stage signal shaping system and a *discriminator*. The discriminator compares the shaped signal current against a threshold fixed globally across the pixel chip. The binary output of the discriminator corresponds to the hit decision: signals above the threshold are considered hits (logic bit 1). Three latches are used to fine-tune the threshold on a pixel-by-pixel basis to maximise the signal-to-noise ratio. Another latch can be used to inject a test pulse into the analogue front-end. The test pulse is generated by a circuit on the pixel chip, which is triggered by an external pulse and transmitted to all pixels on the chip. A final latch is located after the discriminator, which can be used to *mask* (deactivate) the pixel. This is useful for rejecting signals from malfunctioning or excessively noisy pixels.

The digitised signals are stored in two L0 pipeline buffers during the $4\,\mu\mathrm{s}$ L0 latency. The strobe trigger signal, which is by default one clock cycle long, then drives the L0 accepted

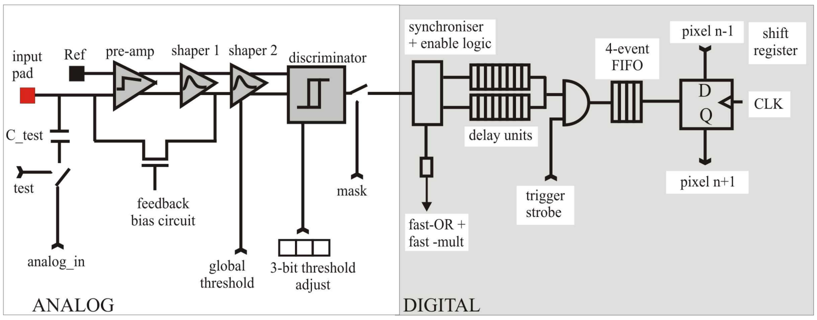


Figure 3.1: Simplified circuit diagram of a pixel on the LHCbPIX1 chip. The diagram has been split into the analogue and digital parts.

signals into a 4-event FIFO (first-in first-out) derandomizer buffer. The FIFO output is then loaded into a flip-flop, with the flip-flops of each pixel column forming a shift register. As discussed in Sec. 2.5.4, each pixel chip comprises 256×32 individual pixels, which can either be read out individually (ALICE mode) or as a series of 32×32 "super-pixels" (LHCb mode). The pixel logic for these two modes is illustrated in Fig. 3.2. In the case of LHCb mode, used during normal data-taking, a fast logical OR operation is performed on the digitised output from eight pixels to provide the single super-pixel signal. The sixteen delay buffers are configured as an array, and four of the FIFO buffers are combined to form a sixteen event FIFO. The FIFO output is then loaded onto the flip-flop of the top pixel, bypassing the other seven flip-flops. Finally, the thirty two 32-bit column registers are multiplexed and transmitted via kapton cables to the L0 interface module.

The delay between the signals entering the pipeline buffers and the strobe trigger signal being sent is controlled by the 8-bit *delay_control* register. If n is the decimal representation of the delay_control register contents, then the total delay will be $2n + 2$ clock cycles. Thus, the trigger delay can be between 2 and 512 clock cycles (50 ns – 12.8 µs). The delay_control register for all LHCbPIX1 chips has been set to 160 clock cycles, which corresponds to the L0 latency. The strobe signal is synchronised to Clock40Des1, so further delays due to time-of-flight, cable lengths, signal propagation and so on are dealt with by the TTCrx coarse and fine delay generators.

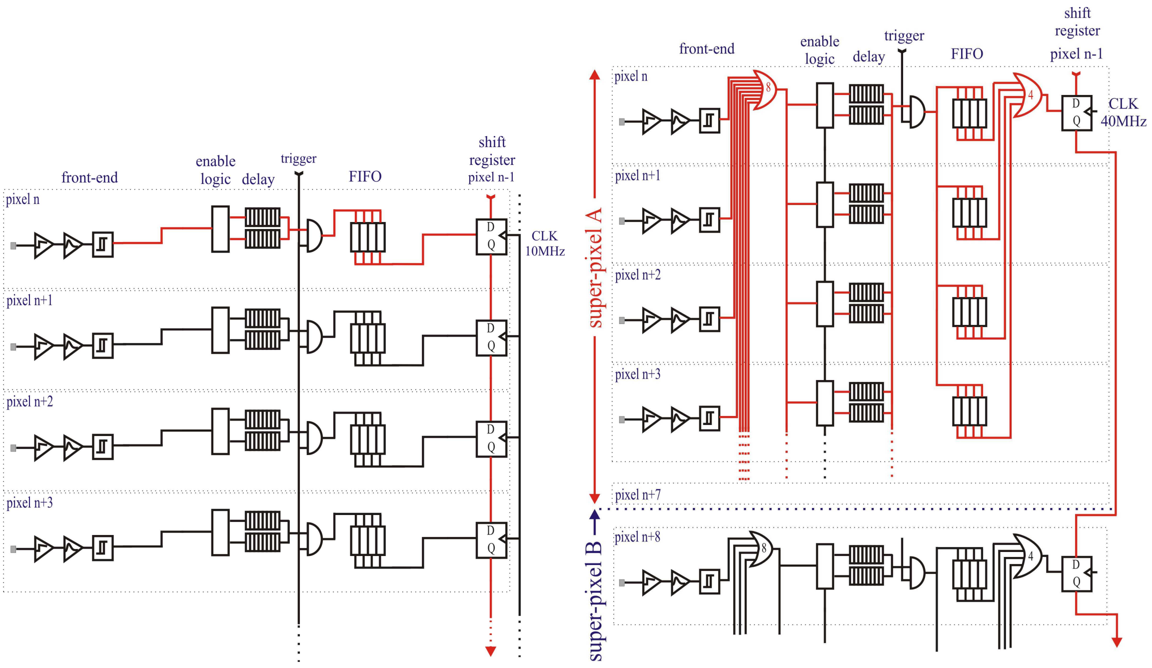


Figure 3.2: Diagrams indicating the LHCbPIX1 pixel logic in (a) ALICE and (b) LHCb readout modes.

3.2.2 L0 interface module

The L0 interface module is illustrated in Fig. 3.3. The cable plugs for the two LHCbPIX1 chips which are connected to the module can be seen at the base of the board. At the heart of the L0 interface module is the *Pixel Interface* (PINT) chip. The TFC signals are received through the TTCrx by the *Truelight* receiver, and thence interpreted by the PINT, which includes setting the correct mode of operation for the front-end electronics, *e.g.* configuring the pixels for readout in ALICE or LHCb mode. The PINT generates the strobe signal and transmits it to the LHCbPIX1 chips, synchronises and formats the signal data from the LHCbPIX1 chips, and adds additional information to the event data received from the TTCrx such as the BXID and L0ID before transmitted the data to the GOLs. The GOLs then serialise the data and perform a second level of multiplexing, resulting in a an output rate of 1.6 Gbytes/s. The data are transferred to the L1 electronics via optical fibres, with the optical signals generated by *Vertical Cavity Surface Emitting Lasers* (VCSELs). The VCSEL currents and bias voltages are controlled by the GOLs.

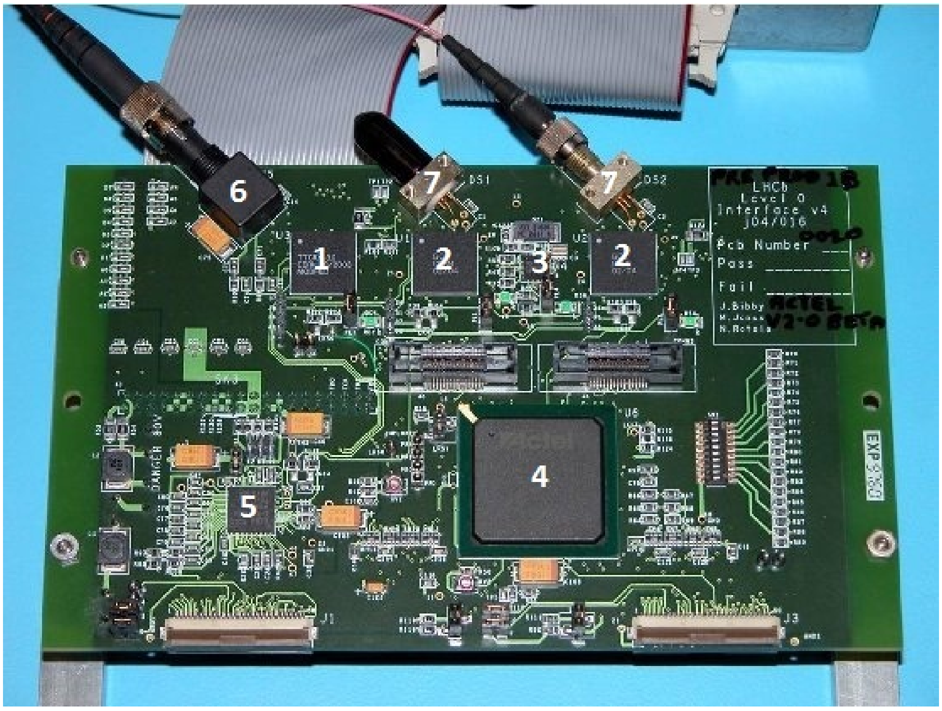


Figure 3.3: Photograph of an L0 interface module, comprising the following components: TTCrx (1), GOLs (2), QPLL (3), PINT (4), PILOT (5), Truelight receiver (6) and VCSELs (7).

The TTCrx Clock40Des1 clock signal has a relatively low jitter of around 300 ps peak-to-peak. However, to ensure high-speed transmission of data through the GOLs, the clock input received from the TTCrx to the GOLs must have a jitter of 40 ps or less. This is achieved by a special jitter filter circuit called a *Quartz Phase Locked Loop* (QPLL) [63].

The final component of the L0 interface module is the analogue PILOT chip, which provides reference voltages for the various LHCbPIX1 Digital Analogue Converters (DACs), which includes the voltage bias used to set the global threshold and the delay control register. Details of typical PILOT reference voltages can be found in [64]. The PILOT also monitors its own temperature and that of the LHCbPIX1 and PINT chips, using four temperature-monitoring sensors.

The PINT uses the JTAG (Joint Test Action Group) protocol[1] to configure and monitor all of the components of the L0 interface module. As with the SPECS and CAN protocols,

[1] The majority of the JTAG interface logic adheres to the IEEE 1149.1 boundary scan architecture standard. See `http://standards.ieee.org/findstds/standard/1149.1-1990.html` and `http://www.jtag.com` for further details.

the JTAG protocol is not radiation hard. To avoid accidental changes of detector state due to SEUs (single-event upsets), the hardware reset is kept in an active state when the boards are configured. This means that the board states cannot be altered via the JTAG protocol after configuration; however, this also prevents the monitoring of the registers during data-taking.

3.2.3 RICH L0 columns

The L0 boards are mounted in columns comprising eight (RICH1) or nine (RICH2) boards. Two L0 boards share a single LV regulator. The SPECS protocol is used to control the LV regulators, and each column has a single *SPECS slave* from which commands are propagated to the other LV regulators on the column.

3.3 RICH L0 DAQ ECS system

Each of the four HPD enclosures: the two RICH1 enclosures (the *upper* and *lower* detector *boxes*) and the RICH2 enclosures (the *left* and *right* detector *sides*) have an associated control PC located in the LHCb control room. Two of these PCs (one for each RICH detector) run the PVSS RICH L0 DAQ ECS *projects*. One software *SPECS server* runs on each control PC, which receives commands from the DCS PVSS projects. The SPECS server also relays the status of the LV boards via a number of *SPECS master* drivers located on the control PC and SPECS slaves located on the L0 columns. In addition, each PC runs multiple DIM server instances, one for each L0 column it controls. The DIM servers monitor the status of the L0 electronics during the power up sequence, control the configuration of the L0 components and HPDs, receive commands from the L0 DAQ ECS and check the status of the boards after data-taking. The DIM servers are also used to perform special calibration runs, such as those used to perform the coarse and fine time alignment of the L0 boards.

3.3.1 ECS software architecture

The ECS software architecture comprises two components. The high-level interaction with the detector software and hardware components is performed by the *Control Units*

(CUs). The system is *hierarchical*, meaning that each CU has a single parent node from which it receives commands, and itself can have a number of children (of which it is the only parent node), which it can issue commands to. The operator can take control of a CU at any stage of the hierarchy, which allow commands to be issued to individual detector components without affecting the rest of the detector. The *Device Units* (DUs) are low-level components which interface directly with the various hardware and software devices. Examples of DUs in the RICH ECS architecture are the LV and HV power supplies, and the individual L0 boards. Each DU is connected to a single CU controller, and the operator can only command the DUs through the CU which controls it.

Commands propagate from the operator or control program connected to the CU down the command tree to the DUs. State changes and alarms propagate up from the DU that changes its state or issues the alarm up to the top-level CU under the operator's control. Each DU can only be in one of several definite states, and can only change state to/from particular states. Thus, the ECS and its sub-component CUs are referred to as *Finite State Machines* (FSMs). LHCb uses the *SMI++* framework [65] for the software implementation of its FSMs, which is based on the state manager used by the DELPHI experiment. The physical devices communicate with PVSS via the *Distributed Information Management* (DIM) system [66], also designed for DELPHI.

3.3.2 PVSS datapoints, recipes and panels

The device data in PVSS is structured in terms of *datapoints* (DPs), each with a specific *datapoint type* (DPT), such as string (text), floating-point number, integer and so on. Each device can be associated with a number of DPs. These DPs generally either represent detector readings, such as temperature or voltage monitors, or settings that controls how the device operates, such as the TTCrx coarse and fine time delays. In the former case, the DP is updated by the device communicating that its value has changed, whilst in the latter case, if the DP value is changed then the new setting is propagated to the device.

A key part of the ECS is the *configuration databases*. Each PVSS manager has access to one or more configuration databases. The configuration databases store values for a number of DP settings that can be used to ensure that each device is correctly configured, such as the voltage settings of the LV boards and the values to use for the TTCrx delay generators and pipeline delays. The operator can retrieve information from the configuration database and set the relevant PVSS DPs by issuing a *CONFIGURE* command from the relevant CU. This configuration step is an essential part of the power on process for the RICH L0 front-end boards, since they don't have the ability to store such settings in non-volatile memory, and the default settings are often far from optimal. The configuration settings are stored in *recipes*. The CONFIGURE command takes an argument which tells PVSS which recipe(s) to apply. This allows the ECS to be able to automatically configure the detector (or sub-detector) for a given configuration. For example, there are recipes to configure the HPDs to run in ALICE or LHCb mode and to perform the coarse and fine time alignment scans. These recipes give the ECS a great deal of control as to how the experiment is configured, and makes it trivial to change configuration settings when needed, *e.g.* if the detectors need to be time aligned, or if a detector component is replaced.

Since there are a vast number of DPs to set for the L0 boards, the recipes are cached on the control PC, so that subsequent CONFIGURE commands will cause the recipes to be loaded from the cache rather than from the database. When the recipes are altered, the old cache files are removed by issuing an *UNLOAD* command.

The manipulation of the CUs and DPs is achieved using a number of *graphical user interface* (GUI) windows called *panels*. Each CU has its own ECS panel, with icons representing the CU and its daughters. The CU state icon can be used to change the state of the CU, *e.g.* to switch on or configure the detector or sub-detector (and its daughters, granddaughters and so on) and also to take and release control of the CU. The daughter sub-system icons can be used to navigate down to the next level in the ECS hierarchy. These icons also show which state the CU is currently in. Each CU panel can have its own sub-panel, which can be used to display and manipulate the status of the detector, sub-detector or sub-detector components by querying or updating the PVSS DPs. These panels can contain a variety of GUI objects, including plots, histograms, text fields, buttons and drop-down menus, as indicated by Fig. 3.4. Such panels make it easy for operators to check the detector status and to react to or fix any issues.

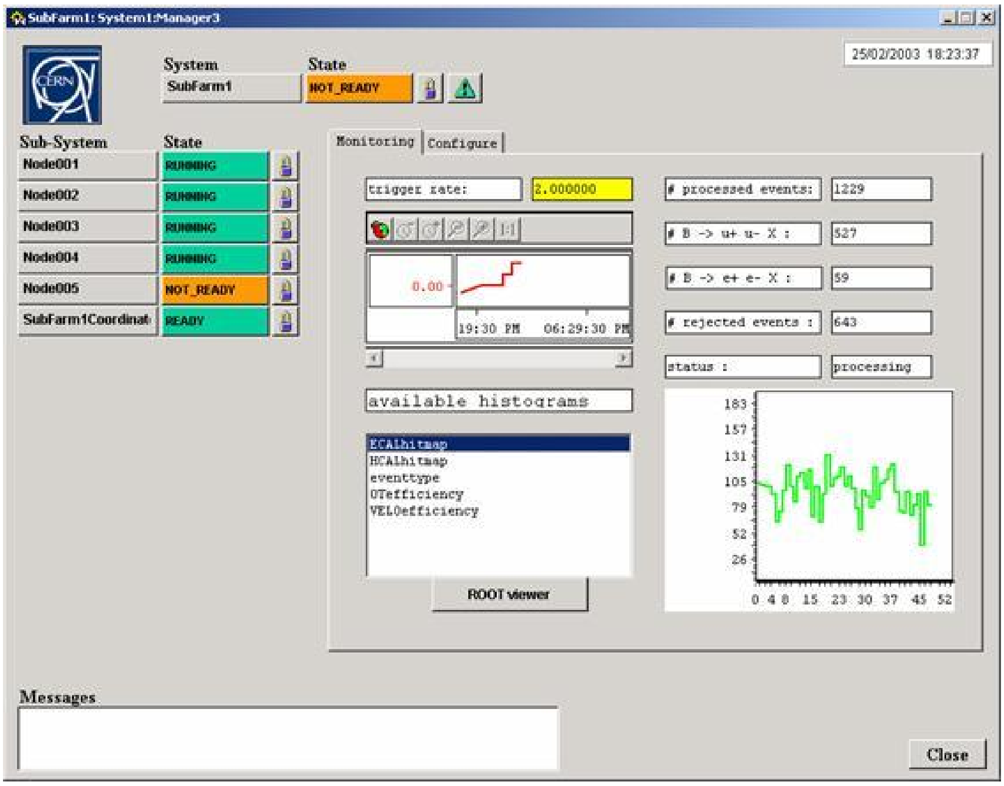


Figure 3.4: Example of a PVSS ECS panel, illustrating many of the possible graphical features. The same basic panel is used for all ECS CUs (referred to as the *System* on the panel), and contains the name of the system and its state (top left), and its direct descendants (referred to as *Sub-Systems*) and their state (left). This type of panel is called a *reference panel*, since it uses reference variables to determine what the correct system name should be, and the number and names of the sub-systems. Commands can be sent to the system or one of its sub-systems by clicking on the relevant *State* box. The padlocks are used to take or release control of the system or sub-system. Each CU panel can have a sub-panel (right), which is used to display and manipulate parameters relevant to the system and its sub-systems. This sub-panel may also contain references, allowing the same sub-panel to be used for multiple CUs.

3.3.3 RICH L0 DAQ recipe configuration

As explained in Sec. 3.3.2, a string is passed to the PVSS managers during a CONFIGURE command informing the manager of which recipes need to be loaded. This string, referred to as the *Activity*, is set via the ODIN TFC control panel. In the case of the RICH L0 DAQ systems, the main Activities are:

- **DEFAULT**, which is the default Activity that loads the recipes needed to configure the L0 boards for normal data-taking;

- **ALICE**, which is used to configure the L0 boards for readout in ALICE mode.

The RICH L0 DAQ managers load a number of recipes during configuration, each holding different types of DPs. These recipe types are:

- **CONFIGURE_MAIN**, which sets the majority of DPs, including the PILOT DACs, and the pixels which need to be masked;

- **CONFIGURE_TIMING**, which sets the TTCrx coarse-time delay and the fine-time delays of the two deskewed clocks;

- **CONFIGURE_PINTINIT**, which sets the strobe and test pulse registers, as well as the *config register* that configures the L0 electronics to readout in either ALICE or LHCb mode;

- **CONFIGURE_TEST**, which is used to store settings used during certain calibration runs such as the coarse and fine time alignment.

During the board configuration, recipes of the form *Activity/Recipe Type* are searched for, *e.g.* ALICE/CONFIGURE_PINTINIT. If such a recipe does not exist, then the default recipe is loaded instead, such as DEFAULT/CONFIGURE_PINTINIT. The exception is the CONFIGURE_TEST recipe, which is only loaded if it is found for the specified Activity.

By convention, the calibration Activities are prefixed with "CALIB|", to indicate that they are to be used for calibration purposes only. For example, the Activity used to perform a coarse time scan is called *CALIB|COARSETIMEALIGN*. Usually, only the CONFIGURE_TEST and CONFIGURE_PINTINIT recipe types are implemented for the calibration Activities, to ensure that the remaining settings are the same as those used during normal data-taking. The CONFIGURE_PINTINIT recipe needs to be implemented for calibration runs which need to be performed in ALICE mode.

3.4 Time alignment to pulsed laser data

The timing profiles of the HPDs were determined prior to installation using a test vessel with a pulsed laser installed, triggered by a TTC calibration broadcast command [67]. A pulsed laser has been installed in the radiation-free LHCb barracks in order that similar timing studies can be performed *in situ*. The laser light is introduced into the RICH detectors via optical fibres, using optical splitters to distribute the light to each HPD enclosure. These optical fibres are included in the same fibre bundles used to send data from the GOLs to the L1 electronics. The fixed time for the light to travel through the optical fibres introduces an additional delay to the system compared to *pp* collision data. This delay is not necessarily the same for the two RICH1 boxes nor the two RICH2 detector sides, so the pulsed laser system is used to align the L0 boards in a particular HPD enclosure relative to one another. A coarse time alignment is performed first to ensure the strobe signal and event information (BXID and L0ID) are sent in the correct clock cycle. The fine time alignment is then performed to maximise the number of pixel hits which are read out by the two HPDs on each L0 board.

Since the optical fibres introduce artificial delays, the detectors need to be realigned for data-taking. However, once the artificial delay between the detector boxes/sides is removed, it is sufficient to shift the coarse and fine delays of all L0 boards in the RICH detector simultaneously and integrate hits over the entire detector rather than individual L0 boards.

Using the pulsed laser system for the relative time alignment of the L0 boards and integrating hits to determine the optimal delay for data-taking means that the alignment can be performed with significantly fewer events than would be needed if the coarse and fine time scans were performed for individual L0 boards using *pp* collision data.

The coarse and fine time scans follow similar procedures. In both cases, after a set number of events, the ODIN issues a *STEP* command. This sends an event ID reset, and increments an additional counter called the *step number*. During calibration runs, the RICH L0 DIM servers interpret the STEP command as a signal to adjust settings on the L0 boards, such as incrementing the coarse time delay. The number of events per step and the total number of steps can be adjusted from the ODIN TFC control panel. The step number (along with the event number and BXID) is stored in the event packet, so it can be accessed during offline analysis of the data.

For the pulsed laser time scans, the ODIN is configured to generate *calibration C triggers*. The calibration C trigger broadcast command is used to pulse the laser. The trigger delay parameter on the TFC control panel is then used to set the number of clock cycles between the broadcast trigger command and the L0 accept, to account for the signal propagation time, L0 latency and arrival time of the pulses.

To inform the DIM server of what type of calibration is being performed, a *CalibrationType* DP is set for each L0 board. These DPs are of integer type, with each value corresponding to a different calibration type. Various other DPs are used to configure the calibration steps, and the DPs which are queried depends on the calibration type.

The coarse and fine time alignment of the RICH detectors using the pulsed laser system was originally performed by Nick Styles (University of Edinburgh) [2], using the PVSS calibration recipes produced by Sean Brisbane (University of Oxford) [1]. However, as previously mentioned, several HPDs were subsequently replaced, and the RICH group routinely performs timing scans after major interventions. Thus, coarse and fine time scans using the pulsed laser system were performed shortly before the first proton beams were collided in the LHC in late 2009. These are discussed in the remainder of this section.

3.4.1 Coarse time scan

For the coarse time scan, the relevant configuration DPs are the *CoarseTimeStart* and *CoarseTimeStop*, which have units of 25 ns clock cycles. In the first step, the coarse time delay of all boards is set to CoarseTimeStart. After the first STEP command is received, the coarse time delay is incremented by one clock cycle. This procedure is repeated until CoarseTimeStop is reached.

After some trial-and-error, the calibration C trigger delay was set to 165 clock cycles, and the number of events per step was set to 2,000 (30,000 events in total).

The CALIB|COARSETIMEALIGN/CONFIGURE_TEST recipe is used to set the CalibrationType, CoarseTimeStart and CoarseTimeStop DPs. To cover the full range of possible coarse delays, the CoarseTimeStart and CoarseTimeStop were set to 0 and 15 respectively.

3.4.2 Fine time scan

For the fine time scan calibration type, the DPs *FineTimeStart*, *FineTimeStop* and *FineTimeStep* are used, which all have units of ps. In the first step, the coarse time delay is set to the value in the CONFIGURE_TTCRX recipe, and the fine time delay is set to FineTimeStart, which can be a positive or negative number (or 0). If the fine time is below 0, or above 25 ns, the coarse time is decremented or incremented. After the first STEP command is received, the fine time delay is incremented by FineTimeStep. For each subsequent STEP command, the fine time delay is incremented by the same amount until FineTimeStop is reached (again, incrementing the coarse time delay where appropriate).

The CalibrationType, FineTimeStart, FineTimeStop and FineTimeStep are set in the CALIB|FINETIMEALIGN/CONFIGURE_TEST recipe. To encapsulate the full timing profile of the L0 boards, the FineTimeStart and FineTimeStop were chosen to be -25 ns and +50 ns. The FineTimeStep can be as small as 100 ps, but was chosen to be 1 ns as a compromise between the sensitivity of the scan and the time needed to perform the scan. The number of events per step was chosen to be 20,000 (15 million events in total). Test runs with fewer events had an insufficient number of events to determine the optimal delay for the majority of HPDs in the low illumination region of RICH2. The issue of poor laser illumination in RICH2 is discussed in more detail in Sec. 3.4.4.

3.4.3 Offline analysis of the timing scans

The offline analysis of the timing-scan runs was performed using the PANOPTES software project, which is used for the online monitoring and commissioning of the RICH detectors and can also be used to perform offline analyses. A dedicated package called *RichTimingScan-Analysis* was created by Conor Fitzpatrick (University of Edinburgh) for this purpose [68], developed from code written by Nick Styles. The code analyses the event number, step number and number of hits for each L0 board, and determines the optimal timing for each board based on the timing profile. The individual timing profiles for the boards are also stored in histograms, so that the analyst can check the results and make any necessary adjustments to the suggested optimal coarse and fine time delay. The results are saved as a text file in a format which can be interpreted by a PVSS panel, used to set the coarse and fine time delay DPs.

The ideal timing profile of the L0 boards is a flat-top distribution, with a 25 ns width corresponding to the length of the strobe signal. In reality, the time taken to pass the discriminator threshold depends on the charge deposited on the pixel, which is reduced by charge sharing and backscattering. These effects give the timing profile of the L0 boards a distinctive structure: a flat or sloped summit and asymmetric tails. The majority of the timing profile lies within a single clock cycle, though the tails may extend beyond this. When analysing the timing scan results, it is important to have a method for determining the optimal delay. The *turn on* of the timing profile is defined as the point at which the number of hits reaches 90% of the maximum, *i.e.* $0.9 \times (N_{\mathrm{max}} - N_{\mathrm{min}})$, where N_{max} is the maximum number of hits, and N_{min} is the background level (the number of hits far from the tails, which should ideally be close to zero). Similarly, the *turn off* is the point at which the number of hits drops back down to 90% of the maximum. The *midpoint* is defined as the point equidistant from the turn on and turn off. The turn on, turn off and midpoint are indicated for one of the L0 boards in Fig. 3.5. The midpoint is taken to be the optimal operation point on the timing curve.

3.4.4 Results of the pulsed laser timing scans

The turn on, midpoint and turn off for each L0 board are shown in Figs. 3.6a and 3.6b. The results indicate that all RICH1 boards turn on in the same clock cycle, 170 clock cycles after the laser pulse. The majority of RICH2 boards turn on three clock cycles later. This is an artificial relative delay, caused by the aforementioned difference in the optical fibre lengths.

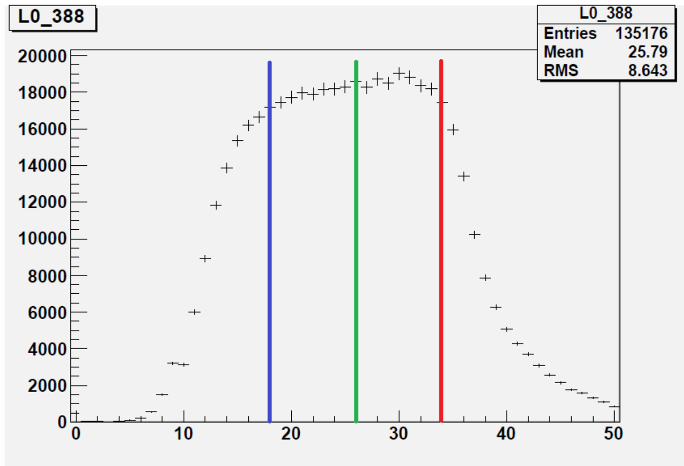
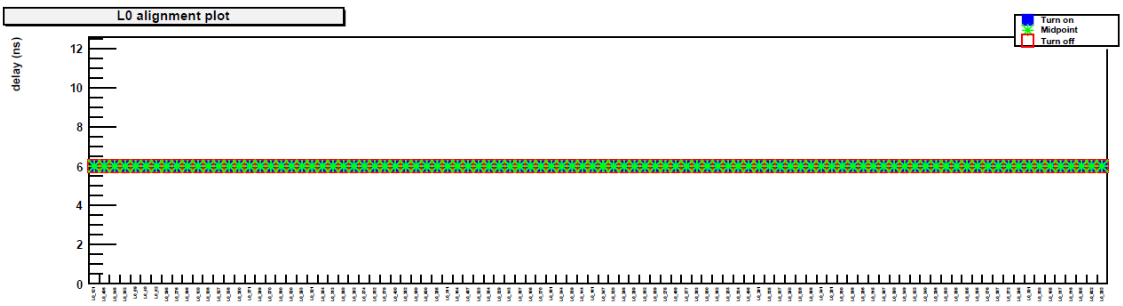
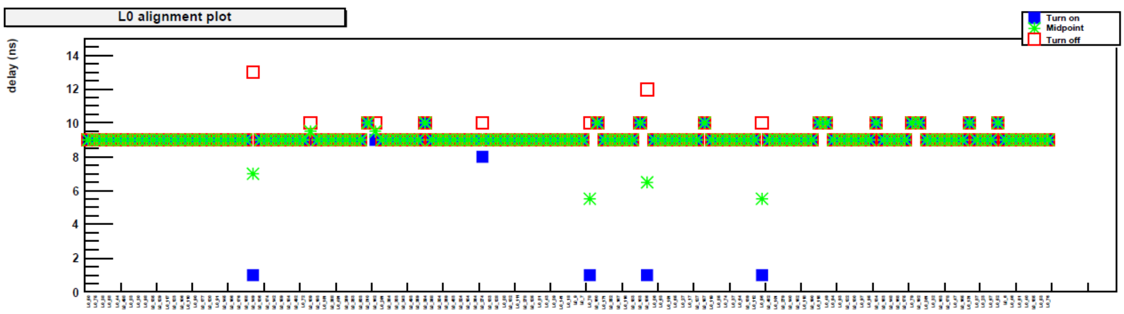


Figure 3.5: Timing profile (number of pixel hits *vs.* time delay) of a RICH L0 board. The turn on (blue), midpoint (green) and turn off (blue) of the timing profile are indicated by the vertical lines. The x-axis indicates the delay in ns.



(a)



(b)

Figure 3.6: Coarse time scan alignment plots for the L0 boards of (a) RICH1 and (b) RICH2. The turn on, midpoint and turn off of the timing profile for each board are shown.

Several of the remaining RICH2 boards have relatively large time differences between the turn on and turn off. In most cases, this is a result of poor laser illumination. To understand why, it is informative to look at the online monitoring. Figure 3.7 shows a snapshot of the HPD hit maps for the two RICH2 detector sides taken during the coarse time scan. Each point on the plot represents the accumulated hits for a cluster of pixels. The optical fibres need to be located outside of the active part of the detector to avoid disrupting the path of the photoelectrons during normal data-taking. However, the location of the optical fibres in RICH2 means that the photon funnel and magnetic shielding, which are indicated at the bottom of Fig. 2.8b, cast a shadow on some of the HPDs. This results in the lack of illumination seen at the periphery of the detector in Fig. 3.7. An example of a board with poor illumination can be seen in Fig. 3.8a. In this case, any hits from the laser light are completely swamped by the background hits. The number of boards for which a valid turn on could not be determined was minimised by using a large number of events per step, to counteract the low hit efficiency of the poorly illuminated boards. For the remaining boards such as the one shown in Fig. 3.8a, the coarse and fine delays were set to the average of the boards which had sufficient hits above the background level to determine the turn on.

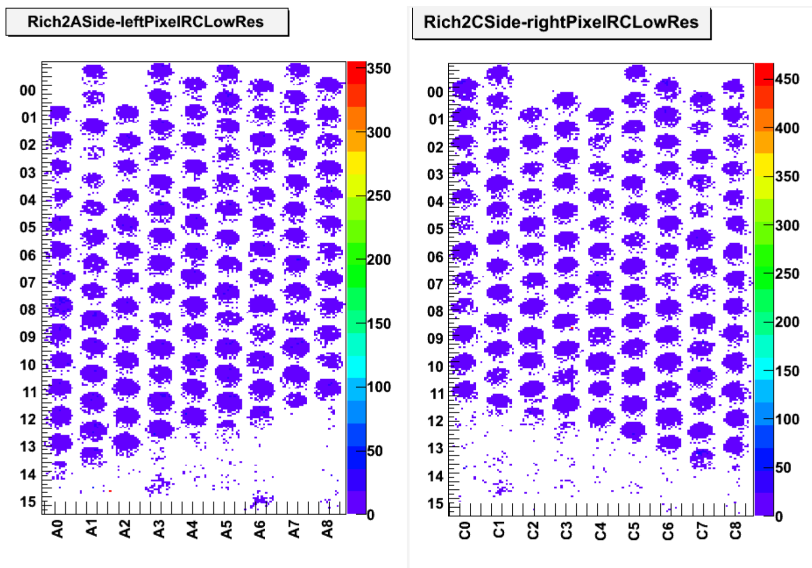
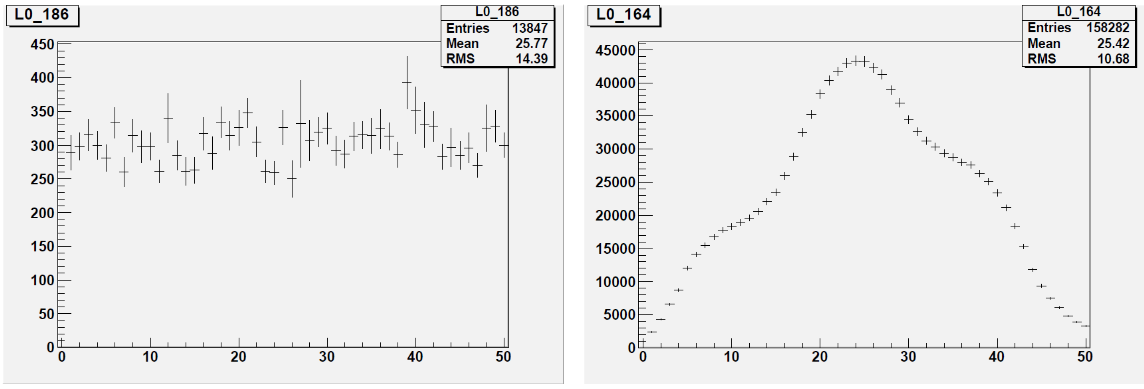


Figure 3.7: Snapshot from the RICH online monitoring showing hit maps of the HPDs of the left and right RICH2 detector sides.

Another issue with the time alignment was that, as previously mentioned, some HPDs had to be replaced prior to performing the scans. Due to the limited number of spare HPDs, it was not always possible to match the timing profiles of the two HPDs on the L0 board. An example of the timing profile of an L0 board with badly matched HPDs can be seen in Fig. 3.8b. These boards were dealt with on a case-by-case basis. In most cases, the delay was set to the midpoint determined by the offline analysis code. In a few cases where one HPD received substantially more hits than the other and the difference in turn on between the two HPDs was particularly large, the delay was shifted by a few ns to align closer to the midpoint of the HPD with the higher hit efficiency.



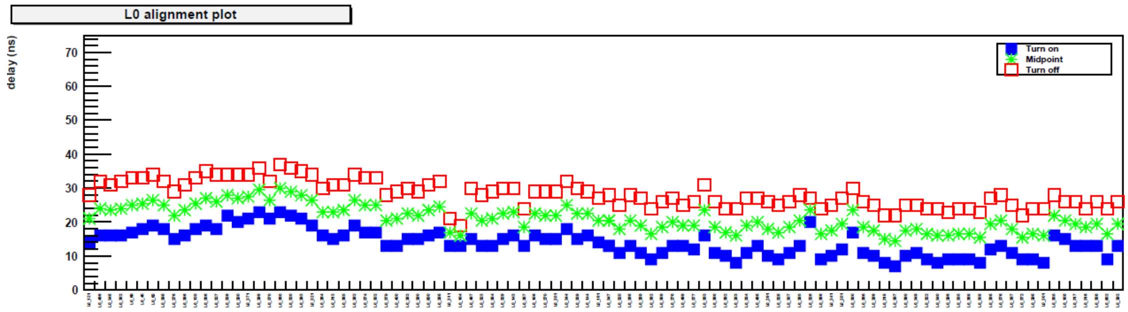
(a) (b)

Figure 3.8: Timing profile of (a) a RICH2 L0 board with poor laser illumination, and (b) a RICH2 L0 board where the two HPDs have a large difference in turn on times. In both cases, the x-axis has units of ns.

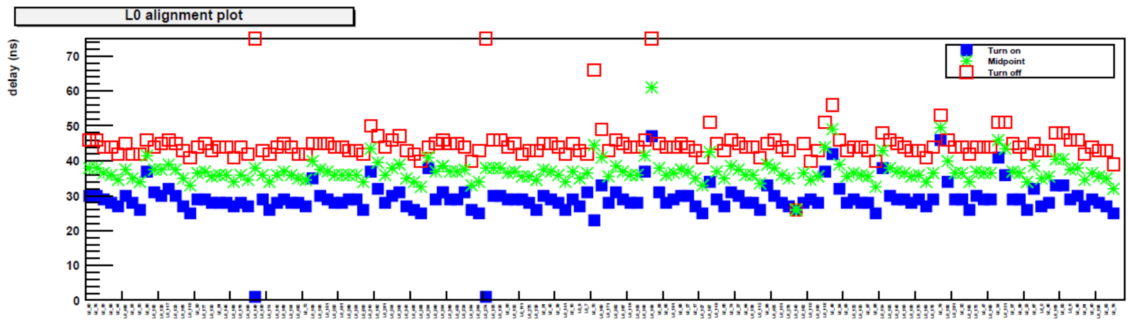
The alignment plots from the fine time scan can be seen in Fig. 3.9. The zero point on the vertical axis corresponds to one clock cycle earlier than the nominal coarse time delay according to the results of the coarse time scan.

During the processing of the timing scan results, it was decided to produce a new type of calibration scan similar to the fine time scan, but in which the delays are adjusted relative to the coarse **and** fine time delays in the DPs, rather than just the coarse time delay. This type of scan is useful to confirm that the results of the coarse and fine time scans were applied successfully. The DIM server code was updated to implement this new type of scan, and a new calibration recipe called *CALIB|TIMESCANATNOMINAL* was created to set the scan parameters.

In the CALIB|TIMESCANATNOMINAL recipe, the FineTimeStart was set to -25 ns,



(a)



(b)

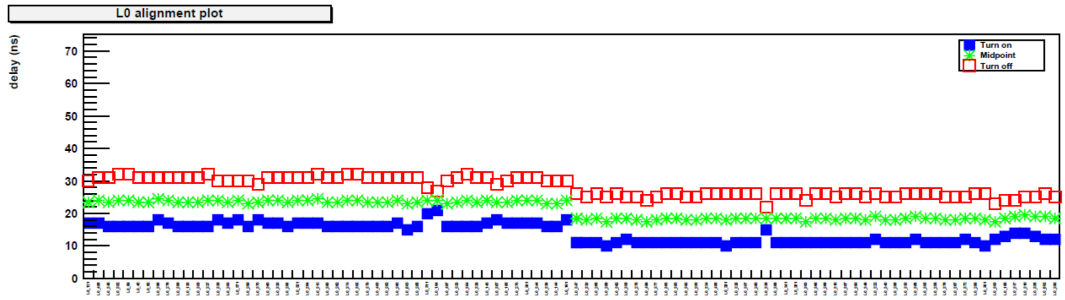
Figure 3.9: Fine time scan alignment plots for the L0 boards of (a) RICH1 and (b) RICH2. The turn on, midpoint and turn off of the timing profile for each board are shown.

the FineTimeStop to +50 ns and the FineTimeStep to 1 ns. The results of this scan can be seen in Fig. 3.10. The zero point on the vertical axis corresponds to 25 ns earlier than the nominal coarse+fine time delay. The results indicate that apart from the few problematic boards in RICH2 that had particularly low laser illumination, the maximum deviation of the midpoints of the boards from the median is around ± 2 ns, with the majority of boards aligned to within 1 ns.

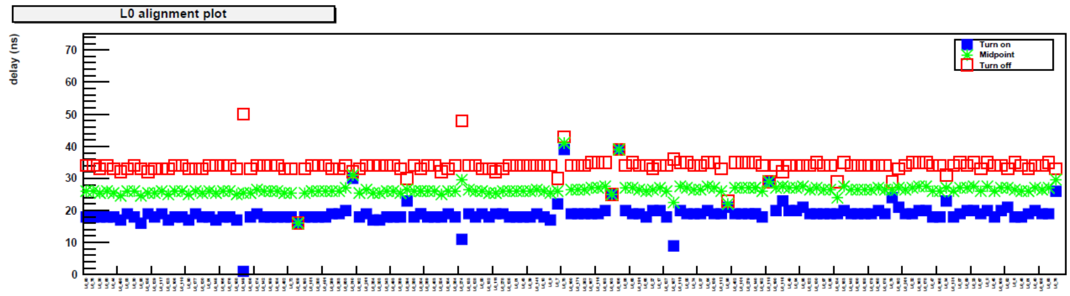
The 6 ns shift between the left and right hand side of Fig. 3.10a is a result of the differing fibre lengths of the upper and lower RICH1 detector boxes, which was corrected for prior to the time alignment to *pp* collision data.

Once the nominal coarse and fine delays have been calculated, the PVSS DPs need to be set and stored in the recipes. The PANOPTES package outputs a text file which is used to apply the time-scan results to the DPs.

For the coarse time scan, each line in the file specifies the unique L0 board ID and the coarse time delay. The board ID is converted by a PVSS panel into the name of the PVSS



(a)



(b)

Figure 3.10: "Time scan at nominal" alignment plots for the L0 boards of (a) RICH1 and (b) RICH2. The turn on, midpoint and turn off of the timing profile for each board are shown.

DP which needs to be set, and the specified delay is then stored in the DP.

For the fine time scan, each line in the file specifies the board ID and the fine time delay that needs to be applied to shift this board to the average midpoint.

Finally, for the "time scan at nominal", each line in the text file indicates the board ID and the fine time shift relative to the current value stored in the DPs which needs to be applied to shift the board to the new nominal delay according to the scan. This is then converted to the new coarse and fine time values which need to be set to apply this shift.

3.5 Time alignment to *pp* collision data

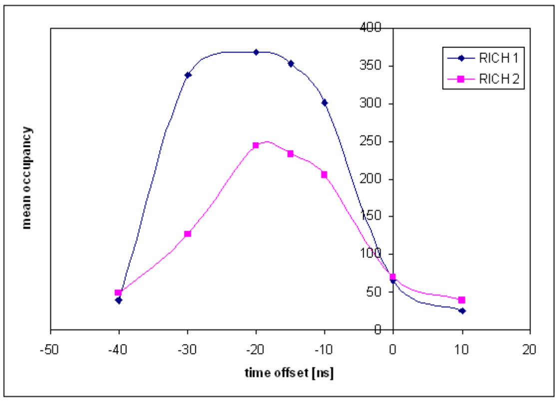
When the LHC successfully collided the first proton beams at the LHCb IP, the RICH detectors were aligned to the LHC clock by incrementing the coarse time delay for all L0 interface boards simultaneously until Cherenkov rings were detected by the RICH online monitoring system, using a PVSS panel designed by the author. The RICH detectors were controlled and monitored separately to the rest of the LHCb detector by using a separate

ODIN (see the explanation of the detector partitioning in Sec. 2.9.3).

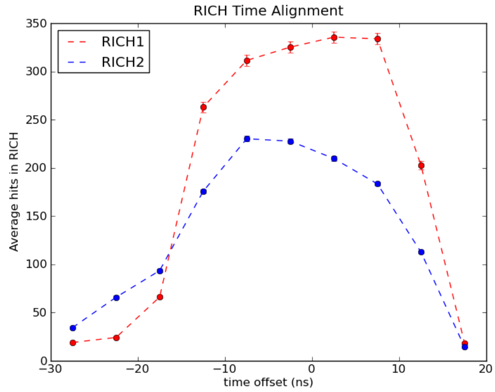
After the RICH detectors were coarsely aligned to beam collisions, a series of runs were taken in which the timing of all the L0 boards was shifted simultaneously using the aforementioned PVSS panel. After each run, the delay was shifted by 5 or 10 ns, with a total scan range $[-50, 20]$ ns relative to the delays stored in the PVSS recipe. The data were analysed using a simplified version of the RichTimingScanAnalysis code, which simply recorded the number of hits in each RICH detector. This information was then collated, divided by the total number of events in each run and plotted. The results are shown in Fig. 3.11a, where the offset is the shift in the delay relative to the values stored in the recipe. Based on these results, the delays of the L0 boards in RICH1 were shifted 20 ns earlier and the boards in RICH2 were shifted 15 ns earlier.

After this initial analysis, a series of runs with approximately ten times the number of events of the first were taken to confirm that the previously applied fine time shifts resulted in the optimal timing. The data were taken with 5 ns increments in the delays over a range $[-27.5, 17.5]$ ns. The results are shown in Fig. 3.11b. As before, the offset is the shift in the delay relative to the values stored in the recipe. These results indicate that the RICH1 L0 boards turned on 5 ns too late, whilst RICH2 turned on around 5 ns too early.

The data were subsequently analysed using the RichTimingScanAnalysis package. The results can be seen in Fig. 3.12, which confirm the previous conclusions for RICH1, but indicate a wide variation in the midpoints for RICH2. This variation is mainly due to boards which have low occupancy in pp collision data, *i.e.* those boards in the large acceptance angle region, on the top and bottom edges of the RICH2 detector sides, where the particle flux is low. These results were obtained shortly before the LHC ceased operation over the winter. It was decided to confirm the alignment of RICH2 after the winter shutdown by performing an automated time scan at the nominal values.

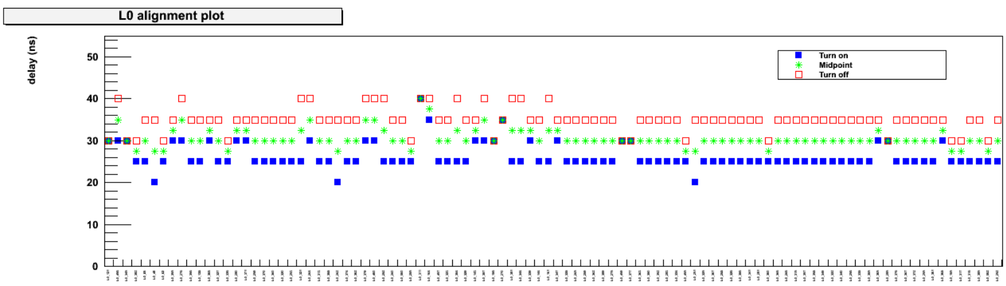


(a)

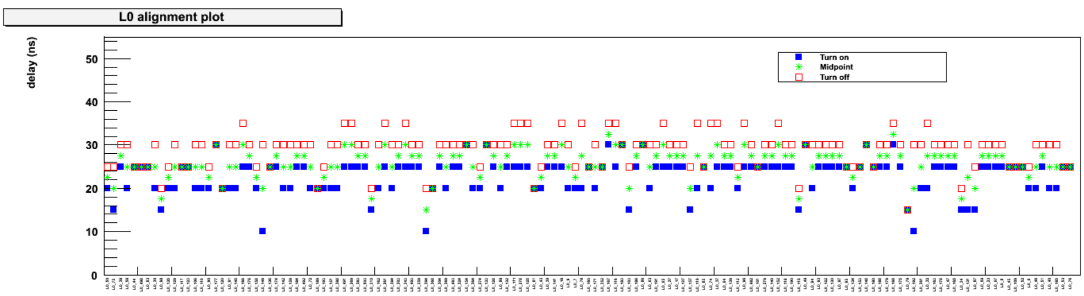


(b)

Figure 3.11: Results of the runs used to time align the RICH detectors to proton beam collisions. Each data point indicates the average number of hits per event for a given fine-time offset relative to the delays stored in the PVSS recipe. (a) shows the results of the initial low event-number runs (350 events per data point), whilst (b) shows the results of later runs with a greater number of events (3,000 events per data point, apart from the left-most point, which corresponds to 1,920 events).



(a)

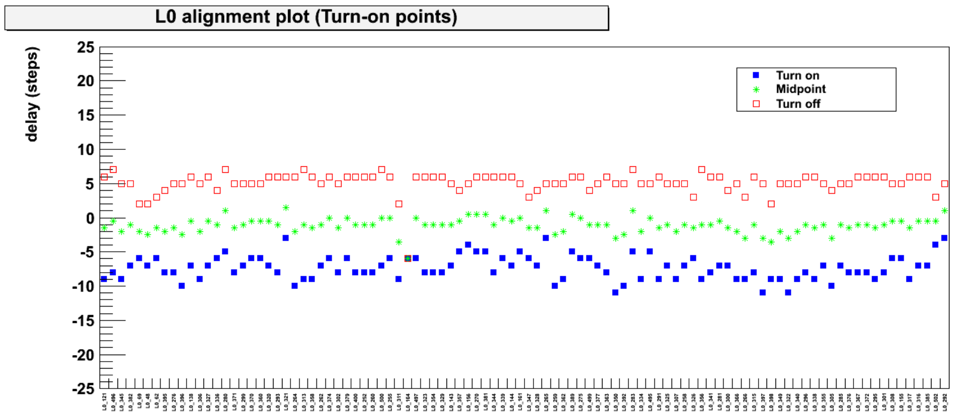


(b)

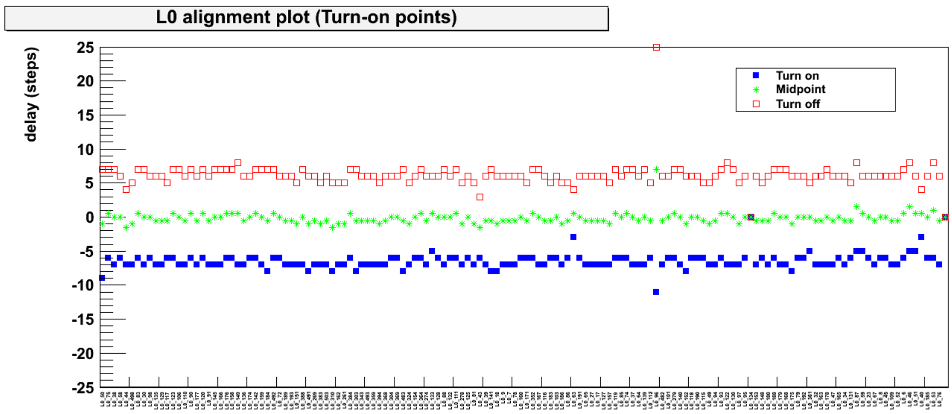
Figure 3.12: Time scan at nominal delay alignment plots for the L0 boards of (a) RICH1 and (b) RICH2 determined from a manual time-scan using $\sqrt{s} = 900\,\text{GeV}$ *pp* collision data. The turn on, midpoint and turn off of the timing profile for each board are shown.

3.6 Time scan at nominal delays using pp collision data

After the 2009 winter shutdown, two sets of time scans were performed after setting the nominal delays from the pulsed laser scans and applying the shifts discussed in Sec. 3.5. Each scan comprised 5,000 $\sqrt{s} = 7\,\text{TeV}$ pp collision events per step, and the accumulated results are presented in Fig. 3.13. The results indicate that both RICH1 and RICH2 were well aligned as a result of the pulsed laser scans, with the RMS spread of the midpoints of the boards corresponding to 1.1 ns for RICH1 and 0.8 ns for RICH2.



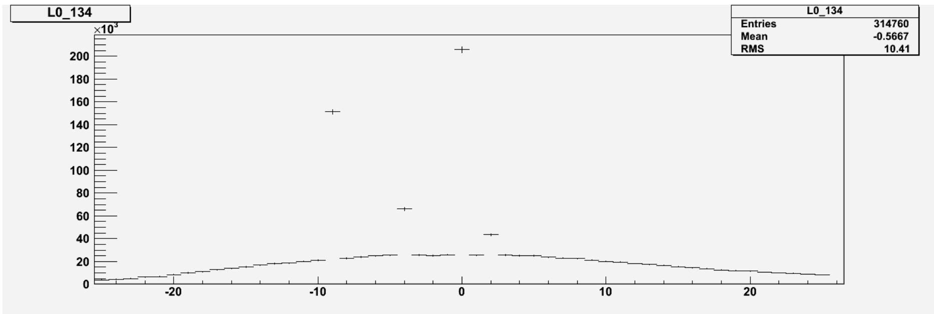
(a)



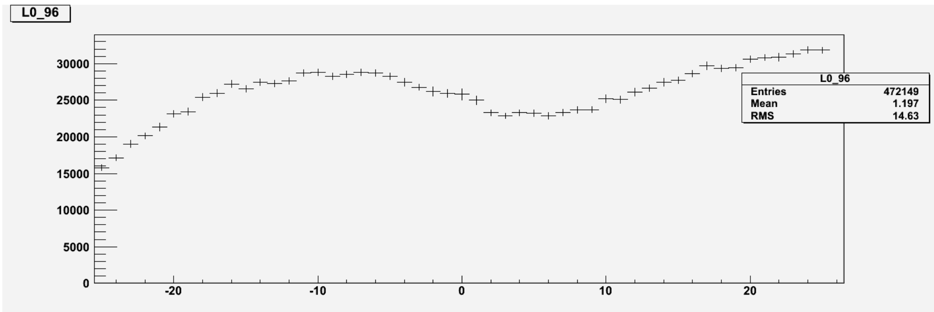
(b)

Figure 3.13: Time scan at nominal delay alignment plots for the L0 boards of (a) RICH1 and (b) RICH2 determined using $\sqrt{s} = 7\,\text{TeV}$ pp collision data. The turn on, midpoint and turn off of the timing profile for each board are shown.

However, there were three problematic boards that needed to be individually investigated. Two of these boards suffered from significant detector noise during at least one time step. The timing profiles for one these boards can be seen in Fig. 3.14a. Despite the noise, there were enough "good" time scan points to conclude that both boards were correctly timed.



(a)



(b)

Figure 3.14: Timing profile of two badly behaving L0 boards according to the timing scan at nominal delays using $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data. The L0 board shown in (a) suffered from significant detector noise, whilst for the L0 board shown in (b), the turn on of the two HPDs differed by more than one clock cycle. The x-axes have units of ns.

The final board was more problematic. The timing profile for this board is indicated in Fig. 3.14b. As can be seen, the difference in the turn on of the two HPDs on this board was significantly more that one clock cycle (around 40 ns peak-to-peak). The problem was solved by reducing the delay_control of the later HPD by two clock cycles. This reduced the difference in turn on times between the two HPDs to 10 ns, with the late HPD now turning on earlier than the other HPD. Finally, the board delay was shifted 10 ns earlier in the CONFIGURE_TTCRX recipe to bring the board into alignment with the other L0 boards.

3.7 Summary

The RICH detectors have been successfully time aligned, using a combination of pulsed laser timing scans to align the L0 boards relative to one another, and manual and automated time scans with *pp* collision data around the nominal delays to ensure that the boards are correctly time aligned for beam collisions.

Several issues were identified during the analysis. Provision had to be made for low laser illumination in some of the RICH2 HPDs during the pulsed laser scans, low occupancy in some HPDs in the high angular acceptance region of RICH2 in *pp* collision data, detector noise and badly paired HPDs. However, following manual intervention in some cases, the results demonstrate that the L0 boards of the two RICH detectors have been aligned with an RMS spread of around 1 ns.

Chapter 4

Measurement of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A_{fav}^{K3\pi}$ from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ decays

Details of the analysis to measure the ratio of branching factions $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ decays from data collected by the LHCb experiment are outlined. A result of $0.057 \pm 0.007 \pm 0.004$ is found. In addition, the rate asymmetry $A_{fav}^{K3\pi}$ is measured, and a value of $0.01 \pm 0.11 \pm 0.06$ is determined.

4.1 Chapter outline

This chapter outlines the measurement of the ratio of branching fractions $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ decays[1] using data collected by the LHCb experiment.

The data samples analysed are outlined, and the choice of selection cuts are indicated. The consistency of the distributions of variables that are cut on in the final selection according to the MC signal candidates (used to determine the optimal cuts and measure the various efficiency corrections) and the candidates from data (used to measure $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A_{fav}^{K3\pi}$) is then considered. The efficiencies of the triggers at each level are then analysed, and the choice of trigger lines is discussed. The *probability density functions* (PDFs) used to estimate the signal and background $B^{\pm}$ mass distributions are detailed, as is the calibration of the *particle*

[1] Here and subsequently, h indicates that decays involving both a kaon or pion *bachelor* (the final-state track from the B decay) are considered, whilst D refers to either a D^0 or $\overline{D}^0$ meson decaying to the same final state.

identification (PID) selection and the fit procedure. Finally, a summary of the systematic uncertainties associated with the yields, efficiency corrections and fit procedure are given, and the final results presented.

4.2 Introduction

As described in Chap. 1, the interference between $B^{\pm} \to DK^{\pm}$ decays in which the D^0 and $\overline{D}^0$ decay to the same final state is potentially a powerful method for determining the CKM angle γ. In particular, the series of "wrong sign" (WS) decays (decays in which the bachelor kaon and the kaon from the D decay have opposite charge) are expected to have a strong sensitivity to γ.

The determination of γ from the WS decay $B^{\pm} \to (K^{\mp}\pi^{\pm})_D K^{\pm}$ can be extended to that of multi-body decays of the D, such as $D^0 \to K^+\pi^-\pi^+\pi^-$ (and its charge conjugate). However, in this case the D^0 and $\overline{D}^0$ can decay to the same final state through a multitude of intermediate resonances, each of which can have a different amplitude and strong phase, significantly complicating the analysis. However, it has been found [28] that the issue can be simplified to a pseudo two-body problem involving a single additional parameter called the *coherence factor*, C_f, which represents the interference between the various intermediate resonances of the $D^0/\overline{D}^0$. The coherence factor and its associated average strong phase, defined in Sec. 1.6.2, have been measured using $D^0 \to K^+\pi^-\pi^+\pi^-$ (and $D^0 \to K^+\pi^-\pi^0$) quantum-correlated D^0 decays by the CLEO-c collaboration [29].

At the time this analysis was performed, LHCb did not have a sufficient sample of WS $B^{\pm} \to (K^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_D K^{\pm}$ decays to make a measurement of γ. However, several measurements can be made using the kinematically equivalent "right sign" (RS) decay mode (in which the bachelor kaon and the kaon from the D decay have the same charge) $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$, which is expected to be at least two orders of magnitude more abundant than the WS mode.

This chapter describes the measurement of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ decays from $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data accumulated at LHCb in the first full year of operation (2010). The rate asymmetry for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays, $A_{fav}^{K3\pi} = \frac{\Gamma(B^{-}\to(K^{-}\pi^{+}\pi^{-}\pi^{+})_D K^{-})-\Gamma(B^{+}\to(K^{+}\pi^{-}\pi^{+}\pi^{-})_D K^{+})}{\Gamma(B^{-}\to(K^{-}\pi^{+}\pi^{-}\pi^{+})_D K^{-})+\Gamma(B^{+}\to(K^{+}\pi^{-}\pi^{+}\pi^{-})_D K^{+})}$, is also measured. As discussed in Sec. 1.6.3, the ratio of branching fractions can also be measured from $B^{\pm} \to (K^{\pm}\pi^{\mp})_D h^{\pm}$ decays, and $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ has been determined at LHCb using these decays, and also using a combination of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp})_D h^{\pm}$ decays [3]. The analysis strategy described in the following sub-section is similar to that employed by the analysis in Ref. [3], and the results from the two analyses are compared in the final section of this chapter.

4.2.1 Analysis strategy

$\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ requires the determination of the ratio of the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ yields, whilst $A_{fav}^{K3\pi}$ requires the yields of $B^{+} \to (K^{+}\pi^{-}\pi^{+}\pi^{-})_D K^{+}$ and $B^{-} \to (K^{-}\pi^{+}\pi^{-}\pi^{+})_D K^{-}$ to be measured separately. These yields are extracted from maximum likelihood fits to the B invariant mass spectra of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ mass combinations in the data. This is referred to as the "final fit" (to distinguish it from intermediate mass fits used to study the components of the final fit) and is described in detail in Sec. 4.9. The analysis is complicated by several factors, which will now be discussed.

Firstly, due to the large number of tracks in the final state, the level of *combinatoric* background (the combination of random tracks from the event reconstructed to form a fake vertex) is expected to be large for the multi-body modes (compared to $B^{\pm} \to (K^{\pm}\pi^{\mp})_D h^{\pm}$ for example). The contamination of this background in the sample of reconstructed $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ candidates is significantly reduced by applying a series of selection requirements, hereby referred to as *cuts* prior to performing the fits. This brings its own complications, since any difference in the *selection efficiency* (the fraction of correctly identified signal candidates passing all cuts) between $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ needs to be corrected for when extracting $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ from the ratio of the yields. The procedure to determine the optimal cuts and selection efficiencies is discussed in Sec. 4.3.

Secondly, as discussed in Sec. 2.8, LHCb utilises a series of hardware and software triggers in order to reduce the LHCb data rate from around 40 MHz to $\approx$ 2 kHz before the events are written to mass storage. Inevitably, this process will not be efficient at retaining all decays of physics interest. In addition, the trigger selection requirements are often changed depending on the detection conditions. Thus, an understanding of the efficiency of the trigger is an important part of the analysis. Since $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ are kinematically equivalent modes, the *trigger efficiency* should be largely independent of which of the two modes was triggered. However, as discussed in Sec. 4.4, some difference in trigger efficiency is seen, and this is also corrected for in the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement.

Thirdly, since the branching fraction for $B^{\pm} \to D\pi^{\pm}$ is an order of magnitude larger than $B^{\pm} \to DK^{\pm}$, $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ candidates where the bachelor pions are incorrectly reconstructed as kaons are a major source of background for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$. This is dealt with by utilising the discriminating power of the LHCb RICH detectors and by including a "mis-ID" component in the final fit, which describes the B invariant-mass distribution of this type of background. The choice of signal and mis-ID components are discussed in Sec. 4.7. The determination of the performance of the *particle identification* (PID) requirements used to distinguish between bachelor kaons and pions is outlined in Sec. 4.8.

Finally, in addition to the previously mentioned combinatoric and mis-ID backgrounds, various other sources of background can contaminate the reconstructed B mass distribution. These need to be thoroughly understood, and either be removed/minimised through additional cuts, or accounted for in the final fit. The study of these backgrounds is described in Sec. 4.6.

4.2.2 Physics motivation

$\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ has been measured to a high degree of accuracy by both BELLE [69] and BABAR [70], but the two results differ by approximately 2σ (see Fig. 4.1). The current world average of the ratio of branching fractions, determined by the Particle Data Group

(PDG) [23], is $(7.6 \pm 0.6)\%$, corresponding to a relative uncertainty of $\sim 8\%$. This analysis aims to provide a competitive measurement of this observable.

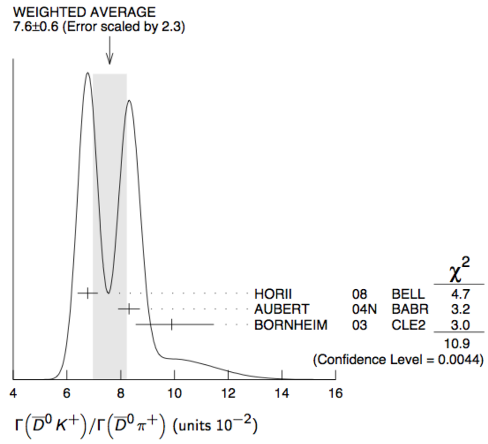


Figure 4.1: Likelihood profile of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ provided by the PDG.

The $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ RS decay has very little sensitivity to γ, so the rate asymmetry between $B^+ \to (K^+\pi^-\pi^+\pi^-)_D K^+$ and $B^- \to (K^-\pi^+\pi^-\pi^+)_D K^-$ due to *CP*-violation, $A^{K3\pi}_{fav}$, should be ≈ 0. However, it has the same form as the γ-sensitive asymmetry $A^{K3\pi}_{ADS}$ (see Sec. 1.6.2) with the RS $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ decays exchanged with the WS $B^\pm \to (K^\mp\pi^\pm\pi^\mp\pi^\pm)_D K^\pm$. The analysis of $A^{K3\pi}_{ADS}$ is therefore expected to be similar to the $A^{K3\pi}_{fav}$ measurement presented here, with many of the same challenges (effective suppression of combinatoric background events, as well as a necessarily good understanding and modelling of the various background sources). As such, this analysis lays down much of the ground-work for the determination of γ from the study of $B^\pm \to (K^\mp\pi^\pm\pi^\mp\pi^\pm)_D K^\pm$ decays.

4.2.3 Variable definitions

The variables used throughout this chapter and Chap. 5 to discriminate signal from background are described below.

IP χ^2

The χ^2 significance of the reconstructed impact parameter (*IP*) of the particle with respect to the primary vertex (PV). The *IP* χ^2 will tend to be small for particles originating from the PV (such as the $B^\pm$ in this analysis) compared to background processes, whilst particles originating from a vertex displaced from the PV (*e.g.* the decay products of a B or D meson) will tend to have a larger $IP\,\chi^2$. If there is more than one PV in the event (due to multiple pp collisions per bunch crossing) then the $IP\,\chi^2$ is evaluated with respect to the "best" PV according to the vertexing tool.

$sIP\,\chi^2$

The $IP\,\chi^2$ of the particle with respect to the PV that returns the smallest $IP\,\chi^2$. This parameter is sometimes used instead of the $IP\,\chi^2$ with respect to the best PV when applying selection requirements to particles originating from secondary vertices.

$VD\,\chi^2$

The χ^2 significance of the reconstructed distance between the best PV and the decay vertex of the particle. Cuts on the minimal value of the $VD\,\chi^2$ of B and D mesons are useful for rejecting background decays in which the final-state tracks originate from the PV.

$B^{\pm}$ $DoCA$

The distance of closest approach ($DoCA$) between the trajectories of the reconstructed D^0 and the *bachelor* track (the K or π from the top-level of the $B^{\pm}$ decay). A cut on the $B^{\pm}$ $DoCA$ is generally applied in stripping selections prior to fitting the B decay vertex to reduce the number of track-vertex combinations that need to be considered by the vertexing tool.

D^0 max ($DoCA$)

The maximum distance of closest approach ($DoCA$) of the various two-track combinations of the D^0 decay products.

Vertex $\chi^2/N_{\rm dof}$

The χ^2 per degree of freedom of the reconstructed $B^{\pm}$ or D^0 decay vertex.

Track $\chi^2/N_{\rm dof}$

The χ^2 per degree of freedom of the reconstructed $K^{\pm}$ or $\pi^{\pm}$ track.

DIRA

The cosine of the angle between the direction of flight of the $B^{\pm}$ or D^0 (according to its production and decay vertices) and its momentum vector. If the tracks and vertices are perfectly reconstructed, then the *DIRA* will be unity.

Mass window

The range of $B^{\pm}$ or D^0 invariant mass values in which the majority of reconstructed signal candidates are expected to be distributed. The mass window is generally kept quite loose in the stripping so that sources of background can be investigated and either removed by further cuts or parametrised in a fit to the data. In this thesis, the mass window is generally specified with respect to the *nominal mass* (the B or D invariant mass according to the PDG).

Side-bands

The regions of the $B^{\pm}$ or D^0 invariant-mass distribution either side of the mass window, which are expected to be populated only by background decays and the random combination of tracks.

4.3 Selection

The final (offline) selection comprises a series of cuts applied to the $K^{\pm}$ and π^- track kinematic distributions, as well as cuts to the reconstructed $B^{\pm}$ and D vertices. The selection cuts are optimised to maximise the figure of merit $S/\sqrt{S+B}$, where S and B are the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ signal yield and background contamination in data respectively. As will be discussed below, the signal yield is determined using MC signal, so S is estimated based on the number of generated MC events and the integrated luminosity of the data sample.

A recursive optimisation procedure has been employed to determine the selection cuts using the *CROP* tool developed by Conor Fitzpatrick [71]. Each cut is optimised separately to maximise $S/\sqrt{S+B}$. Once this has been achieved, each cut is varied again to find the cut that maximises $S/\sqrt{S+B}$, keeping the other cuts fixed to the value determined from the previous step. This procedure is performed multiple times until the value of $S/\sqrt{S+B}$ converges. The samples used to perform the optimisation are outlined in Secs. 4.3.1 and 4.3.2.

4.3.1 MC samples

A large sample of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ MC has been used as the signal sample for the cut optimisation. This sample, in addition to a similarly sized $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ sample, is also used to determine the various efficiency corrections (apart from the PID efficiency, which is determined using a data-driven approach outlined in Sec. 4.8). In addition, MC samples of $B^{\pm} \to (K^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp})_D \pi^{\pm}$, as well as various background samples are used to parametrise some sources of background.

Each $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ (and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$) event is generated with an average of 2.5 *pp* interactions per bunch crossing. As well as the D decaying directly to $K\pi\pi\pi$, the decay model allows the D to decay via several intermediate resonances, with the appropriate branching fractions according to the PDG. However, this decay model neglects interference between the intermediate resonances. The full list of resonances included and their branching fractions are given in Tab. 4.1. In addition, the products of each signal decay in the MC are required to have been generated within the 400 mrad LHCb angular acceptance in the lab frame. After these *generator-level* cuts, the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ MC samples comprise approximately 800,000 events each. The efficiency of the generator-level cuts is referred to as the *generator efficiency* (ε_{acc}). Any differences between ε_{acc} for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ is corrected for in the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ calculation.

After the generator-level cuts have been applied, the MC samples are reconstructed using the LHCb reconstruction software packages Brunel and DaVinci (described in Sec. 2.10.2) using the same reconstruction version[1] as the data. As discussed in Sec. 2.10.3,

[1] Reconstruction version *Reco08*.

after the LHCb events are triggered and reconstructed, a series of selection algorithms known as *stripping* lines are applied, each of which is optimised for specific physics needs. The MC samples analysed consist of the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ candidates passing the same *StrippingB2DXWithD2hhhhLine* stripping line (see Tab. 4.2) used to strip the data.

The signal MC samples used in this chapter and Chap. 5 consist of reconstructed candidates that are associated as signal according to the MC-truth information, unless otherwise stated. Hereby, these are referred to as *MC-associated* or *truth-matched* candidates.

Decay mode	BF (%)
$D^0 \to a^+(1260)K^-$	3.60
$D^0 \to \overline{K}^{*0}(892)\rho^0(770)$	1.06
$D^0 \to \overline{K}^{*0}(892)\omega^0(782)$	0.01
$D^0 \to K_1^-(1270)\pi^+$	0.29
$D^0 \to \overline{K}^{*0}(892)\pi^+\pi^-$	1.60
$D^0 \to K^-\pi^+\rho^0(770)$	0.51
$D^0 \to K^-\pi^+\omega^0(782)$	0.01
$D^0 \to K^-\pi^+\pi^-\pi^+$ non. res.	1.89

Table 4.1: Summary of the resonance model of the D^0 (and charge conjugate) used in the decay of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ candidates in the LHCb MC. $\rho^0(770)$ and $\overline{K}^{*0}(892)$ are assumed to exclusively decay to $K^-\pi^+$ and $\pi^+\pi^-$ respectively.

4.3.2 Data samples

The data samples analysed comprise the entirety of the $\sqrt{s} = 7\,\text{TeV}$ pp collision data collected in 2010, which corresponds to an integrated luminosity of $\mathcal{L}_{\text{int}} = 35.6 \pm 3.6\ \text{pb}^{-1}$.

The samples consist of candidates that pass the *StrippingB2DXWithD2hhhhLine* stripping line in the end-of-year reprocessing of the 2010 data[1]. This stripping selection is described in more detail in Sec. 4.3.3.

4.3.3 Stripping selections

The full list of cuts employed in the *StrippingB2DXWithD2hhhhLine* and *StrippingB2DX-WithD2hhLine* lines are shown in Tab. 4.2. Candidates passing the latter line are used at

[1] Stripping version *Stripping12b*.

various stages of the analysis, in particular to determine the "low mass" background PDFs, as discussed in Sec. 4.6.2. Here, and subsequently, the final-state tracks of a D decay are referred to as the D^0 *daughter* tracks.

	Stripping line cuts	
Cut variable	$B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, h^\pm$	$B^\pm \to (K^\pm\pi^\mp)_D\, h^\pm$
$B^\pm$ mass window (MeV/c^2)	± 500	± 500
D^0 mass window (MeV/c^2)	± 100	± 100
D^0 p_T (GeV/c)	> 2	> 1
D^0 daughter $K^\pm$ p_T (MeV/c)	> 250	> 250
D^0 daughter $\pi^\pm$ p_T (MeV/c)	> 150	> 250
Bachelor p_T (MeV/c)	> 500	> 500
D^0 daughter $K^\pm$ $\lvert\vec{p}\rvert$ (GeV/c)	> 3	> 2
D^0 daughter $\pi^\pm$ $\lvert\vec{p}\rvert$ (GeV/c)	> 2	> 2
Bachelor $\lvert\vec{p}\rvert$ (GeV/c)	> 5	> 5
$B^\pm$ *IP* χ^2	< 25	< 25
D^0 daughter *sIP* χ^2	> 4	> 4
max(D^0 daughter) *sIP* χ^2	> 40	> 40
Bachelor *sIP* χ^2	> 16	> 16
D^0 *VD* χ^2	> 36	> 36
D^0 max (*DoCA*) (mm)	< 1.5	< 1.5
$B^\pm$ *DIRA*	> 0.9998	> 0.9998
D^0 *DIRA*	> 0.9	> 0.9
$B^\pm$ lifetime (ps)	> 0.2	> 0.2
$B^\pm$ vertex χ^2/N_dof	< 12	< 12
D^0 vertex χ^2/N_dof	< 10	< 12
D^0 daughter track χ^2/N_dof	< 5	< 5
Bachelor track χ^2/N_dof	< 5	< 5

Table 4.2: Summary of the selection cuts applied in the stripping.

4.3.4 Cut optimisation

As previously mentioned, this analysis uses a cut-based optimisation for the offline selection of $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, K^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, \pi^\pm$ decays.

The previous selection study of the $B \to DK$ decay mode [72] was totally MC-driven, whereas this study considers $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, K^\pm$ mass combinations from the 2010 data that reside in the D^0 mass side-bands as a source of background for the optimisation. This region is expected to principally contain combinatoric background (*i.e.* the random combination of tracks in the event), which, as has been shown in the aforementioned MC study, is expected to be the dominant source of background for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays prior to the application of offline selection cuts (not including the mis-identification of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ decays as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$, which will be discussed in later sections).

The signal sample is assigned a global weight, which corresponds to the expected number of candidates in data normalised to the number of truth-matched signal candidates in the MC. The signal weight, w, is defined by:

$$w = \frac{2 \cdot \sigma_{b\bar{b}} \cdot f(b \to B^-) \cdot \mathcal{L}_{\text{int}} \cdot \mathcal{B} \cdot \varepsilon_{\text{acc}}}{N_{\text{gen}}} \ , \tag{4.1}$$

where $\sigma_{b\bar{b}}$ is the total $b\bar{b}$ production cross-section, $f(b \to D^0)$ is the fraction of b ($\bar{b}$) quarks that hadronise to a B^- (B^+), $\mathcal{L}_{\text{int}} = 35.6$ pb^{-1} is the integrated luminosity of the data samples, $\mathcal{B}$ is the branching fraction for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$, ε_{acc} is the generator efficiency and N_{gen} is the number of events passing the generator-level cuts. The factor two is a consequence of the fact that both B^+ and B^- decays are considered. The generator efficiency of 0.15360 ± 0.00021 is extracted from the Gauss MC production. The hadronisation factor $f(b \to B^-)$ is assumed to be 0.405, which is the value input into the Gauss MC generation algorithms. The $b\bar{b}$ cross-section has been measured by LHCb to be $(0.284 \pm 0.020 \pm 0.049)$ mb [33]. The branching fraction of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ is determined by multiplying the branching fractions for $B \to DK$ and $D^0 \to K^-\pi^+\pi^-\pi^+$, which correspond to $(3.68 \pm 0.33) \times 10^{-4}$ and $\left(8.09^{+0.21}_{-0.19}\right) \times 10^{-2}$ respectively according to the PDG [23].

The background sample is weighted according to the level of background expected to contaminate the D^0 invariant-mass signal region in the data. The weighting procedure is as follows:

- $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ mass combinations are classified as residing in one of the two side-band regions if the reconstructed D^0 mass is between $-98\,\text{MeV}/c^2$ and $-60\,\text{MeV}/c^2$, or between $+40\,\text{MeV}/c^2$ and $+98\,\text{MeV}/c^2$ of the nominal D^0 mass, and the reconstructed

$B^{\pm}$ mass is $\pm 500\,\mathrm{MeV}/c^2$ of the nominal $B^{\pm}$ mass;

- The signal mass window is chosen to be between $-40\,\mathrm{MeV}/c^2$ and $+30\,\mathrm{MeV}/c^2$ of the nominal D^0 mass, and $-80\,\mathrm{MeV}/c^2$ and $+70\,\mathrm{MeV}/c^2$ of the nominal $B^{\pm}$ mass;

- A two-dimensional fit to the $B^{\pm}$ and D^0 mass distributions is performed using the RooFit [73] software to $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ mass combinations in either of the side-band regions;

- The (normalised) two-dimensional PDF is integrated over the signal and side-band regions, and the ratio of the integrals is assigned as a global weight to the background sample used for the optimisation.

The D^0 mass component of the side-bands is fitted with a first-order *Chebychev polynomial* (see Eq. (A.5)) whilst the $B^{\pm}$ mass component is fitted with a second-order Chebychev polynomial. Whilst an asymmetric signal D^0 mass window has been used in the optimisation to account for the small difference in the tails seen in the MC (see Fig. 4.5) a symmetric signal mass window of $\pm 25\,\mathrm{MeV}/c^2$ is employed in the final fit.

To reject *prompt* D backgrounds in which a random kaon/pion track and a correctly reconstructed D^0 from the PV or a short-lived charm resonance are combined to form a fake B vertex, a cut of $D^0 IP\,\chi^2 > 16$ is applied prior to performing the cut optimisation.

The final choice of the PID cuts for the bachelor kaon/pion, as well as their efficiency in data are determined from a separate PID calibration study, outlined in Sec. 4.8. Additional cuts on the maximum p_{T} and momentum of the bachelor of $< 10\,\mathrm{GeV}/c$ and $< 100\,\mathrm{GeV}/c$ respectively are included in the final selection. These correspond to the fiducial cuts applied to the PID calibration samples. In addition to the optimised cut variables and the aforementioned fiducial cuts, it has been found that a charmonium veto is necessary to remove contamination from $c\bar{c}$ background. This is described in more detail in Sec. 4.3.6.

4.3.5 Cut efficiencies

Tables 4.3 and 4.4 summarise the efficiencies of the offline selection cuts for truth-matched signal MC and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ candidates from the D^0 mass side-bands in data (used as the background sample in the optimisation). The *preselected* candidates refers to the number of reconstructed candidates passing the stripping (and the D^0 $IP\,\chi^2$ cut). The *exclusive* efficiency refers to the fraction of preselected candidates that pass the cut shown in the first column of the table. The *inclusive* cut efficiency refers to the fraction of candidates passing this cut after all other offline cuts have been applied. Finally, the *total* efficiency is the fraction of candidates passing all previous cuts in the order shown in the table that also pass the cut in the first column. Without correlations, the exclusive and inclusive efficiencies should be equal. The bachelor PID cut is not shown, since as previously mentioned, its efficiency is determined from a separate study using data calibration samples. The last row in the table shows the efficiency after selecting a $B^{\pm}$ mass window of $\pm 50\,\mathrm{MeV}/c^2$, however, the $B^{\pm}$ mass is fitted to obtain the signal and background yields in the data, and so this cut is not used in the final selection. Since the reconstructed $\to$ MC particle association occasionally fails, $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ candidates with no MC-associated particle for one or more tracks are considered in the final selection efficiency calculation (see Sec. 4.3.7).

Tables 4.3 and 4.4 show that 63% of MC-associated $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ candidates that pass the stripping also pass the offline cuts, whilst less than 1% of the background in the D^0 side-bands is retained in the full $B^{\pm}$ mass range, and the retention rate of the D^0 side-band data is below 1‰ in the $B^{\pm}$ mass signal region.

4.3.6 Charmonium veto

The LHCb trigger is highly efficient at selecting $\mu^+\mu^-$ (dimuon) pairs from the decays of charmonium resonances, such as J/ψ and $\psi(2S)$. Various B decay modes which include such resonances (*e.g.* $B^{\pm} \to J/\psi\, K^{\pm}$) have almost identical decay kinematics to the signal under investigation. Without the use of PID, such modes act as a background to the signal selection. Whilst the RICH provides a limited discrimination between hadrons ($K^{\pm}$ and

Cut name	Cut efficiency Exclusive	Inclusive	Total
Preselected candidates		63069	
D^0 daughter $K^{\pm}$ $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 0$	$97.32 \pm 0.06\%$	$98.81 \pm 0.08\%$	$97.32 \pm 0.06\%$
$p_{\mathrm{T}} > 240\,\mathrm{MeV}/c$ (≥ 3 D^0 dau.)	$99.49 \pm 0.03\%$	$99.79 \pm 0.03\%$	$96.83 \pm 0.07\%$
$p_{\mathrm{T}} > 400\,\mathrm{MeV}/c$ (≥ 2 D^0 dau.)	$99.74 \pm 0.02\%$	$99.83 \pm 0.03\%$	$96.62 \pm 0.07\%$
$sIP\,\chi^2 > 16$ (≥ 3 D^0 dau.)	$94.57 \pm 0.09\%$	$97.55 \pm 0.11\%$	$91.50 \pm 0.11\%$
$sIP\,\chi^2 > 30$ (≥ 2 D^0 dau.)	$97.10 \pm 0.07\%$	$99.63 \pm 0.04\%$	$90.44 \pm 0.12\%$
$B^{\pm}$ $IP\,\chi^2 < 9$	$95.58 \pm 0.08\%$	$97.13 \pm 0.12\%$	$86.41 \pm 0.14\%$
Bachelor $K^{\pm}$ $sIP\,\chi^2 > 28$	$97.57 \pm 0.06\%$	$99.16 \pm 0.07\%$	$84.47 \pm 0.14\%$
D^0 $\max(DoCA) < 0.3\,\mathrm{mm}$	$94.92 \pm 0.09\%$	$96.41 \pm 0.13\%$	$80.77 \pm 0.16\%$
$B^{\pm}$ $DoCA < 0.1\,\mathrm{mm}$	$98.88 \pm 0.04\%$	$99.08 \pm 0.07\%$	$79.98 \pm 0.16\%$
$B^{\pm}$ $VD\,\chi^2 > 76$	$99.44 \pm 0.03\%$	$99.94 \pm 0.02\%$	$79.89 \pm 0.16\%$
D^0 $VD\,\chi^2 > 252$	$91.89 \pm 0.11\%$	$95.80 \pm 0.14\%$	$76.13 \pm 0.17\%$
$B^{\pm}$ $DIRA > 0.99995$	$95.15 \pm 0.09\%$	$94.94 \pm 0.15\%$	$74.09 \pm 0.17\%$
D^0 $DIRA > 0.992$	$99.52 \pm 0.03\%$	$99.43 \pm 0.05\%$	$73.77 \pm 0.18\%$
D^0 vertex $\chi^2/N_{\mathrm{dof}} < 6$	$97.15 \pm 0.07\%$	$98.49 \pm 0.09\%$	$72.32 \pm 0.18\%$
J/ψ veto	$100.00 \pm 0.00\%$	$100.00 \pm 0.00\%$	$72.32 \pm 0.18\%$
$\psi(2S)$ veto	$100.00 \pm 0.00\%$	$100.00 \pm 0.00\%$	$72.32 \pm 0.18\%$
Bachelor $K^{\pm}$ $p_{\mathrm{T}} < 10\,\mathrm{GeV}/c$	$97.55 \pm 0.06\%$	$98.75 \pm 0.06\%$	$70.39 \pm 0.18\%$
Bachelor $K^{\pm}$ $\lvert\vec{p}\rvert < 100\,\mathrm{GeV}/c$	$90.50 \pm 0.12\%$	$92.08 \pm 0.13\%$	$64.64 \pm 0.19\%$
D^0 mass window $= \pm 25\,\mathrm{MeV}/c^2$	$97.68 \pm 0.06\%$	$99.15 \pm 0.05\%$	$63.44 \pm 0.19\%$
$B^{\pm}$ mass window $= \pm 50\,\mathrm{MeV}/c^2$	$97.07 \pm 0.07\%$	$98.72 \pm 0.06\%$	$62.62 \pm 0.19\%$

Table 4.3: $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ cut efficiencies for combined magnet up and magnet down, truth-matched MC signal candidates.

Cut name	Cut efficiency: Exclusive	Inclusive	Total
Preselected candidates		752012	
D^0 daughter $K^\pm \log(\mathcal{L}_K/\mathcal{L}_\pi) > 0$	$40.83 \pm 0.06\%$	$38.75 \pm 1.72\%$	$40.83 \pm 0.06\%$
$p_\mathrm{T} > 240\,\mathrm{MeV}/c$ (≥ 3 D^0 dau.)	$91.17 \pm 0.03\%$	$99.04 \pm 0.55\%$	$37.10 \pm 0.06\%$
$p_\mathrm{T} > 400\,\mathrm{MeV}/c$ (≥ 2 D^0 dau.)	$91.86 \pm 0.03\%$	$99.36 \pm 0.45\%$	$34.85 \pm 0.05\%$
$sIP\,\chi^2 > 16$ (≥ 3 D^0 dau.)	$34.88 \pm 0.05\%$	$78.28 \pm 2.07\%$	$12.89 \pm 0.04\%$
$sIP\,\chi^2 > 30$ (≥ 2 D^0 dau.)	$49.07 \pm 0.06\%$	$98.41 \pm 0.70\%$	$10.93 \pm 0.04\%$
$B^\pm$ $IP\,\chi^2 < 9$	$58.35 \pm 0.06\%$	$77.11 \pm 2.10\%$	$6.01 \pm 0.03\%$
Bachelor $K^\pm$ $sIP\,\chi^2 > 28$	$80.67 \pm 0.05\%$	$91.45 \pm 1.52\%$	$4.84 \pm 0.02\%$
D^0 max$\,(DoCA) < 0.3\,\mathrm{mm}$	$48.52 \pm 0.06\%$	$72.60 \pm 2.16\%$	$2.89 \pm 0.02\%$
$B^\pm$ $DoCA < 0.1\,\mathrm{mm}$	$87.22 \pm 0.04\%$	$92.26 \pm 1.46\%$	$2.64 \pm 0.02\%$
$B^\pm$ $VD\,\chi^2 > 76$	$88.96 \pm 0.04\%$	$98.73 \pm 0.63\%$	$2.59 \pm 0.02\%$
D^0 $VD\,\chi^2 > 252$	$31.59 \pm 0.05\%$	$67.83 \pm 2.19\%$	$1.62 \pm 0.01\%$
$B^\pm$ $DIRA > 0.99995$	$49.91 \pm 0.06\%$	$67.39 \pm 2.19\%$	$1.31 \pm 0.01\%$
D^0 $DIRA > 0.992$	$88.42 \pm 0.04\%$	$96.88 \pm 0.97\%$	$1.28 \pm 0.01\%$
D^0 vertex $\chi^2/N_\mathrm{dof} < 6$	$57.60 \pm 0.06\%$	$78.28 \pm 2.07\%$	$1.02 \pm 0.01\%$
J/ψ veto	$99.60 \pm 0.01\%$	$98.73 \pm 0.63\%$	$1.01 \pm 0.01\%$
$\psi(2S)$ veto	$99.96 \pm 0.00\%$	$100.00 \pm 0.00\%$	$1.01 \pm 0.01\%$
Bachelor $K^\pm$ $p_\mathrm{T} < 10\,\mathrm{GeV}/c$	$99.20 \pm 0.01\%$	$99.05 \pm 0.39\%$	$0.98 \pm 0.01\%$
Bachelor $K^\pm$ $\lvert\vec{p}\rvert < 100\,\mathrm{GeV}/c$	$94.57 \pm 0.03\%$	$91.79 \pm 1.05\%$	$0.90 \pm 0.01\%$
$B^\pm$ mass window $= \pm 50\,\mathrm{MeV}/c^2$	$9.59 \pm 0.03\%$	$9.29 \pm 0.35\%$	$0.08 \pm 0.00\%$

Table 4.4: Cut efficiencies for combined magnet up and magnet down data side-bands from *StrippingB2DXWithD2hhhhLine* reconstructed as $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$.

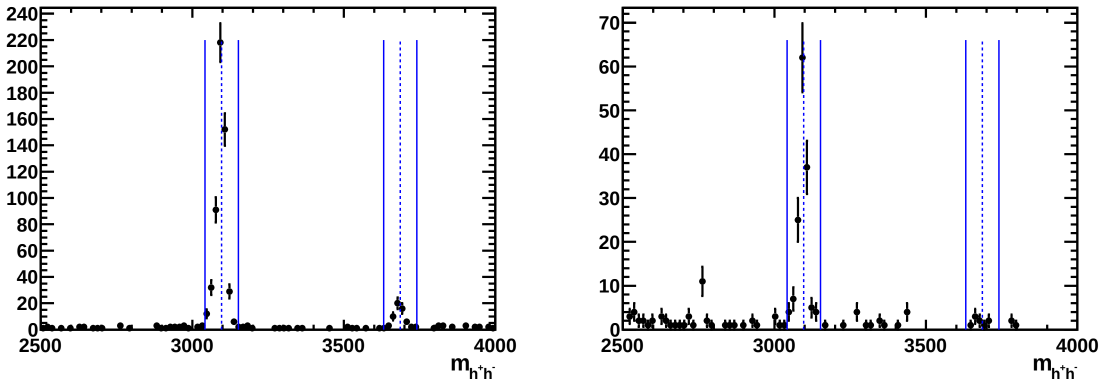
$\pi^{\pm}$) and muons, a veto is employed on oppositely charged track pairs which pass a loose muon identification requirement, whilst also being consistent with the mass of either the J/ψ or $\psi(2S)$. The four-momentum of the two tracks are re-calculated assuming a muon mass hypothesis, and the invariant mass of the dimuon association is calculated.

Figure 4.2 shows the invariant mass of dimuon associations for $B^{\pm} \to (K^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ offline-selected data. As can be seen, the level of this background is around four times larger for $B^{\pm} \to (K^{\pm}\pi^{\mp})_D K^{\pm}$ candidates compared to $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$. This is because these backgrounds involve the additional reconstruction of two pion tracks which, along with the mis-identified dimuons from the J/ψ or $\psi(2S)$ decays, form a fake D vertex. The extra pions can either come from two random pions in the event, or from the decays $B^{\pm} \to J/\psi K^{\pm}\pi^{\mp}\pi^{\pm}$ or $B^{\pm} \to \psi(2S)K^{\pm}\pi^{\mp}\pi^{\pm}$, where the pions are missed in the $B^{\pm}$ vertex reconstruction, and are instead reconstructed along with the dimuons. Whilst these backgrounds could be minimised or removed by tightening the vertex quality cuts, this would result in a loss of signal candidates and, as explained below, a veto on these backgrounds is a more efficient solution. Based on the aforementioned dimuon mass distributions, $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ candidates are rejected from the final selection if one or more of the reconstructed dimuon mass combinations lies within 55 MeV/c^2 of the nominal mass of the J/ψ or $\psi(2S)$.

The charmonium vetoes have been applied to a sample of approximately 1.5 million $B^{\pm} \to J/\psi K^{\pm}$ MC decays reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp})_D K^{\pm}$, using the offline selection shown in Tab. 4.5. Of the $B^{\pm} \to J/\psi K^{\pm}$ decays passing all other offline cuts, 99.99% are rejected by the vetoes. By comparison, Table 4.3 indicates that these cuts lose no $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ truth-matched MC candidates.

4.3.7 The results of the offline selection

The D^0 mass distribution for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ candidates from the full data and MC datasets after applying the offline selection (summarised in Tab. 4.5) are shown in Figs. 4.4 and 4.5 respectively. In addition to these cuts, a $B^{\pm}$ mass window of 50 MeV/c^2 is



(a) (b)

Figure 4.2: Invariant mass of h^+h^- muon-like pairs for (a) $B^\pm \to (K^\pm\pi^\mp)_D\,\pi^\pm$ and (b) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ offline-selected candidates. All offline cuts are applied except the D^0 mass window and bachelor $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts. The dashed lines represent the nominal J/ψ and $\psi(2S)$ masses, whilst the solid lines represent $\pm 55\,\mathrm{MeV}/c^2$ from these nominal masses.

applied. The D^0 mass peak is fitted with a *Cruijff* PDF, as defined in Eq. (A.2). For the data, the combinatoric background is fitted with a first order Chebychev polynomial. Below the mass plots, the pull distributions[1] have been determined by binning the distribution and the PDF, where the expected value for each pull is the integral of the PDF evaluated over the bin range. In the case of the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ candidates from data, the sum of the signal and combinatoric background PDFs (the latter of which is modelled by a first order Chebychev polynomial) is used to determine the pulls. Table 4.6 shows the D^0 mass fit parameters for the offline-selected data and MC. The mean D^0 mass is lower in data than MC by $\sim 2\,\mathrm{MeV}/c^2$, whilst the width (σ) is wider by approximately $1\,\mathrm{MeV}/c^2$. The mean and width are floated in the final fit to the data, so these data-MC differences are not expected to impact on the results.

Table 4.7 shows the efficiencies of selecting $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ MC candidates. $\varepsilon_{\mathrm{strip|acc}}$ is defined as the number of reconstructed candidates in the MC datasets passing the stripping divided by the total number of candidates

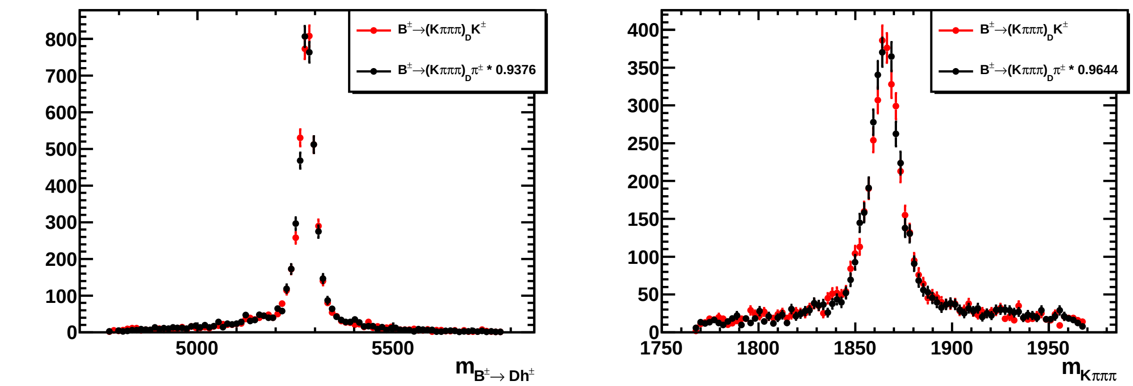
[1] The pull for a data point x with statistical uncertainty σ is defined by $\dfrac{x-\bar{x}}{\sigma}$, where $\bar{x}$ is the expectation of x, usually determined from a fitted function or theoretical model. If the fit function or theoretical model perfectly describes the data, then the pulls will be distributed according to a *standard Gaussian* (a Gaussian with zero mean and a width of unity).

passing the generator-level cuts. This incorporates both the efficiencies of the stripping cuts, and the efficiency of reconstructing the candidate. $\varepsilon_{\mathrm{sel|strip}}$ is defined as the number of candidates in the MC passing the offline selection (excluding the bachelor PID cut) divided by the number of candidates passing the stripping cuts. Finally, $\varepsilon_{\mathrm{sel|acc}}$, hereby referred to as the *selection efficiency*, is the number of candidates passing the offline selection divided by the total number of MC candidates passing the generator-level cuts.

As previously discussed, both MC-associated candidates and candidates in which one or more tracks have no associated MC particle are included when determining the efficiencies. This procedure is not perfect, since some of the non-associated candidates could be combinatoric background rather than signal, and so the efficiencies could be overestimated. However, this analysis considers only the ratio of efficiencies for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$, and this is only problematic if the distributions of non-associated signal and background differ between the two decay modes. Figure 4.3 shows the $B^{\pm}$ and D^0 mass distributions for non-associated $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ signal MC, in which the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ distribution is normalised such that the integral is the same as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$. The results show good agreement in both the $B^{\pm}$ and D^0 mass distributions, indicating that considering non-associated MC candidates when determining the selection efficiency is a reasonable approach.

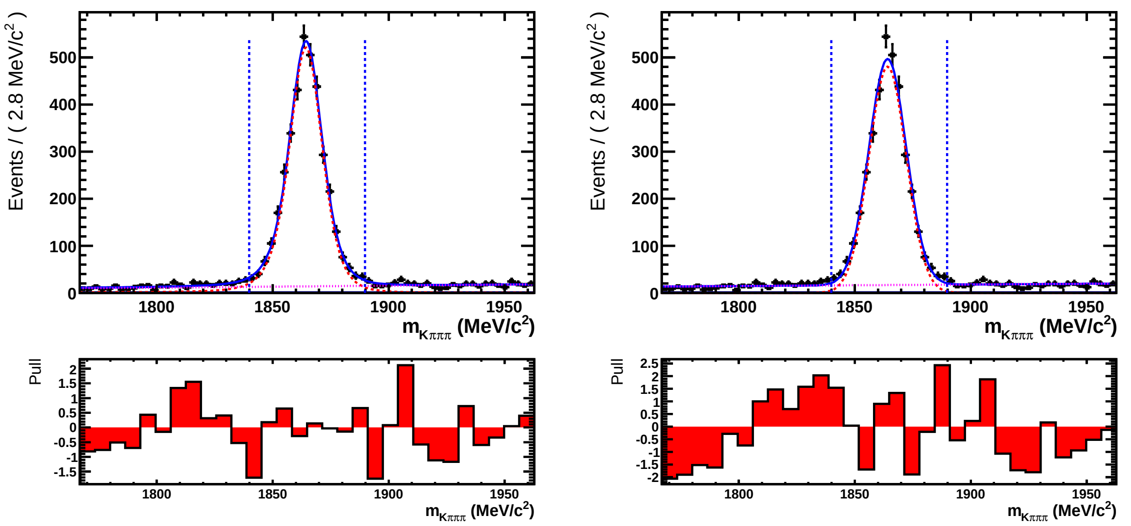
4.4 Comparison of offline cuts for data and MC

As discussed in Sec. 4.3, the optimisation of the selection cuts uses MC signal samples to represent the distributions of the cut variables. In this section, the distributions of each of the variables used in the offline selection from the MC and data samples are compared. Figures 4.6 to 4.9 show the distributions for signal MC and "raw" data (*i.e.* without utilising any background-subtraction techniques). In each case, the offline cuts outlined in Tab. 4.3 are applied, except for the variable being considered. The charmonium vetoes and cuts related to the PID performance (maximum bachelor p_{T} and momentum cuts) are also not applied. A $B^{\pm}$ mass window of $\pm 50\ \mathrm{MeV}/c^2$ has been imposed and, for the data, a bachelor PID cut



(a) (b)

Figure 4.3: $B^{\pm}$ and D^0 mass distributions for offline-selected $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ (red) and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ (black) MC candidates with failures in the MC truth matching. (a) shows the $B^{\pm}$ mass distribution with a D^0 mass window of $\pm 25\,\mathrm{MeV}/c^2$, whilst (b) shows the D^0 mass distribution with a $B^{\pm}$ mass window of $\pm 50\,\mathrm{MeV}/c^2$. In both cases, the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ distribution has been normalised such that the integral is the same as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$.



(a) (b)

Figure 4.4: D^0 mass distribution for offline-selected $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ data. The D^0 mass peak is fitted with a Cruijff PDF in (a) and a Gaussian PDF in (b). The dashed vertical lines indicate the $\pm 25\,\mathrm{MeV}/c^2$ signal mass-window. The pull values in uniform bins are shown below the mass plots.

Cut variable	Offline cuts $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, h^\pm$	$B^\pm \to (K^\pm\pi^\mp)_D\, h^\pm$
D^0 mass window (MeV/c^2)	± 25	± 25
D^0 daughter $K^\pm$ $\log(\mathcal{L}_K/\mathcal{L}_\pi)$	> 0	> 0
D^0 daughter p_T (MeV/c)	$> 240^*, > 400^\dagger$	> 330
Bachelor p_T (GeV/c)	< 10	< 10
Bachelor $\lvert\vec{p}\rvert$ (GeV/c)	< 100	< 100
$B^\pm$ *IP* χ^2	< 9	< 9
D^0 *IP* χ^2	< 16	< 16
D^0 daughter *sIP* χ^2	$> 16^*, > 30^\dagger$	> 21
Bachelor *sIP* χ^2	> 28	> 28
$B^\pm$ *VD* χ^2	> 76	> 76
D^0 *VD* χ^2	> 252	> 252
$B^\pm$ *DoCA* (mm)	< 0.1	< 0.1
D^0 max (*DoCA*) (mm)	< 0.3	< 0.3
$B^\pm$ *DIRA*	> 0.99995	> 0.99995
D^0 *DIRA*	> 0.992	> 0.992
D^0 vertex χ^2/N_dof	< 6	< 6
J/ψ mass veto (MeV/c^2)	± 55	± 55
$\psi(2S)$ mass veto (MeV/c^2)	± 55	± 55

Table 4.5: Summary of the offline selection cuts, excluding the bachelor $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut. These cuts are applied to candidates selected by the appropriate stripping lines.

* At least three D^0 daughters must satisfy this criterion.
† At least two D^0 daughters must satisfy this criterion.

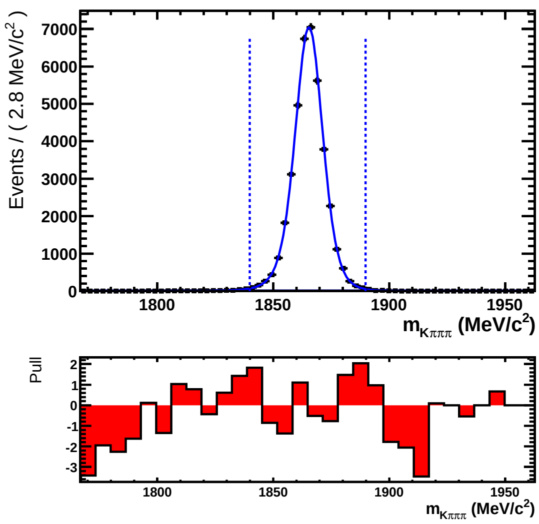


Figure 4.5: D^0 mass distribution for offline-selected $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ signal MC. The distribution is fitted with a Cruijff PDF. The dashed vertical lines indicate the $\pm 25\,\mathrm{MeV}/c^2$ signal mass-window. The pull values in uniform bins are shown below the mass plot.

Fit variable	Data Cruijff fit	MC Cruijff fit	Data Gaussian fit
Mean (MeV/c^2)	1864.5 ± 0.2	1865.41 ± 0.04	1864.2 ± 0.2
Sigma (MeV/c^2)	6.292 ± 0.007	5.73 ± 0.03	8.0 ± 0.1
α_{L} (MeV/c^2)	0.21 ± 0.02	0.114 ± 0.002	
α_{R} (MeV/c^2)	0.17 ± 0.01	0.094 ± 0.002	
n_{sig}	3569 ± 112		3446 ± 62
n_{bkg}	182 ± 28		297 ± 10
S/B	20 ± 3		12.6 ± 0.4

Table 4.6: Cruijff fit parameters from a fit to the D^0 mass distribution of $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ candidates in data and MC, with Gaussian fit parameters as a comparison for the data. The quantities n_{sig} and n_{bkg} correspond to the signal and background yields from the fit (data only) determined by fitting the PDFs in the full mass range, then integrating in the $\pm 25\,\mathrm{MeV}/c^2$ mass window. The signal-background ratio (S/B) is also indicated for the data.

Efficiencies	$B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$	$B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$
$\varepsilon_{\mathrm{strip}\vert\mathrm{acc}}$	0.0510 ± 0.0002	0.0520 ± 0.0002
$\varepsilon_{\mathrm{sel}\vert\mathrm{strip}}$	0.530 ± 0.003	0.527 ± 0.003
$\varepsilon_{\mathrm{sel}\vert\mathrm{acc}}$	0.0270 ± 0.001	0.0274 ± 0.0001

Table 4.7: Selection efficiencies for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ signal MC.

of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 4$ is applied to decays reconstructed under the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ hypothesis to reduce the $B \to D\pi$ contamination, and a bachelor cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 0$ is applied to $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ candidates to reduce the mis-identified $B \to DK$ background.

Figures 4.6 to 4.9 demonstrate that in most cases, the MC and data agree quite well. However, there are a few exceptions, particularly the IP χ^2 and vertex χ^2 variables (Figs. 4.6c, 4.6d, 4.7c and 4.7d). However, the differences primarily occur at very tight cut values, whilst the selection cuts are much looser, so the loss of signal efficiency from this cut is anyway small.

4.4.1 Validation of D^0 daughter $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts

It is known that the RICH $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distributions used to distinguish between pions and kaons are poorly described in the MC. The final choice of the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut of the bachelor pion/kaon and the efficiency of the PID cut are determined from data. The procedure for determining the efficiency of the PID cuts is outlined in Sec. 4.8.

Since the D final state is the same for the two bachelor particle hypotheses ($B \to DK$ and $B \to D\pi$) any inaccuracy in the efficiency of the D-daughter track $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts should cancel in the ratio (assuming they are uncorrelated with the kinematics of the bachelor). Cuts on the D-daughter pion $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ variables were included at an early stage in the analysis, but have been removed in the final selection, since they do not provide a strong discrimination between signal and background. However, the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut on the kaon remains, and the following study investigates whether the choice of this cut from the optimisation procedure is a valid one, in that the signal significance ($S/\sqrt{S+B}$) is optimal according to the data. Data comprising the output of the stripping reconstructed under the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ hypothesis are considered, since the branching fraction is much larger for this channel than $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$. All offline cuts have been applied apart from the D^0 daughter PID cuts, and a $B^\pm$ mass window of ± 50 MeV/c^2 and a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 0$ are applied to reduce the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ contamination. The data are fitted with

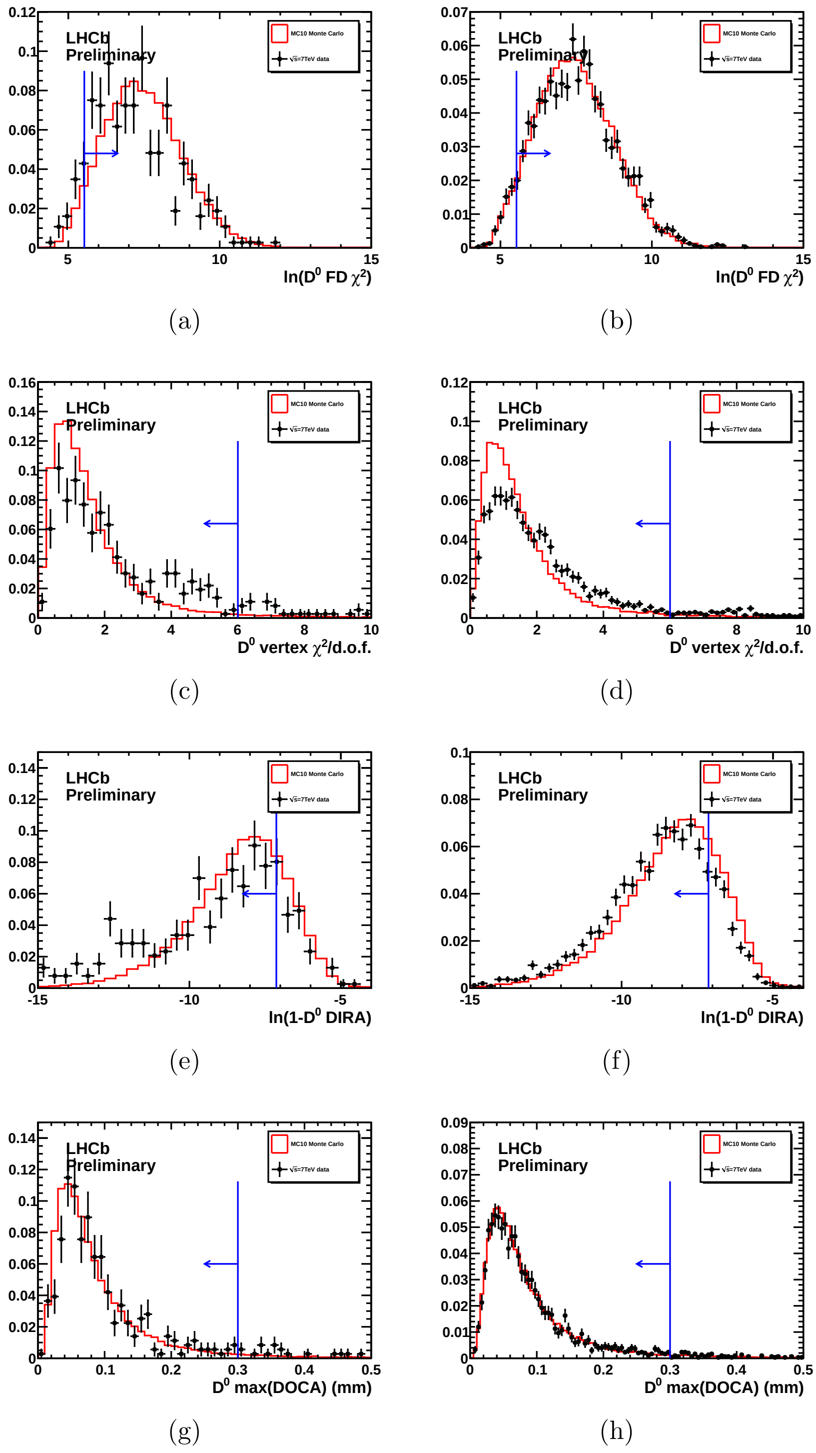


Figure 4.6: Comparison of the cut-variable distributions in data (black) and MC (red) for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ (left plots) and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ (right plots) decays. The variables are the D^0 (a)(b) $VD\,\chi^2$, (c)(d) vertex $\chi^2/N_{\rm dof}$, (e)(f) $DIRA$ and (g)(h) $\max(DoCA)$.

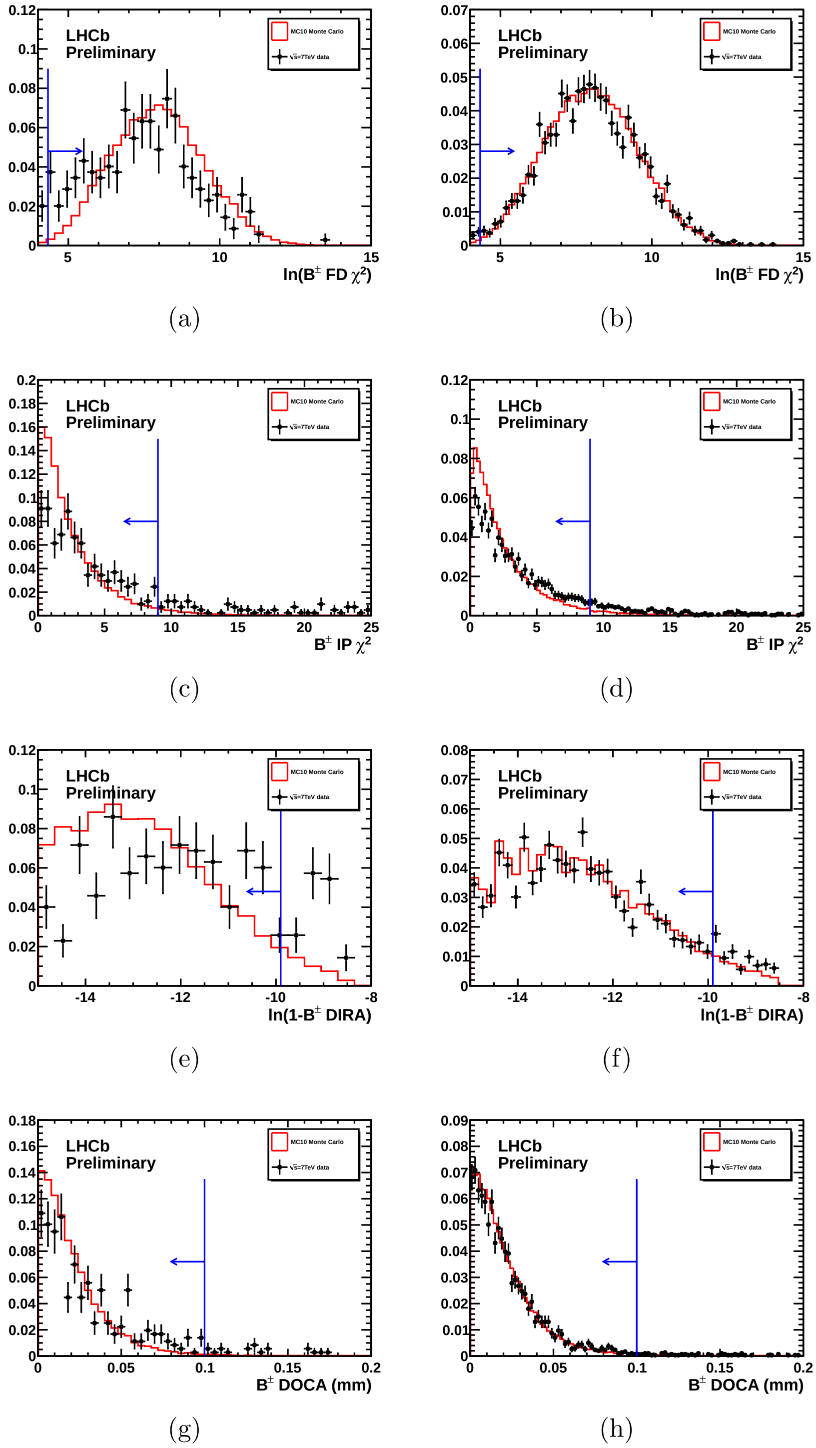


Figure 4.7: Comparison of the cut-variable distributions in data (black) and MC (red) for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, K^\pm$ (left plots) and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, \pi^\pm$ (right plots) decays. The variables are the $B^\pm$ (a)(b) *VD* χ^2, (c)(d) *IP* χ^2, (e)(f) *DIRA* and (g)(h) *DoCA*.

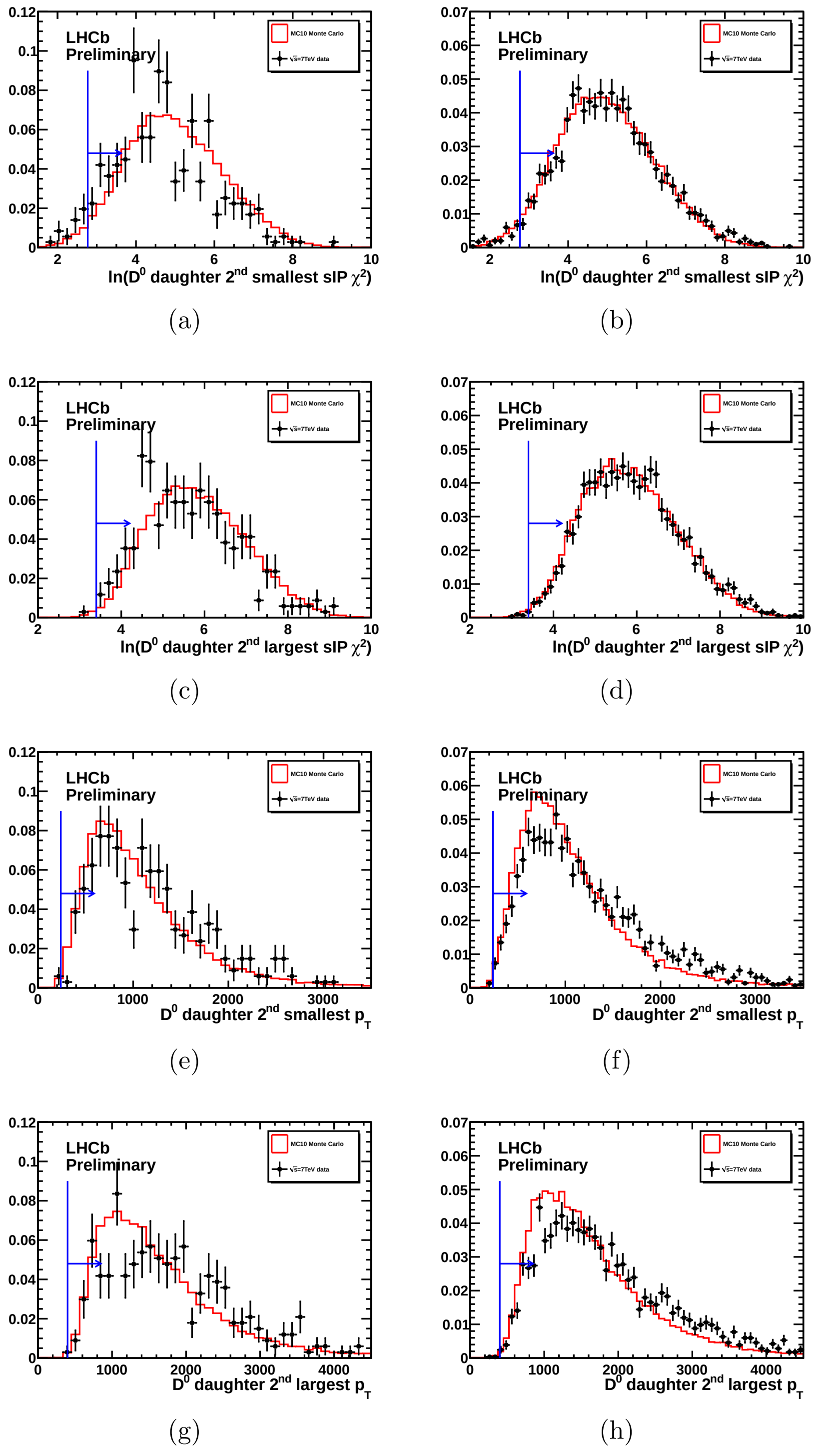
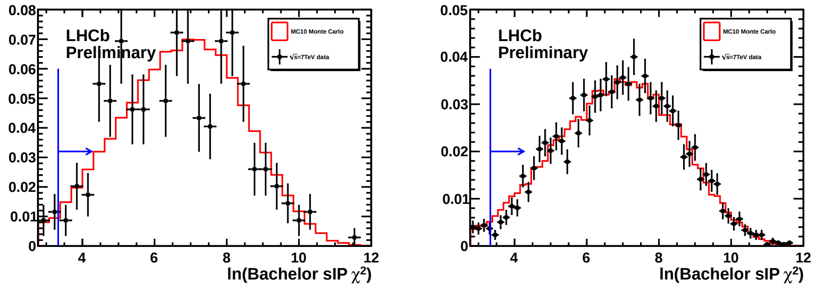


Figure 4.8: Comparison of the data (black) and MC (red) for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, K^\pm$ (left plots) and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, \pi^\pm$ (right plots) decays. The variables are the D^0 daughter (a)–(d) $sIP\,\chi^2$ and (e)–(h) p_{T}.

a straight line for the combinatoric background PDF and a double Crystal Ball with fixed powers for the signal shape. A kaon $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut has then been applied and the data fitted again. The PID cut is varied using a step size of two over the range $-16 < \log(\mathcal{L}_K/\mathcal{L}_\pi) < 16$. In each case, $S/\sqrt{S+B}$ is determined from the signal and background yields in the D^0 mass range ($[-40\,\mathrm{MeV}/c^2, 30\,\mathrm{MeV}/c^2]$) before and after applying the PID cut. The results are show in Fig. 4.10. They indicate that the cut chosen by the optimisation procedure ($\log(\mathcal{L}_K/\mathcal{L}_\pi) > 0$) is very close to optimum according to the data, and so the cut has not been altered in the final offline selection.



(a) (b)

Figure 4.9: Comparison of the cut-variable distributions for (a) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ and (b) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ decays in data (black) and MC (red). Both plots show the $sIP\,\chi^2$ of the bachelor track.

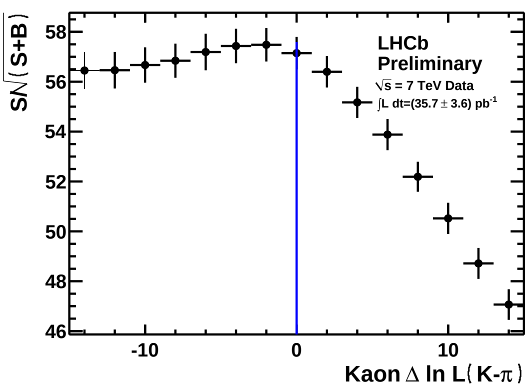


Figure 4.10: Validation study of the D^0 daughter K PID efficiency. The D^0 mass for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ is fitted before and after applying all offline cuts apart from the D^0 daughter K PID cut. In addition, a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 0$ has been applied. The plot shows $S/\sqrt{S+B}$ determined in the D^0 mass signal region for various D^0 daughter kaon $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts. The "optimal" cut is indicated by the vertical line.

4.5 Performance of the trigger lines

As explained in Chap. 2, the LHCb triggering system consists of a hardware level known as L0 followed by two software levels known as HLT1 and HLT2. HLT1 comprises generic trigger lines, whilst HLT2 trigger lines are more focused on individual physics interests. For the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement, it is important to ascertain whether there is any difference in the ratio of the trigger efficiencies (with respect to candidates passing the offline selection) between $B \to D\pi$ and $B \to DK$. If so, then this must be taken into account in the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement.

When investigating the trigger efficiency, we must review what is causing the event to be triggered by the trigger lines. For each trigger line considered, an event can be classified as:

- **Triggered on signal (TOS)**, which indicates that the decay products of the offline-reconstructed signal candidate were sufficient to cause the trigger line to select the event;
- **Triggered independently of signal (TIS)**, which means that information from the rest of the event (*i.e.* excluding the decay products of the signal candidate) was sufficient to cause the trigger line to select the event;
- **Triggered on both (TOB)**, which indicates that *both* the decay products of the signal candidate and information from the rest of the event were necessary for the trigger line to select the event.

The above classifications are determined by comparing the decay products of the offline-reconstructed signal candidate to the tracks, track segments or calorimeter hits used by the trigger line to decide whether to accept or reject the event.

Considering candidates which pass all trigger lines when performing the final fit can, in principle, lead to biases. In particular, including candidates which are TOS with respect to trigger lines other than those designed to select the decays of interest can distort the distributions of the kinematic variables (assuming that they are not also TOS with respect to the trigger lines of interest). This is because different trigger lines have different cuts applied to quantities such as the invariant mass, momentum, vertex and track-fit quality. In most cases, this should not be an issue for this analysis, since we deal with the ratio of two kinematically similar decay modes. However, if these distortions are not exactly the same for $B \to DK$ and $B \to D\pi$, or for B^+ and B^- decays, then this could lead to biases in the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$ results. Furthermore, for the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement, kinematic distortions could also affect the PID performance (see Sec. 4.8) meaning the kaon and pion PID efficiencies and mis-ID rates could depend on which trigger line is passed. If a large fraction of triggered events are TOB, then this can also lead to biases, since a combination of tracks from the signal candidate(s) and other tracks from the event are necessary to trigger the event. Again, if these issues affect $B \to DK/B \to D\pi$, or B^+/B^- decays differently, then this may bias the measurements of the observables of interest.

The specific trigger-line selection cuts and prescales (the fraction of events processed) used by the trigger lines are defined by the *Trigger Configuration Key* (TCK). The data were taken with several different TCKs, many of which differ only by the prescales. The majority of events selected were taken with a single TCK[1]. In this case, the minimum E_{T} threshold of the L0 hadron trigger (see Sec. 2.8.1) was set to $3.6\,\mathrm{GeV}/c^2$ and the maximum SPD multiplicity to 900.

To evaluate the trigger efficiency, a subset of the MC signal samples (3,000 magnet up and 3,000 magnet down events from each of the $B \to DK$ and $B \to D\pi$ samples) has been processed using the LHCb HLT software known as MOORE[2] using the same software versions and TCKs used in the data. The efficiencies quoted in this section are for the TCK that the majority of data were taken with; however, the conclusions found are broadly independent of the TCK considered [3].

Since the final states contain only hadrons, the majority of events should be TOS from trigger lines designed to select fully hadronic states, namely the L0 hadron, HLT1 1track, and

[1] TCK `0x002e002a`.
[2] The MOORE software also simulates the L0 hardware trigger.

HLT2 topological triggers. Table 4.8, shows the proportion of offline-selected truth-matched $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \, \pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \, K^{\pm}$ MC decays passing these trigger lines (third column). At each subsequent stage of the trigger, the events (in the numerator of the efficiency calculation) are required to have been triggered by the previous trigger stages. In addition to the hadronic lines, the trigger efficiencies of the logical OR of all lines designed to select *physics* events (*i.e.* not including lines used for debugging and monitoring purposes) at each trigger stage are given. The fourth, fifth and sixth columns of Tab. 4.8 show the proportion of offline-selected events that are TOS, TIS and TOB respectively with regards to the specified trigger line(s).

Trigger	Decay type	$\varepsilon_{\text{trig}\vert\text{sel}}$	$\varepsilon^{\text{TOS}}_{\text{trig}\vert\text{sel}}$	$\varepsilon^{\text{TIS}}_{\text{trig}\vert\text{sel}}$	$\varepsilon^{\text{TOB}}_{\text{trig}\vert\text{sel}}$
L0 physics	$B^{\pm} \to D\pi^{\pm}$	46.7 ± 0.6	30.2 ± 0.6	16.5 ± 0.5	0.0 ± 0.0
	$B^{\pm} \to DK^{\pm}$	47.6 ± 0.6	31.1 ± 0.6	16.5 ± 0.5	0.0 ± 0.0
L0 hadron	$B^{\pm} \to D\pi^{\pm}$	39.6 ± 0.6	29.7 ± 0.6	9.9 ± 0.4	0.0 ± 0.0
	$B^{\pm} \to DK^{\pm}$	40.4 ± 0.6	30.5 ± 0.6	9.9 ± 0.4	0.0 ± 0.0
HLT1 physics	$B^{\pm} \to D\pi^{\pm}$	35.1 ± 0.6	33.5 ± 0.6	1.4 ± 0.1	0.2 ± 0.1
	$B^{\pm} \to DK^{\pm}$	36.5 ± 0.6	34.4 ± 0.6	1.7 ± 0.1	0.5 ± 0.1
HLT1 1track	$B^{\pm} \to D\pi^{\pm}$	34.9 ± 0.6	33.6 ± 0.6	1.1 ± 0.1	0.2 ± 0.1
	$B^{\pm} \to DK^{\pm}$	35.9 ± 0.6	34.3 ± 0.6	1.3 ± 0.1	0.3 ± 0.1
HLT2 physics	$B^{\pm} \to D\pi^{\pm}$	30.5 ± 0.6	29.2 ± 0.6	0.5 ± 0.1	0.7 ± 0.1
	$B^{\pm} \to DK^{\pm}$	31.1 ± 0.6	29.9 ± 0.6	0.5 ± 0.1	0.7 ± 0.1
HLT2 topo.	$B^{\pm} \to D\pi^{\pm}$	29.2 ± 0.6	28.5 ± 0.6	0.2 ± 0.1	0.5 ± 0.1
	$B^{\pm} \to DK^{\pm}$	29.8 ± 0.6	29.1 ± 0.6	0.2 ± 0.1	0.5 ± 0.1

Table 4.8: Trigger efficiencies (in %) at different stages of the trigger for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \, K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \, \pi^{\pm}$ offline-selected MC.

Table 4.8 indicates that just under half of all MC events pass the hardware trigger stage, of which around 80% are triggered by the hadron trigger. The TOS efficiency at the L0 level is approximately independent of whether the event was triggered by the L0 hadron trigger or not, and none of the L0-triggered events are found to be TOB. These last two observations imply that the majority of signal events are either TIS, or TOS with respect to the L0 hadron trigger. Thus, no significant trigger bias is expected at the L0 level.

At the two HLT stages, a large proportion of events (around 93-94%) are TOS with respect to the fully-hadronic trigger lines. In the HLT2 trigger in particular, many of the remaining events are TOB. Furthermore, there is a statistically significant difference between the TOB fractions in the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ samples; however, such a difference is not seen at the HLT2 level.

The findings above indicate that there is potentially a bias from considering all trigger decisions. However, using more stringent requirements would result in a significant loss of events, so all trigger decisions are used in the final fit, and a systematic uncertainty is assigned to account for any possible trigger bias. The method for determining the systematic uncertainties due to the trigger requirements is explained in Sec. 4.10.9.

Table 4.9 shows the trigger efficiencies for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ MC samples split by the charge of the B hadron. The results indicate a systematically higher efficiency for triggering on $B^+ \to DK^+$ signal decays compared to $B^- \to DK^-$, although the asymmetry is never larger than 2σ at any trigger stage. Since the same systematic asymmetry is not seen for events classified as TIS and TOB, is possible that K^+ and K^- or π^+ and π^- decays have different interaction lengths, resulting in more detector hits that can be triggered on for B^+ decays compared to B^-. As discussed in Sec. 4.9.3, the pion detection asymmetry is assumed to be zero in the $A^{K3\pi}_{fav}$ measurement, whilst the kaon detection asymmetry is corrected for in the final fit.

4.6 Background study

Even with an optimal selection, there will always be some candidates other than the decays of interest passing all selection cuts. A good understanding of these backgrounds is essential in order to accurately measure the signal yield. Some background types can be eliminated by cuts (such as the charmonia vetoes discussed in Sec. 4.3.6) whilst others must be accounted for either within the final fit or through the analysis of specific background samples.

Figure 4.11 shows the $B^{\pm}$ mass distribution of offline-selected $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ candidates in data. As well as a clear peak centred around the nominal B mass, the following

Trigger	Decay type	$\varepsilon_{\text{trig}\vert\text{sel}}$	$\varepsilon^{\text{TOS}}_{\text{trig}\vert\text{sel}}$	$\varepsilon^{\text{TIS}}_{\text{trig}\vert\text{sel}}$	$\varepsilon^{\text{TOB}}_{\text{trig}\vert\text{sel}}$
L0 physics	$B^+ \to DK^+$	48.3 ± 0.9	31.9 ± 0.8	16.4 ± 0.6	0.0 ± 0.0
	$B^- \to DK^-$	46.9 ± 0.9	30.5 ± 0.8	16.5 ± 0.7	0.0 ± 0.0
L0 hadron	$B^+ \to DK^+$	41.3 ± 0.9	31.4 ± 0.8	9.9 ± 0.5	0.0 ± 0.0
	$B^- \to DK^-$	39.6 ± 0.9	29.7 ± 0.8	9.9 ± 0.5	0.0 ± 0.0
HLT1 physics	$B^+ \to DK^+$	36.9 ± 0.9	34.8 ± 0.8	1.6 ± 0.2	0.5 ± 0.1
	$B^- \to DK^-$	36.3 ± 0.9	33.9 ± 0.8	1.8 ± 0.2	0.5 ± 0.1
HLT1 1track	$B^+ \to DK^+$	36.4 ± 0.9	34.8 ± 0.8	1.4 ± 0.2	0.3 ± 0.1
	$B^- \to DK^-$	35.4 ± 0.9	33.7 ± 0.8	1.3 ± 0.2	0.3 ± 0.1
HLT2 physics	$B^+ \to DK^+$	31.4 ± 0.8	30.1 ± 0.8	0.6 ± 0.1	0.7 ± 0.1
	$B^- \to DK^-$	30.8 ± 0.8	29.7 ± 0.8	0.4 ± 0.1	0.8 ± 0.2
HLT2 topo.	$B^+ \to DK^+$	30.1 ± 0.8	29.4 ± 0.8	0.2 ± 0.1	0.4 ± 0.1
	$B^- \to DK^-$	29.4 ± 0.8	28.8 ± 0.8	0.1 ± 0.1	0.5 ± 0.1

Table 4.9: Trigger efficiencies (in %) at different stages of the trigger for $B^+ \to (K^+\pi^-\pi^+\pi^-)_D\, K^+$ and $B^- \to (K^-\pi^+\pi^-\pi^+)_D\, K^-$ offline-selected MC.

background contributions can be seen:

- combinatoric;
- low mass;
- semi-leptonic.

In the cut optimisation, the $B^\pm$ invariant-mass distribution of the combinatoric background is modelled by a second-order Chebychev polynomial. However, since the level of this background type has been significantly reduced by the offline selection cuts, the remaining candidates are roughly linearly distributed in B mass.

The low-mass backgrounds correspond to fully-hadronic decay modes which have a similar topology to the signal, but with one or more particles in the decay not accounted for in the reconstruction. They are characterised by a wide peak at a lower mass than the nominal $B^\pm$ mass. Many of these low-mass backgrounds involve D^* resonances, where the D^0 (or $\overline{D}^0$) decays to the correct final state. These are characterised by the double-peak structure seen in Fig. 4.11, which is a result of the $\cos^2\theta$ dependence of the decay amplitude of the D^*

resonance. Here, θ is the angle between the momentum vector of the $B^{\pm}$ bachelor track and the either of the D^{*} daughters in the D^{*} rest frame.

The semi-leptonic backgrounds comprise a combination of leptonic and hadronic tracks. The nature of these decays is such that they have missing energy from one or more neutrinos and are reconstructed with at least one mis-identified lepton. Consequently, the majority of these events are reconstructed with an invariant mass much lower than the nominal B mass, though the tail of the B mass distribution may extend into the signal mass window.

Another source of background not considered here comprises candidates in which a $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ decay is incorrectly reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ and vice-versa (mis-ID background). This background is particularly problematic for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ due to the significantly larger branching fraction of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$. The contamination from these background decays can be minimised by applying a PID cut on the bachelor track. The study of the efficiency and mis-ID rate for various values of the PID cut is detailed in Sec. 4.8. The remaining mis-ID background is accounted for in the final fit, as discussed in Sec. 4.7.

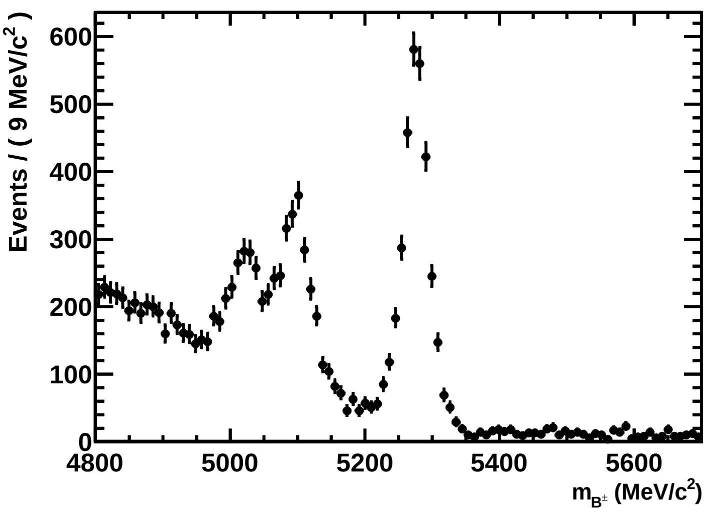


Figure 4.11: $B^{\pm}$ mass distribution of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ candidates in LHCb pp collision data.

The remainder of this section describes the three dominant types of background in more detail. This includes investigations of the principle inclusive and exclusive physics backgrounds (*i.e.* low-mass and semileptonic) expected to contaminate the wide $B^{\pm}$ mass window. Inclusive $b\bar{b}$ MC samples used in previous MC studies of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ [72] contained too few events after offline selection to identify precisely which physics backgrounds are expected to contaminate the B mass signal region, so the choice of which background modes to consider is based on the branching fractions of the decay modes (according to the PDG world averages [23]), the number and types of missing particles in the decay, and, where appropriate, the types of particles which have been mis-identified. This section also outlines how the three background types are quantified in the final fit.

4.6.1 Combinatoric background

The $B^{\pm}$ invariant-mass distribution of combinatoric background decays is determined from a fit to the D invariant-mass side-bands of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ combinations in the data sample. Since the level of contamination of combinatoric background in the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decay channel after offline cuts is small, and these data are used in the cut optimisation, the D side-bands for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ are used to construct the combinatoric background PDF, with the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ background assumed to have the same distribution.

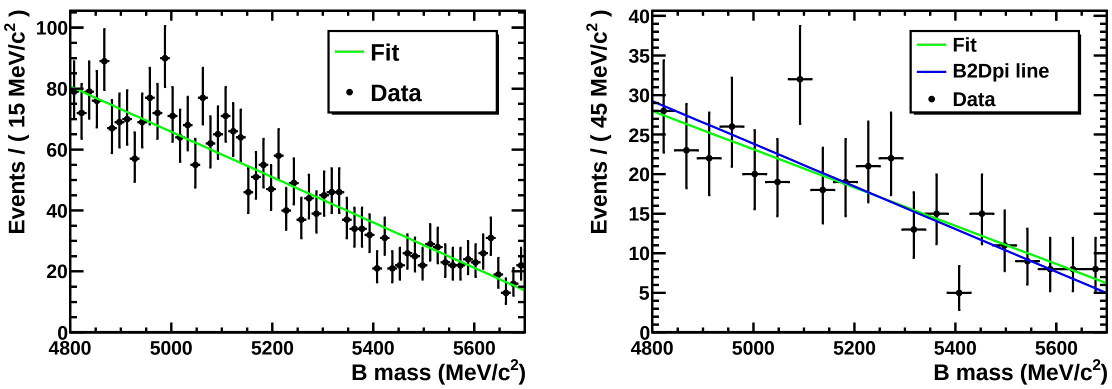
The $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ mass combinations are fitted with a first order Chebychev polynomial. Table 4.10 shows the fitted slope parameter and its fit error for various PID cuts. The results indicate that there is some variation with PID cut, the slope tending to steepen slightly as the PID cut is loosened. However, the results for PID cuts of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 8$ and < 0 are consistent with the median PID cut $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 4$ within uncertainties.

Figure 4.12a shows the fit to the B invariant-mass distribution of combinatoric background reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ residing in the D invariant-mass side-bands for a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 4$. Figure 4.12b shows the PDF determined from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ mass combinations superimposed on the B invariant-mass distribution of combinatoric background reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ for a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 4$, where the PDF has been normalised to the level of

$\log(\mathcal{L}_K/\mathcal{L}_\pi)$	∇	$\sigma(\nabla)$
< 8	-0.67	0.03
< 4	-0.71	0.03
< 0	-0.73	0.03

Table 4.10: PDF parameters from fits to the B invariant-mass distribution of $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ mass combinations in the D invariant-mass data side-bands for various PID cuts. These reconstructed decays are assumed to be formed from the random combination of tracks. The first column indicates the PID cut applied, whilst the second and third columns show the fitted slope and its fit error respectively.

$B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ background. In Fig. 4.12b, the upper side-band (*i.e.* candidates with a D invariant mass larger than the D signal mass window) is not considered, since the B invariant-mass distribution in this region shows large fluctuations (due to a low number of events and/or the presence of non-combinatoric backgrounds). Figure 4.12b indicates a good agreement between the distribution of the combinatoric background reconstructed as $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ and the chosen PDF parameters. A fit to the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ side-band data is also shown in Fig. 4.12b and the result of -0.64 ± 0.08 is also consistent with the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ fit.



(a) (b)

Figure 4.12: $B^\pm$ invariant-mass distribution for the (a) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ and (b) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ candidates residing in the D invariant mass side-bands. A PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 4$ has been applied to the data in (a), whilst a PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 4$ is applied in (b). In both plots, the green lines correspond to linear fits to the data. In (b), the fit from $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ candidates normalised to the number of $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ candidates is represented by the blue line.

4.6.2 Low mass background

As previously mentioned, the low mass backgrounds are decays in which one or more particles are missed in the reconstruction. Several exclusive decay modes have been considered in which there is a single missing particle, and for each of these decay modes MC samples have been produced and reconstructed. The list of decay modes investigated is summarised in Tab. 4.11.

The reconstructed events are modelled with an *adaptive Kernel Density Estimation* (KDE) PDF (see Eq. (A.10)). Each background type is assigned a weight, as was done for the signal MC sample used in the cut optimisation, which represents the number of expected decays in the $\mathcal{L}_{\text{int}} = 35.6\ \text{pb}^{-1}$ data sample. The weight for a given decay mode, w, is calculated by

$$w = 2 \cdot \sigma_{b\bar{b}} \cdot f(b \to B^-) \cdot \mathcal{L}_{\text{int}} \cdot \mathcal{B} \cdot \varepsilon_{\text{acc}} \cdot \varepsilon_{\text{sel|acc}} \varepsilon_{\text{trig|sel}} \cdot \varepsilon_{\text{PID|sel}} \tag{4.2}$$

where $\sigma_{b\bar{b}}$ is the total $b\bar{b}$ production cross-section, $f(b \to D^0)$ is the fraction of $b/\bar{b}$ quarks that hadronise to a B^-/B^+, $\mathcal{L}_{\text{int}}$ is the integrated luminosity of the data samples, $\mathcal{B}$ is the branching fraction of the decay mode and $\varepsilon_{\text{PID|sel}}$ is the efficiency of the PID cut on the bachelor track, or the mis-ID rate if the reconstruction of the decay mode requires the bachelor to be mis-identified. The trigger efficiencies are calculated from the trigger cuts simulated in the MC, whilst the PID efficiencies are determined from the efficiency tables in Sec. 4.8. All other efficiencies are indicated in Tab. 4.11, as are the branching fractions of the decay modes. The weight for each background type is normalised to the number of low-mass background decays in the MC sample and these normalised weights are then used to construct the PDF. The $B^+ \to K^{*+}\overline{D}^0$ and $B^0 \to K^{*0}\overline{D}^0$ components are not included in the final fit, since their contamination of both the $B \to DK$ and $B \to D\pi$ signal modes is negligible.

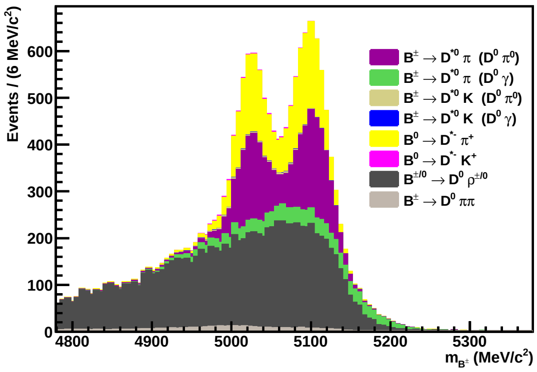
It is important to note that the backgrounds considered involve decay modes for which the D decays to $K^\pm\pi^\mp$ and *not* $K^\pm\pi^\mp\pi^\pm\pi^\mp$, since at the time the MC samples were produced, only decay models in which $D^0 \to K^-\pi^+$ were available in GAUSS. Consequently, the MC

Decay mode	BF ($\times 10^{-4}$)	ε_{acc}	$\varepsilon_{\text{trig\|sel}}$	$\varepsilon_{\text{sel\|acc}}$
$B^+ \to \pi^+ \overline{D}^{*0}$ ($\overline{D}^{*0} \to \pi^0 \overline{D}^0$)	1.25	0.1477	0.37	0.0585
$B^+ \to K^+ \overline{D}^{*0}$ ($\overline{D}^{*0} \to \pi^0 \overline{D}^0$)	0.10	0.1509	0.37	0.0562
$B^+ \to \rho^+ \overline{D}^0$ ($\rho^+ \to \pi^+ \pi^0$)	5.21	0.1516	0.36	0.0270
$B^+ \to K^{*+} \overline{D}^0$ ($K^{*+} \to K^+ \pi^0$)	0.07	0.1578	0.37	0.0296
$B^+ \to \pi^+ \overline{D}^{*0}$ ($\overline{D}^{*0} \to \gamma \overline{D}^0$)	0.77	0.1477	0.37	0.0357
$B^+ \to K^+ \overline{D}^{*0}$ ($\overline{D}^{*0} \to \gamma \overline{D}^0$)	0.06	0.1509	0.37	0.0343
$B^0 \to \pi^+ D^{*-}$ ($D^{*-} \to \pi^- \overline{D}^0$)	0.73	0.1573	0.36	0.0875
$B^0 \to K^+ D^{*-}$ ($D^{*-} \to \pi^- \overline{D}^0$)	0.06	0.1602	0.36	0.0848
$B^0 \to \rho^0 \overline{D}^0$ ($\rho^0 \to \pi^+ \pi^-$)	0.12	0.1516	0.36	0.0270
$B^0 \to K^{*0} \overline{D}^0$ ($K^{*0} \to K^+ \pi^-$)	0.01	0.1578	0.37	0.0296
$B^0 \to \overline{D}^0 \pi^+ \pi^-$	0.33	0.1499	0.35	0.0257

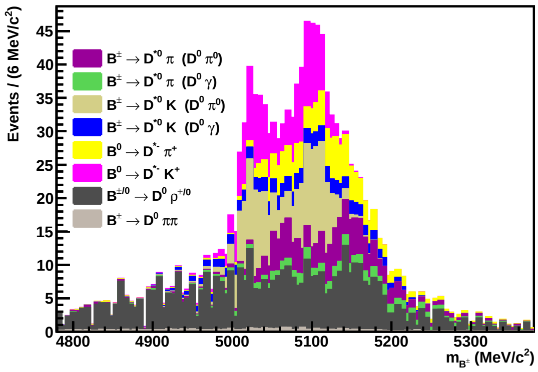
Table 4.11: Summary of simulated backgrounds used to model the B mass distribution of the low-mass background contamination of $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ decays.

samples are stripped using the cuts employed in *StrippingB2DXWithD2hhLine*, shown in Tab. 4.2 and selected with the $B^\pm \to (K^\pm\pi^\mp)_D h^\pm$ offline selection cuts shown in Tab. 4.5. The branching fractions quoted in Tab. 4.11 have been calculated assuming that the D decays to $K^\pm\pi^\mp$, though this has no impact on the background model used in the final fit, since the yield of the low mass backgrounds is a free parameter in the fit.

This background study assumes that the *relative contribution* (*i.e.* the weight of the background type divided by the sum of weights) of each background type will be independent of the D decay mode, and the difference in $B^\pm$ mass resolution between the two-body and four-body D decays will have a negligible impact on the low-mass shapes. The first assumption is equivalent to asserting that the ratio of efficiencies between any two background modes is the same regardless of the D final-state. It is of course, not possible to test this assumption due to the lack of low-mass background MC samples for which the D decays to the four-body mode. However, samples do exist for the signal modes. The product of the generator, trigger and selection efficiencies for $B^\pm \to (K^\pm\pi^\mp)_D K^\pm$ divided by the product of efficiencies for $B^\pm \to (K^\pm\pi^\mp)_D \pi^\pm$ is 1.028 ± 0.016 [3]. The corresponding ratio of efficiencies for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ decays is determined from the efficiencies in Tab. 4.17 to be 1.005 ± 0.018. The difference between these two efficiency ratios



(a)



(b)

Figure 4.13: Offline-selected MC samples of various low-mass backgrounds. (a) shows the B invariant-mass distribution of the offline-selected truth-matched MC candidates from each background mode reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, \pi^{\pm}$, with a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 4$, whilst (b) shows the same background modes reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, K^{\pm}$, with a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 4$.

is 0.023 ± 0.024, which is consistent with zero within statistical uncertainties. This gives a good indication that the relative contributions of the background modes are likely to be independent of the D decay mode. A cross-check of the effect of $B^{\pm}$ mass resolution indicates that data/MC differences have a more significant impact on the mass resolution than the D decay mode, and the data/MC differences are not expected to have a significant impact on the low-mass parametrisation. This cross-check discussed in more detail in Appendix C.

Figure 4.13 shows the B invariant-mass distribution of the low-mass background modes considered, in which the aforementioned weights have been applied to the MC samples and the decays have been reconstructed under $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, K^{\pm}$ mass hypotheses. Comparing with Fig. 4.11, the background model agrees well with the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ data, except below $\approx 5000\,\mathrm{MeV}/c^2$. This is partly a result of semi-leptonic background modes which are fitted separately, as discussed in Sec. 4.6.3. The remaining discrepancy is assumed to be from partially-reconstructed decay modes which are not considered in the low mass model, such as decays involving more than one missing particle. The $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement is performed with candidates above $5100\,\mathrm{MeV}/c^2$ in $B^{\pm}$ mass, so the omitted decays should not influence the result.

4.6.3 Semi-leptonic Modes

The semi-leptonic background PDF is determined using a data-driven approach, using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ candidates for which the bachelor track has hits in the muon chambers. The main semi-leptonic background contributions are expected to be from decays for which the D is correctly reconstructed, but the bachelor muon has been mis-identified as a kaon or pion, such as $B^{+} \to \overline{D}^{0}\mu^{+}\nu_{\mu}$, where the $\overline{D}^{0}$ decays to the correct final state and the neutrino has been lost. Such decays have branching fractions of around an order of magnitude larger than $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$, but the reconstruction efficiency is lower due to the missing neutrino mass and the $\mu \to (K, \pi)$ mis-ID fraction. Figure 4.14 shows the distribution of the semi-leptonic B decays reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$. The data are fitted with a Gaussian PDF, with the central value at very low mass, such that only the tail is visible within the fitted mass range. A clear peak is visible around the nominal $B^{\pm}$ mass, and is fitted with a second Gaussian. These are most likely a combination of hadronic *punch-through* (material interactions of a hadron in the muon chambers which causes the hadron to be incorrectly identified as a muon) and *decays in flight* (where the kaon or pion from a true signal candidate subsequently decays to a muon and neutrino). Both of these cases correspond to true signal decays, so the second Gaussian component is not included as a background in the final fit.

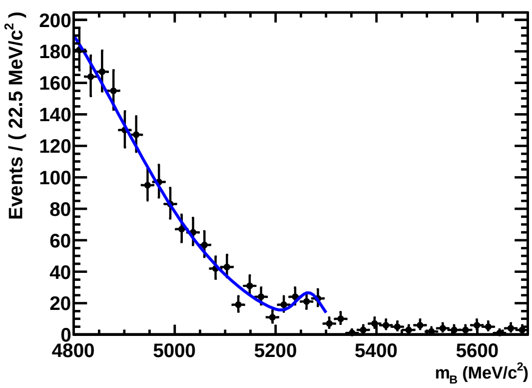
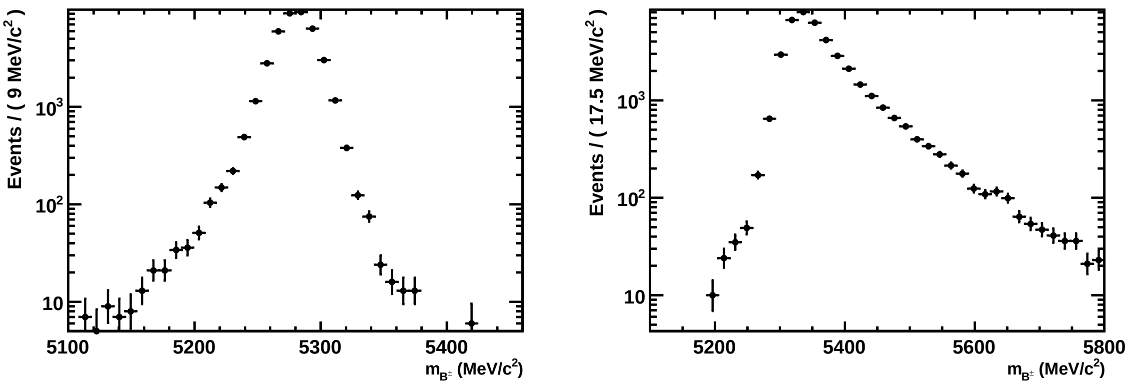


Figure 4.14: Fit of semi-leptonic B decays reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ in which the bachelor has hits in the muon chambers. A small fraction of signal decays is seen to be muon-like and they are taken into account by fitting a Gaussian peak.

4.7 Signal and mis-ID PDFs

The reconstructed $B^{\pm}$ invariant-mass distribution of correctly identified signal decays is modified from its natural distribution due to detector resolution effects. Figure 4.15a shows the reconstructed $B^{\pm}$ mass distribution for truth-matched $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ MC. The results indicate that the mass distribution of the correctly identified signal decays is broadly Gaussian (parabolic for a logarithmic vertical scale). However, the loss of energy due to final-state radiation (FSR) and material interaction results in a non-Gaussian tail at low mass. There is also a less substantial high-mass tail. The origin of the latter tail is less obvious, but may be an indication that the detector resolution is not perfectly Gaussian.



(a) (b)

Figure 4.15: $B^{\pm}$ mass distribution for (a) correctly identified $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ MC decays, and (b) the same decays in which the $B^{\pm}$ mass has been recalculated from the true track four-momenta, assuming a kaon mass hypothesis for the bachelor. Non-Gaussian tails can be seen in both the high- and low-mass regions for the correctly reconstructed decays. For the mis-identified decays, a long tail at high mass can be seen.

It is also important to have a model of the mass distribution of candidates for which the wrong bachelor mass hypothesis has been assigned (*i.e.* $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays mis-identified as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ and vice versa). Even without detector resolution effects, the assignment of the incorrect mass hypothesis means that the $B^{\pm}$ mass distribution can be reconstructed with a range of values. Figure 4.15b shows the $B^{\pm}$ mass distribution for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ truth-matched MC where the $B^{\pm}$ mass has been calculated from the true four-momenta of the tracks (according to the MC information) and the bachelor has been re-assigned a $K^{\pm}$ mass hypothesis. This $B^{\pm}$ mass calculation means that the modification of the $B^{\pm}$ invariant mass from its natural distribution is a result only of the kinematics of the wrong-mass assignment, and not of the detector resolution. By comparison with Fig. 4.15a, it can be seen that the peak value is shifted by about $+50\,\mathrm{MeV}/c^2$, with a very long high-mass tail.

4.7.1 Signal PDF

The signal $B^{\pm}$ mass distribution is modelled by a Cruijff PDF. The choice of PDF to model this distribution was studied in Ref. [3], and the results indicate that of the PDFs studied, the Cruijff PDF provides the best fit to the data. The fitted $B^{\pm}$ mass distribution for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ MC and data (plus background) are shown in Fig. 4.16. The pull distributions, shown below the main plots, indicate that the fits are reasonable.

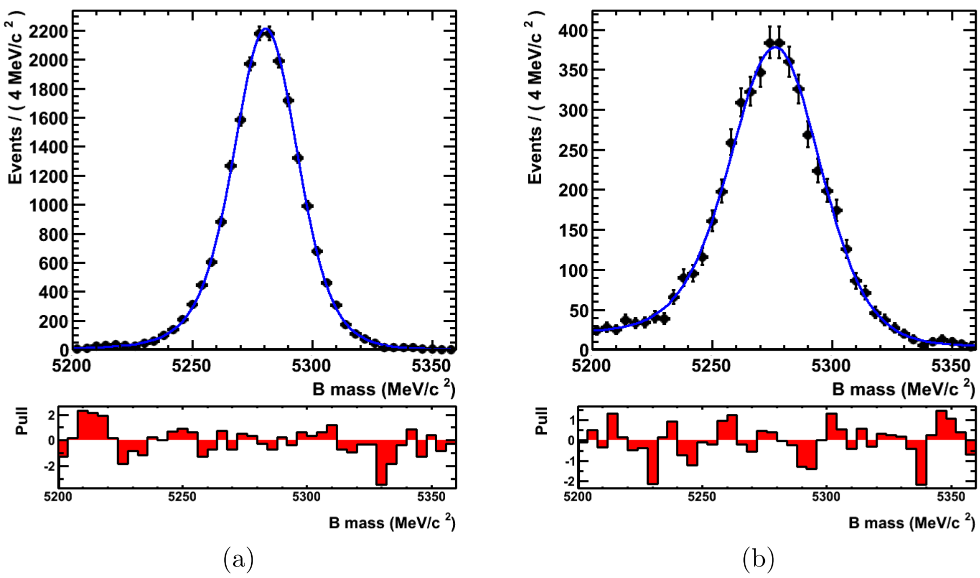


Figure 4.16: Cruijff PDF fits to $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ (a) MC and (b) data. The pull values in uniform bins are shown below the mass plot. The data plot has been fitted with all background PDFs included (not shown) and a bachelor PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < -8$ has been applied.

Table 4.12 shows the Cruijff fit parameters for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ MC and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ data for various PID cuts applied to the bachelor. For the fit to the data, the various background components are considered, including a mis-ID component (using a Cruijff PDF convolved with a Novosibirsk function, as described in Sec. 4.7.2). Encouragingly, the mean $B^{\pm}$ mass in MC is consistent between $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$, and consistent with varying PID cut, though as previously indicated in Sec. 4.3.7, the mean is lower in data than MC (with the MC mean close to the nominal value) and the width of the peak is larger. In addition, the width tends to decrease with tightened PID cut. In general, the tail parameters α_{L} and α_{R} tend to decrease with a tighter PID cut, suggesting that the radiative tails are being suppressed. However, the fact that the tails contain a low number of events and have a lot of freedom to vary without affecting the PDF shape is indicative of the relatively large uncertainties associated with these parameters.

Due to the limited event numbers, the parameters for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ cannot be extracted reliably from the data. Instead, the parameters from the MC and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ data presented in Tab. 4.12 are used to estimate them. Assuming the ratio of the signal PDF parameters for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ in data and MC are the same, then the fit parameters for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ in data can be determined by

$$x_K^{\mathrm{data}} = x_\pi^{\mathrm{data}} \frac{x_K^{\mathrm{MC}}}{x_\pi^{\mathrm{MC}}} \,, \tag{4.3}$$

where x_K and x_π represent the fit parameters for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ candidates respectively, and the superscript indicates whether the candidates are from truth-matched MC or data.

4.7.2 Mis-identified PDF

The PDF parameters for mis-identified $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ decays are determined using a four stage process.

In the first stage, a B invariant-mass distribution fit to $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ candidates is performed with no bachelor PID cut applied. The fit includes the signal PDF, the background components defined in Sec. 4.6, and a *mis-identified* (mis-ID) component. The mis-ID component comprises the convolution of two PDFs. The first is the same Cruijff

Decay type	$\log(\mathcal{L}_K/\mathcal{L}_\pi)$	α_{L}	α_{R}	μ (MeV/c^2)	σ (MeV/c^2)
$B \to D\pi$ MC		0.132±0.005	0.094±0.006	5281.3±0.3	13.3±0.2
$B \to DK$ MC		0.127±0.005	0.114±0.005	5280.2±0.2	12.8±0.2
$B \to D\pi$		0.279±0.047	0.158±0.020	5275.6±0.5	17.1±0.6
$B \to D\pi$	< 0	0.212±0.038	0.153±0.017	5275.6±0.5	16.4±0.6
$B \to D\pi$	< -4	0.157±0.037	0.100±0.031	5276.1±0.5	16.7±0.8
$B \to D\pi$	< -8	0.170±0.037	0.130±0.023	5275.9±0.5	15.9±0.7

Table 4.12: Cruijff PDF parameters from fits to $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ data and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,h^\pm$ MC, with and without PID cuts. $B \to D\pi$ refers to $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ candidates, whilst $B \to DK$ refers to $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$. The $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut indicated is for the bachelor track. If no $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut has been applied, this is indicated by a blank entry.

PDF used to fit the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ signal candidates, which describes the detector resolution part of the mis-ID shape, whilst the second is a Novosibirsk PDF (Eq. (A.3)) used to describe the mis-ID shape. The Cruijff parameters are allowed to float freely in the fit, whilst the Novosibirsk parameters are fixed from a fit to the MC shape seen in Fig. 4.15b. In addition, the fraction of the mis-ID PDF with respect to the signal PDF is fixed to the central value of the PDG world average of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$.

The second stage involves determining fit parameters for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ decays in which the bachelor pion has been mis-identified as a kaon directly from the data (i.e. without input from MC). Firstly, a background subtraction method is applied to the B mass distribution of $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ candidates. Using the fit results from the previous stage, each candidate is assigned a weight $S(m_B)/T(m_B)$, where $S(m_B)$ is the integral of the signal PDF over a narrow range about the reconstructed B mass of the candidate, m_B, and $T(m_B)$ is the equivalent integral for the combined signal plus background PDF. The mass of the bachelor pion is then re-assigned as a kaon for each weighted candidate, and the B mass recalculated. Finally, this weighted B mass mis-ID distribution is fitted with a double Crystal Ball PDF. These values are not used in the final fit, instead, they are used as input to the third stage, in which the parameters for the PDF representing mis-identified $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ decays are estimated.

As with the signal PDF parameters, it is unfeasible to determine the parameters for mis-identified $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays directly from the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ data due to the low number of events. Rather, the mis-ID parameters from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ data from the previous stage are used along with the equivalent parameters from MC, using Eq. (4.3). However, this method cannot be used to estimate the mean parameter of the double Crystal Ball because the mean of the signal Cruijff PDF is different in data and MC. Instead, the ratio of the difference between the mean of the double Crystal Ball, and the mean of the Cruijff PDF for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ candidates is assumed to be equivalent in MC and data. If m is the mean parameter of the double Crystal Ball PDF and μ is the mean parameter of the Cruijff PDF, then the mean of the double Crystal Ball shape for mis-identified $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays, m_K^{data}, to be used to fit the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ data is given by

$$m_K^{\mathrm{data}} = \mu_K^{\mathrm{data}} + \left(m_\pi^{\mathrm{data}} - \mu_\pi^{\mathrm{data}}\right) \frac{m_K^{\mathrm{MC}} - \mu_K^{\mathrm{MC}}}{m_\pi^{\mathrm{MC}} - \mu_\pi^{\mathrm{MC}}} \,, \tag{4.4}$$

where μ_K^{data} is determined from scaling using the MC and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ data, as described in Sec. 4.7.1. The resulting values of the double Crystal Ball PDF parameters for the mis-identified $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays are given in Tab. 4.13.

Parameter	α_{L}	α_{R}	n_{L}	n_{R}	μ (MeV/c^2)	σ (MeV/c^2)
Value	0.708	-2.42	4.06	5.53	5219.9	28.6

Table 4.13: Double Crystal Ball mis-ID parameters for mis-identified $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ decays, determined by scaling the equivalent parameters according to MC and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ data.

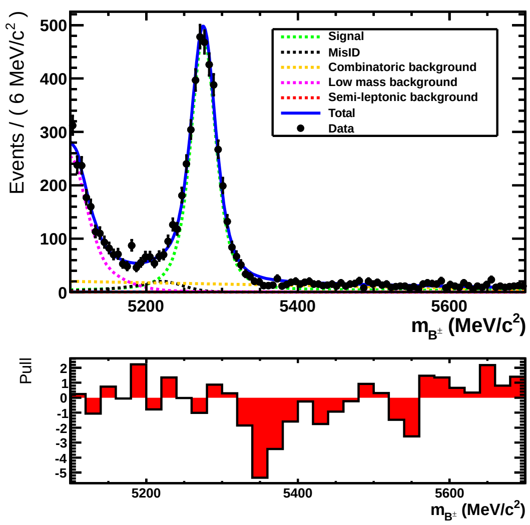
In the final stage, the data are once again weighted using the double Crystal Ball parameters determined from the previous stage. The weighted data are then fitted a final time to determine the parameters of the mis-ID $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ PDF. The results for various bachelor PID cuts are shown in Tab. 4.14.

Figure 4.17 shows the fit to the B mass distribution for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ and

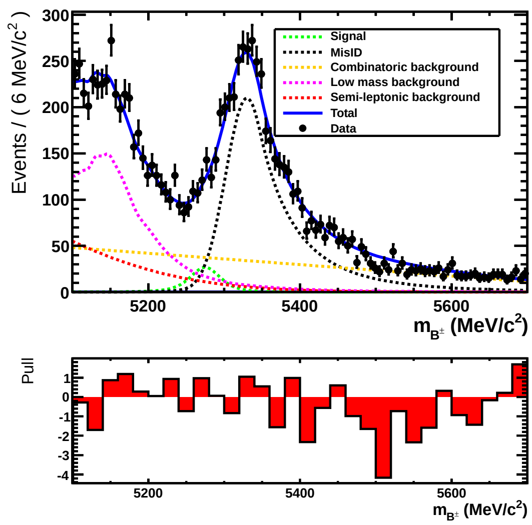
PID cut	$\alpha_{\rm L}$	$\alpha_{\rm R}$	$n_{\rm L}$	$n_{\rm R}$	μ (MeV/c^2)	σ (MeV/c^2)
none	1.97±0.26	-0.64±0.05	3.73±1.83	4.64±0.68	5330.6±1.2	29.2±1.2
> -4	2.09±0.44	-1.05±0.12	3.29±4.57	1.98±0.34	5319.4±1.5	29.0±1.5
> -2	2.25±0.36	-1.03±0.15	2.85±4.27	2.00±0.43	5318.4±2.0	30.6±1.7
> 0	2.33±0.39	-0.89±0.14	2.34±4.85	2.11±0.48	5315.9±2.4	29.8±1.9
> 2	2.21±0.53	-0.74±0.16	3.00±4.18	2.28±0.64	5311.7±3.7	28.9±2.9
> 4	1.89±0.90	-0.45±0.22	4.03±4.17	3.29±2.68	5302.0±16.0	25.0±4.0
> 6	1.83±0.91	-0.37±0.13	3.55±4.31	3.45±1.66	5294.3±5.3	21.0±4.2
> 8	1.75±0.84	-0.24±0.12	2.89±4.30	6.85±8.16	5286.5±7.7	17.5±5.3
> 10	1.71±0.51	-0.26±0.14	2.92±4.42	4.49±3.14	5283.8±7.4	15.8±15.4

Table 4.14: Double Crystal Ball parameters obtained from fits to weighted $B^{\pm} \rightarrow (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ data in which the bachelor pion has been misidentified as a kaon for various bachelor PID cuts.

$B^{\pm} \rightarrow (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ candidates in which all signal and background PDFs have been included (except for the contribution from punch-through and decays in flight, mentioned in Sec. 4.6.3) with no bachelor PID cut applied. The $B^{\pm} \rightarrow (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ and $B^{\pm} \rightarrow (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ mis-ID parameters are taken from the last two stages of the mis-ID parameter study. These parameters represent the final iteration.



(a)



(b)

Figure 4.17: Fit to the B mass distribution for (a) $B^{\pm} \rightarrow (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ and (b) $B^{\pm} \rightarrow (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ data, with all signal and background PDFs included. A double Crystal Ball PDF is used to model the mis-ID background. The pull values of the total PDFs (signal+backgrounds) in uniform bins are shown below the mass plots.

4.8 PID calibration

As mentioned in previous sections, to distinguish between $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ decays, RICH PID information for the bachelor track is used, namely the difference in the log-likelihood of the kaon and pion mass hypotheses, $\log(\mathcal{L}_K/\mathcal{L}_\pi)$. Unlike most selection cuts employed in the analysis, the systematics associated with the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut applied to the bachelor track do not cancel in the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement, so it is vital to know the efficiency of retaining kaons (or pions) and the proportion of pion (kaon) contamination remaining.

Since the PID $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ variables are not well described in the MC, the PID efficiencies and mis-ID rates are determined using $D^{*+} \to \pi^+ D^0(K^-\pi^+)$ decays from $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data. A series of cuts devoid of PID requirements are applied to these data to provide large, high-purity samples of *calibration* pions and kaons.

The distributions of kinematic variables such as $|\vec{p}|$, p_T and η for the bachelor kaon or pion from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ are known to be different from those from the calibration samples. Since the calibration samples pass different triggers than the physics samples, the number of tracks in the event (nTracks) will also be different. Since the RICH PID performance is correlated to the above kinematic variables, a means of *reweighting* these variables is needed. The PID calibration technique involves weighting the calibration samples such that the distributions of variables agree with those of the signal. If the PID performance is perfectly described by the kinematic variables used, then the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distribution of the weighted calibration sample should match that of the signal.

The signal tracks are obtained by fitting the $B^{\pm}$ invariant mass of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ candidates with all signal, background and mis-ID component models included, as indicated by Fig. 4.17a. The ${}_s\mathcal{P}lot$ technique [74] is then used to extract the signal distributions of the $|\vec{p}|$, η and nTrack variables. These ${}_s\mathcal{P}lots$ are then used to weight the calibration kaon and pion tracks, using a weighting scheme comprising eighteen $|\vec{p}|$, four η and four nTrack bins respectively. The weighted and un-weighted $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ calibration distributions, along with the ${}_s\mathcal{P}lot$ of the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distribution of bachelor tracks from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ decays, are shown in Fig. 4.18.

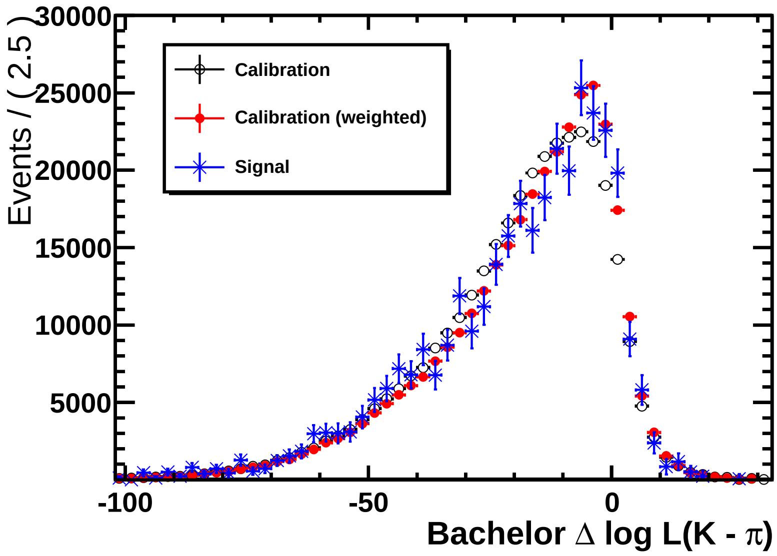


Figure 4.18: The $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distribution of bachelor π^- signal tracks (blue cross) from an ${}_s\mathcal{P}lot$ fit to $B^- \to (K^-\pi^+\pi^-\pi^+)_D\,\pi^-$ compared to the same distribution for π^- calibration tracks, before (black circle) and after (red filled circle) reweighting. A binning scheme of 18 bins in $|\vec{p}|$, 4 bins in η and 4 bins in nTracks has been used to perform the reweighting of the calibration sample.

Tables 4.15 and 4.16 show the PID efficiencies for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ candidates respectively for a wide range of PID cuts. As will be discussed in the next section, the final fit is performed simultaneously in two “slices”. The $K^{\pm}$ and $\pi^{\pm}$ mis-identification rates in the two slices are determined by subtracting the PID efficiencies in Tab. 4.15 and Tab. 4.16 respectively from unity, and inverting the inequality relation (*e.g.* $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 0$ should be replaced by $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 0$).

4.9 Description of the final $B^{\pm}$ mass fit

In previous sections, the parametrisation of the B invariant-mass distribution of the signal decays and various background types has been discussed. This section summarises how the PDFs describing these distributions, as well as the bachelor PID efficiency information, are combined in the final fit to the $B^{\pm}$ invariant-mass distributions of the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ data.

$\log(\mathcal{L}_K/\mathcal{L}_\pi) >$	ϵ_{K^+}	$\sigma(\epsilon_{K^+})$	ϵ_{K^-}	$\sigma(\epsilon_{K^-})$
-8.0	0.9916	0.0002	0.9921	0.0002
-6.0	0.9874	0.0002	0.9883	0.0002
-4.0	0.9797	0.0003	0.9803	0.0003
-2.0	0.9653	0.0004	0.9654	0.0004
0.0	0.9395	0.0005	0.9388	0.0005
2.0	0.8999	0.0006	0.8998	0.0006
4.0	0.8500	0.0008	0.8514	0.0008
6.0	0.7971	0.0009	0.7983	0.0009
8.0	0.7429	0.0009	0.7451	0.0010

Table 4.15: Bachelor $K^{\pm}$ PID efficiencies and statistical uncertainties for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$.

$\log(\mathcal{L}_K/\mathcal{L}_\pi) <$	ϵ_{π^+}	$\sigma(\epsilon_{\pi^+})$	ϵ_{π^-}	$\sigma(\epsilon_{\pi^-})$
-8.0	0.6557	0.0010	0.6532	0.0010
-6.0	0.7131	0.0010	0.7103	0.0009
-4.0	0.7730	0.0009	0.7699	0.0009
-2.0	0.8317	0.0008	0.8287	0.0008
0.0	0.8852	0.0007	0.8818	0.0006
2.0	0.9283	0.0005	0.9246	0.0005
4.0	0.9569	0.0004	0.9537	0.0004
6.0	0.9744	0.0003	0.9716	0.0003
8.0	0.9839	0.0002	0.9820	0.0002

Table 4.16: Bachelor $\pi^{\pm}$ PID efficiencies and statistical uncertainties for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$.

As previously mentioned, the observables $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A_{fav}^{K3\pi}$ are determined from an unbinned maximum-likelihood fit to the $B^{\pm}$ invariant mass performed using the RooFit framework. The observables are included as parameters in the fit so that the statistical uncertainties and fit-parameter uncertainties can be extracted directly. The fit is split into two sub-samples. The *pass slice* selects $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ candidates reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ in which the bachelor is "kaon-like", by applying a bachelor PID cut in which the selected candidates have $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ greater than a chosen cut value. The *fail slice* selects candidates reconstructed as $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ with a pion-like bachelor, by applying a $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut less than the same cut value as the pass slice (*e.g.* $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 4$ for the pass slice and $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 4$ for the fail slice). Both slices will have some contamination from decays with mis-identified bachelor tracks. Fitting the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ yields simultaneously in the two slices, and using the PID efficiencies/mis-ID rates from Sec. 4.8 are crucial to determine the mis-ID, signal and background PDFs, and thus accurately determine the yields and parameters of interest. In order to measure the rate asymmetry, both the pass and fail slices are further separated into B^+ and B^- decays.

4.9.1 Summary of the fit components

Each of the four slices (B^+ and B^-, pass and fail) has five components:

- signal;
- mis-ID;
- combinatoric background;
- partially-reconstructed background;
- semi-leptonic background.

These components are described in detail in Secs. 4.6 and 4.7. The models used for the PDFs are the same for the B^+ and B^- slices. However, in some cases, the parameters are allowed to vary separately in the pass and fail slices, whilst in others, some of the parameters are fixed in one or both slices. This is discussed in more detail below. In addition, an explanation is given of how the yields for the various PDF components are determined.

$B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$

The parameters of the signal (fail slice) and mis-ID (pass slice) PDFs are allowed to vary freely in the fit. The total $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ yield is also a free parameter in the fit. The proportion of the total fitted $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ candidates in the fail slice is fixed to the value determined by the PID calibration (see Sec. 4.8).

$B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$

As discussed in Secs. 4.7.1 and 4.7.2, MC is used to fix the relative size of the parameters of the signal (pass slice) and mis-ID (fail slice) PDFs relative to the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ parameters (excluding the yield). The total $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ yield *relative* to $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ (*i.e.* N_K/N_π, where N_K and N_π are the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ yields in both slices respectively) is a free parameter in the fit. As with $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$, the proportion of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$ decays in the pass slice is fixed to the PID efficiency determined by the PID calibration study.

Combinatoric background

As indicated in Sec. 4.6.1, the coefficient of the combinatoric background PDF is fixed from the D mass side-bands, and is set to the same value for all slices. Since the variation of the slope over the PID cuts considered is consistent within the fit uncertainties, the slope is chosen to be the same for all PID cuts considered, namely $-0.71(\,\mathrm{MeV}/c^2)^{-1}$, which is the central value of the fitted slope for a PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 4$. The yield of this background is a free parameter in the fit, and is independent in all slices.

Partially-reconstructed background

A different PDF is used for the pass and fail slices, and the yield is allowed to vary independently in each slice.

Semi-leptonic background

The mean and width of the semi-leptonic PDF are fixed in the fit, using the values from the combined analysis of $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\, h^{\pm}$, $B^{\pm} \to (K^{+}K^{-})_D\, h^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ decays described in Ref. [3]. The width is fixed to 300 MeV/c^2 in all slices, and the mean is set to 4750 MeV/c^2 for the pass slice and 4650 MeV/c^2 for the fail slice. These numbers are broadly in agreement with those from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ decays alone.

4.9.2 Multiple candidates

It is required that only one $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ candidate per event is used in the final fit (except in the case of events containing multiple *pp* interactions, as discussed below). The decision on which candidate is retained is made after all offline selection cuts are applied, and is chosen across all signal decay modes. In the case of multiple candidates in an event, the one with the smallest B vertex χ^2 is retained. In the case that two candidates have the same vertex χ^2 (within the finite precision of the LHCb vertexing tool) the candidate closest to the nominal value of the D mass is selected. To ensure that events with multiple *pp* interactions are accounted for, two or more signal candidates are allowed in an event provided that the B decay vertices are separated by more than 2 mm in the z direction (along the direction of the beam). Approximately 7% of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, h^{\pm}$ mass combinations passing the offline selection requirements are rejected due to the above requirements.

4.9.3 Asymmetries

In addition to *CP*-violating physics processes, rate asymmetry may be a result of:

- the different hadronic production factors for B^{+} and B^{-} (A_{prod});
- the different interaction lengths of K^{+} and K^{-} in the detector material (A_{Kdet});
- the difference in the efficiency of collecting positive and negative tracks due to a *left–right detector asymmetry* (A_{LR}) for example as a result of an asymmetry in the material budget on the two sides of the detector.

The left–right asymmetry can only be extracted if the fit considers magnet up and magnet down data separately. Due to the low number of events, additional splitting of the data is unrealistic, so this asymmetry is assumed to be zero. The production asymmetry is common to all modes. The interaction length asymmetry is different for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$, since this depends upon the total number of kaons in the final state. Any difference in the interaction lengths of π^+ and π^- ($A_{\pi det}$) is fixed to zero in the final fit in line with expectations.

Combining asymmetries

The asymmetry between two quantities such as the number of B^+ and B^- decays, N_+ and N_- respectively, can be expressed in one of two ways,

$$A = \frac{N_- - N_+}{N_- + N_+} \,, \tag{4.5a}$$

$$a = \frac{N_-}{N_+} \,. \tag{4.5b}$$

The former representation, A, is used to express physical asymmetries, whilst the latter, a, is useful when combining multiple asymmetries. The following equations indicate how to convert between the two forms:

$$A = \frac{a-1}{a+1} \,, \tag{4.6a}$$

$$a = \frac{1+A}{1-A} \,. \tag{4.6b}$$

The total number of $B^{\pm}$ decays, $N_{\rm tot}$ is given by $N_{\rm tot} \equiv N_- + N_+$. Substituting this and Eq. (4.6a) into Eq. (4.5a) and rearranging gives

$$N_- - N_+ = N_{\rm tot} A \equiv N_{\rm tot} \frac{a-1}{a+1} \,. \tag{4.7}$$

Finally, substituting Eq. (4.5b) into Eq. (4.7), and solving for N_+ and N_- gives the number of B^+ and B^- decays expressed in terms of the total rate asymmetry a as

$$N_+ = \frac{N_{\rm tot}}{a+1}\ , \tag{4.8a}$$

$$N_- = \frac{N_{\rm tot}}{\frac{1}{a}+1}\ . \tag{4.8b}$$

Asymmetries in the form of Eq. (4.5b) are combined by multiplication if they are in the same "sense" (*e.g.* both representing an excess of positive particles) or division if they are "opposite" (*e.g.* one represents an excess of positive particles, and the other a negative particle excess). Using this logic, a in Eqs. (4.8a) and (4.8b) can be decomposed into the individual contributions to the rate asymmetry. For $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ candidates, the measured $B^+ \to (K^+\pi^-\pi^+\pi^-)_D K^+$ and $B^- \to (K^-\pi^+\pi^-\pi^+)_D K^-$ yields, N^K_+ and N^K_-, are thus related to the asymmetries by

$$N^K_\pm \equiv \frac{N^K_{\rm tot}}{1+\left(a_p a^K_f a^2_K a^{-1}_\pi\right)^{\pm 1}}\ , \tag{4.9}$$

where $N^K_{\rm tot}$ is the total $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ yield, a_p is the production asymmetry, a^K_f is the physics asymmetry in the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$ system (from *CP*-violating processes), a_K is the kaon detection asymmetry and a_π is the pion detection asymmetry. Similarly, if $N^\pi_{\rm tot}$ is the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ yield, then the number of B^+ and B^- decays in terms of the various rate asymmetries is given by

$$N^\pi_\pm \equiv \frac{N^\pi_{\rm tot}}{1+\left(a_p a^\pi_f a_K\right)^{\pm 1}}\ . \tag{4.10}$$

The pion detection asymmetry is fixed to zero in the fit ($A_{\pi det}$ =0, $a_\pi = 1$), as is the physics asymmetry in $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D \pi^\pm$ decays (A^π_f is expected to be less than 10^{-3}). This leaves three asymmetries from four equations, allowing in principle a simple extraction of all unknowns. The physics asymmetry for $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D K^\pm$, $A^{K3\pi}_{fav} = (a^K_f - 1)/(a^K_f + 1)$ is a key measurement of the analysis, and so it is extracted directly in the fit. However, the number of events in the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ samples is too small to allow all three asymmetries to float in the fit.

The value for A_{prod} is fixed based on the results of an LHCb analysis of $B \to h^{+}h^{-}$, *i.e.* $A_{prod} = -0.024 \pm 0.013 \pm 0.010$ [75]. The value of A_{K}^{det} is also fixed in the fit to $A_{K}^{det} = 0.0153 \pm 0.0194 \pm 0.0029$, as determined from Ref. [3]. However, since the A_{K}^{det} measurement uses the same sample of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ candidates, there is clearly a potential bias. Nevertheless, the statistical uncertainty of the A_{K}^{det} measurement is very large (and is incorporated into the systematic uncertainty of the $A_{fav}^{K3\pi}$ result) and only about 1/3 of the events used to measure A_{K}^{det} come from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ decays, any bias is assumed negligible. Since A_{K}^{det} and A_{prod} are fixed in the fit, their statistical uncertainties are not incorporated into the $A_{fav}^{K3\pi}$ measurement. Instead, they are considered in the systematic uncertainty of $A_{fav}^{K3\pi}$, as described in Sec. 4.10.8.

4.10 Systematics and correction factors

This section details the systematic biases and uncertainties that affect the measurement of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A_{fav}^{K3\pi}$, and correction factors which need to be applied to the values returned by the fit. Since both measurements involve ratios of observables, many systematic effects cancel, but others still need to be accounted for.

4.10.1 Efficiency corrections for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$

The ratio of branching fractions ($\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$) is given by

$$\mathcal{B}_{DK}/\mathcal{B}_{D\pi} = \frac{N_{DK}}{N_{D\pi}} \frac{\varepsilon_{\rm acc}^{D\pi}}{\varepsilon_{\rm acc}^{DK}} \frac{\varepsilon_{\rm sel|acc}^{D\pi}}{\varepsilon_{\rm sel|acc}^{DK}} \frac{\varepsilon_{\rm trig|sel}^{D\pi}}{\varepsilon_{\rm trig|sel}^{DK}} \ . \tag{4.11}$$

Here, N represents the signal yield from the unbinned maximum-likelihood fit, whilst $\varepsilon_{\rm acc}$, $\varepsilon_{\rm sel|acc}$ and $\varepsilon_{\rm trig|sel}$ represent the generator, selection and trigger efficiencies respectively. The superscripts (or subscripts in the case of the yields) DK and $D\pi$ indicate the quantities are determined from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ or $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ mass combinations respectively. In principle, the efficiency of the PID cut also needs to be considered, however, this is absorbed into the fit yields. The generator efficiencies are calculated from the MC production, the selection efficiencies are determined by the selection efficiency study in Sec. 4.3 and the trigger efficiencies are determined from the "default" TCK (`0x002e002a`) simulated in the MC.

Since $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ are topologically similar, the ratio of efficiencies is expected to be approximately unity. However, some differences in the efficiencies are seen (at the % level) as indicated by Tab. 4.17, and these are corrected for in the final $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ result. A systematic uncertainty is assigned by propagating the individual correction-factor statistical uncertainties.

Quantity	$\varepsilon_{D\pi}$	$\delta(\varepsilon_{D\pi})$	ε_{DK}	$\delta(\varepsilon_{DK})$	$\varepsilon_{D\pi}/\varepsilon_{DK}$	$\delta\left(\varepsilon_{D\pi}/\varepsilon_{DK}\right)$
$\varepsilon_{\mathrm{acc}}$	0.1506	0.0002	0.1536	0.0002	0.980	0.002
$\varepsilon_{\mathrm{sel\|acc}}$	0.0274	0.0001	0.0270	0.0001	1.0152	0.007
$\varepsilon_{\mathrm{trig\|sel}}$	0.339	0.004	0.339	0.004	1.00	0.02

Table 4.17: Summary table of generator, selection and trigger efficiencies for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ decays, and the correction factors to be applied to the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement. The first two columns represent the efficiency central values and statistical uncertainties for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$, the second two columns represent the efficiency central values and statistical uncertainties for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$, and the last two columns correspond to the final correction factors and statistical uncertainties for the ratio of branching fractions. The rows in the final column, when summed in quadrature, give the systematic uncertainties associated with the corrections.

4.10.2 PID systematic uncertainty

The efficiency of pion and kaon identification is determined using the method described in Sec. 4.8. The assumption that the $\log\left(\mathcal{L}_K/\mathcal{L}_\pi\right)$ distribution of the bachelor can be considered a function of only three variables, ($|\vec{p}|$, η and nTracks) is only an approximation, as the weighting procedure does not perfectly reproduce the $\log\left(\mathcal{L}_K/\mathcal{L}_\pi\right)$ distribution associated with the signal tracks under study. This results in a systematic uncertainty on the PID efficiencies, determined from the weighted calibration sample. An estimate of the size of this systematic can be made by comparing the efficiencies determined from truth-matched MC with those predicated by the weighting procedure.

The method to determine the above systematic assumes that the *asymmetry* of PID efficiencies is the same in data and MC, where the asymmetry of PID efficiencies is defined by

$$A_{\rm PID}^{\rm (MC,data)} \equiv \frac{\varepsilon_{\rm sig}^{\rm (MC,data)} - \varepsilon_{\rm calib}^{\rm (MC,data)}}{\varepsilon_{\rm sig}^{\rm (MC,data)} + \varepsilon_{\rm calib}^{\rm (MC,data)}} \; . \tag{4.12}$$

Here $\varepsilon_{\rm calib}$ is the PID efficiency determined from the calibration samples, and $\varepsilon_{\rm sig}$ is the signal efficiency. The superscripts "MC" and "data" refer to whether the calibration and signal samples correspond to MC or data respectively. When we assume $A_{\rm PID}^{\rm MC} = A_{\rm PID}^{\rm data}$, then the residual in data, $\varepsilon_{\rm sig}^{\rm data} - \varepsilon_{\rm calib}^{\rm data}$, is given by

$$\left(\varepsilon_{\rm sig}^{\rm data} - \varepsilon_{\rm calib}^{\rm data}\right) = \left(\varepsilon_{\rm sig}^{\rm data} + \varepsilon_{\rm calib}^{\rm data}\right) \left(\frac{\varepsilon_{\rm sig}^{\rm MC} - \varepsilon_{\rm calib}^{\rm MC}}{\varepsilon_{\rm sig}^{\rm MC} + \varepsilon_{\rm calib}^{\rm MC}}\right) \; . \tag{4.13}$$

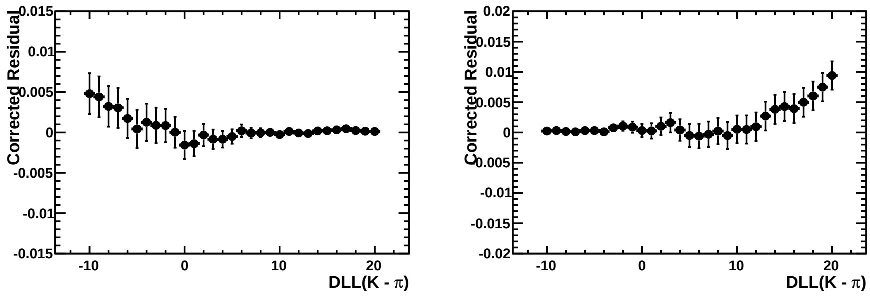
Clearly, the signal PID efficiency in data, $\varepsilon_{\rm sig}^{\rm data}$, is not known. However, providing the discrepancy between the "true" signal efficiency and the efficiency determined from the calibration method is small, then we can make the approximation

$$\left(\varepsilon_{\rm sig}^{\rm data} - \varepsilon_{\rm calib}^{\rm data}\right) \approx 2\, \varepsilon_{\rm calib}^{\rm data} \left(\frac{\varepsilon_{\rm sig}^{\rm MC} - \varepsilon_{\rm calib}^{\rm MC}}{\varepsilon_{\rm sig}^{\rm MC} + \varepsilon_{\rm calib}^{\rm MC}}\right) \; . \tag{4.14}$$

Various assumptions have been made when determining the formula to extract the absolute residuals. In order that the systematic uncertainty is not underestimated, residuals are determined over a range of PID cut values ($\{-10, -8, \ldots, 18, 20\}$) with the largest taken as the systematic. The results are shown in Fig. 4.19. From this, the systematic uncertainties associated with the PID calibration method for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ are taken to be 0.5% and 1.0% respectively.

The associated systematic on the physics observables is determined by performing the fit multiple times. Each time, the efficiencies of the PID cuts are randomly varied by a Gaussian PDF with mean equal to the PID efficiency and width equal to the systematic uncertainty.

The data are fitted 100 times in this way, and the *median absolute deviation* (MAD) defined



(a) (b)

Figure 4.19: Estimates of the residual $\varepsilon_{\rm sig} - \varepsilon_{\rm calib}$ in data over the range $\log(\mathcal{L}_K/\mathcal{L}_\pi) < \{-10, -8, \ldots, 18, 20\}$ for (a) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$, and $\log(\mathcal{L}_K/\mathcal{L}_\pi) > \{-10, -8, \ldots, 18, 20\}$ for (b) $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$, determined using a combination of calibration and signal data and MC samples. The calibration MC has been weighted according to the corresponding truth-matched signal MC.

below, of the central values of the observables of interest (*i.e.* $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$) returned from the fits is assigned as the systematic uncertainty associated with that observable.

The MAD is a method of quantifying the spread of a univariate dataset, and is defined by

$$\text{MAD} \equiv \text{median}_i\left(|x_i - \text{median}_j\left(x_j\right)|\right) \ , \tag{4.15}$$

where $\text{median}_i\left(x_i\right)$ represents the median of data points x_i. In summary, Eq. (4.15) indicates that the MAD is determined by calculating the median of the sample, from which the median of the absolute deviation of each data point from the sample median is determined. This quantity is more robust to outliers than the sample standard deviation. In order to use the MAD as an estimator for the population sample deviation ($\hat{\sigma}$) a scaling factor is applied, *i.e.* $\hat{\sigma} = K \cdot \text{MAD}$, where K depends on the underlying distribution. Assuming the data are sampled from a Gaussian distributed random variable, then $K \approx 1.4826$[1].

4.10.3 Signal PDF

The signal PDF parameters are allowed to float freely in the fit, so the errors returned by the fit comprise the statistical uncertainty of the sample added in quadrature with the

[1] The scaling factor K satisfies $\Phi(K^{-1}) - \Phi(-K^{-1}) = 1/2$, where $\Phi(x)$ is the cumulative density function of the standard Gaussian function, from which this approximate numerical solution can be derived.

uncertainty associated with the fit parameters. In order to isolate the latter contributions, the PDF parameters are fixed to the central values determined from the fit in which they are allowed to float, and the data are re-fitted. Since the two sets of results were taken from the same data sample, the systematic uncertainty associated with the uncertainties of the signal PDF parameters is given by the absolute quadrature difference between the errors from the two fits.

4.10.4 Mis-ID PDF

The systematic uncertainty associated with the mis-ID PDF parameters (other than the yield, which is accounted for by the PID systematic study) is determined by running the fit multiple times, each time randomly varying the PDF parameters within the uncertainties according to the initial fit in a similar way to that of the PID systematics, as outlined in Sec. 4.10.2. As with the PID systematics, the MAD is used to determine the systematic uncertainty. However, the situation is complicated by the fact that the fit variables are correlated with each other. The solution is to use the covariance matrix to generate a set of correlated random variables corresponding to the smeared PDF parameters. Such a method involves a two stage process.

The first stage requires that the covariance matrix of the parameters be calculated. Whilst this can be determined directly from the fit results, due to the limited sample size, the tail parameters (particularly n_{L} and α_{L} defined in Sec. 4.7) have large variations, which leads to problems in the generated fit values. Instead, n_{L} and α_{L} are fixed to the central values returned by the initial fit and the covariance matrix is calculated using the following procedure. First, the B mass distribution is binned and the yield in each bin randomly varied from its central value according to the initial fit within the binned statistical uncertainties. A double Crystal Ball PDF is then used to fit the smeared data, with the values of n_{L} and α_{L} fixed from the initial fit, as previously mentioned. This procedure is repeated many times, and the values of the fit parameters returned are used to construct the covariance matrix.

In the second stage, the covariance matrix is used to generate random variables for the fit parameters assuming a multivariate Gaussian distribution. If $\mathbf{x}$ represents the vector of central values of the parameters (excluding the fixed parameters n_{L} and α_{L}) and $\mathbf{y}$ a vector of standard Gaussian-distributed random variables (one for each fit parameter) then the vector of fit parameters varied within their correlated uncertainties, $\mathbf{z}$, is given by

$$\mathbf{z} = \mathbf{x} + \mathbf{C}\,\mathbf{y}\ , \tag{4.16}$$

where $\mathbf{C}$ is the Cholesky decomposition of the covariance matrix. The modified fit parameters $\mathbf{z}$, in conjuction with the fixed parameters n_{L} and α_{L}, are then used to perform the fit and the observables of interest are calculated. This process is repeated 100 times and the MAD of the generated observables is calculated.

4.10.5 Low mass background PDFs

The uncertainty associated with the low mass and semi-leptonic PDFs is determined by varying the lower limit of the fit range by $\pm 50\,\mathrm{MeV}/c^2$ and re-running the fit. The largest absolute difference between the central values obtained by these modified fit ranges and the central value from the standard fit is taken to be the systematic uncertainty.

4.10.6 Combinatoric background PDF

The uncertainty associated with the combinatoric background PDF is determined by varying the slope parameter by $\pm 1\sigma$ and re-running the fit. As with the low mass background PDF systematic discussed in Sec. 4.10.5, the systematic uncertainty is determined by the largest absolute difference between the two sets of central values.

4.10.7 $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ signal PDF scaling factors

Due to the lack of events containing $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, K^{\pm}$ decays and the significant contamination from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$, the signal PDF fit parameters (excluding the yield) cannot be determined from the fit. Instead, the fit parameters are determined by applying scaling factors to the parameters determined from the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\, \pi^{\pm}$ fit, with the scaling factors fixed from MC, as described in Sec. 4.7. The systematic uncertainty from this method is evaluated by varying each of the scaling factors by $\pm 1\sigma$ and re-running the fit. Again, the largest absolute difference between the two sets of central values is assigned as the systematic uncertainty.

4.10.8 B^+/B^- production and kaon detection asymmetries

The production and kaon detection asymmetries are fixed in the fit, using the results from the LHCb $B \to h^+h^-$ analysis [75], and combined $B^\pm \to (K^\pm\pi^\mp)_D\, h^\pm$ and $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, h^\pm$ results [3] respectively. A conservative systematic uncertainty is determined by propagating the combined statistical and systematic uncertainties of A_{prod} and A_{Kdet}, assuming 100% correlation between the variables, using the formula

$$\sigma_{A^{K3\pi}_{fav}} = \left(1 - \left(A^{K3\pi}_{fav}\right)^2\right)\left(\frac{\sigma_{A_{prod}}}{1 - A^2_{prod}} + 2\frac{\sigma_{A_{Kdet}}}{1 - A^2_{Kdet}}\right) , \tag{4.17}$$

where $\sigma_{A^{K3\pi}_{fav}}$ is the systematic uncertainty associated with the production and kaon detection asymmetries, and $\sigma_{A_{prod}}$ and $\sigma_{A_{Kdet}}$ represent the statistical and systematic uncertainties added in quadrature for A_{prod} and A^{det}_K respectively. The derivation of Eq. (4.17) can be found in Appendix B.

4.10.9 Choice of trigger lines

The analysis has been performed using data comprising all $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\, h^\pm$ candidates that pass the trigger, without discriminating between triggers designed to select hadronic events and those designed to select other decay modes. As explained in Sec. 4.5, this choice of trigger lines could potentially introduce biases in the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$ results.

To evaluate the systematic uncertainty associated with this decision, the default central values are compared with those obtained with a more specific trigger configuration. Namely, each event is required to pass at least one of the following categories at each stage:

L0 L0 hadron TOS **or** L0 global TIS.

HLT1 HLT1 1Track TOS **or** HLT1 global TIS.

HLT2 HLT2 topological (triggered by anything).

Here *global* refers to a trigger decision from any trigger line at the specified trigger stage.

For each physics observable, the absolute difference in central values with all triggers and with the restrictive trigger requirement is taken to be the systematic uncertainty associated with the choice of trigger lines.

4.11 Results

The central values for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$, along with the associated statistical uncertainties, are shown for a range of bachelor PID cuts in Tab. 4.18. The correction factors for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$, summarised in Sec. 4.10.1 have been applied. The statistical uncertainties quoted are those determined when the signal PDF parameters are fixed in the fit (see Sec. 4.10.3).

$\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut	Quantity	$A^{K3\pi}_{fav}$	$\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$
0	Central value	0.0533	0.0554
	Statistical uncertainty	0.1210	0.0069
2	Central value	-0.0142	0.0508
	Statistical uncertainty	0.1220	0.0064
4	Central value	0.0029	0.0545
	Statistical uncertainty	0.1130	0.0063
6	Central value	0.0080	0.0575
	Statistical uncertainty	0.1070	0.0063
8	Central value	-0.0028	0.0589
	Statistical uncertainty	0.1050	0.0064

Table 4.18: Central values and statistical uncertainties for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$ measurements, for various $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts.

4.11.1 Pull study

A *toy study* has been performed in which a series of approximately 10,000 MC datasets (toys) are produced. For each toy, the value of the B mass is randomly generated from the signal- and background-model PDFs used in the final fit, and the toys are then fitted to determine the parameters of interest. This study is used to determine the following:

1. the optimal $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut to use to determine the final fit results;

2. whether there are any biases in the fit parameters, and if so, correct for them.

Table 4.19 shows the average fitted value of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ for the PID cuts under investigation and the average fit error associated with the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ parameter. From Tab. 4.19, it can be seen that the mean fit error does not vary significantly with PID cut, except for the PID cut $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 4$, which intriguingly has a much lower fit error than the other cuts, albeit with a much larger uncertainty on the fit error. The mean value of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ for a PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 6$ is closest to the central value of 0.07 assumed when generating the toys, so this cut has been chosen to determine the final results.

$\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut	Mean value	Mean fit error
0	0.0722 ± 0.0228	0.0068 ± 0.0008
2	0.0713 ± 0.0191	0.0067 ± 0.0008
4	0.0711 ± 0.0221	0.0040 ± 0.0030
6	0.0704 ± 0.0206	0.0067 ± 0.0009
8	0.0710 ± 0.0200	0.0068 ± 0.0009

Table 4.19: Mean value and fit error for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ from $\sim 10,000$ toys for various $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts. The results for the preferred $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut are highlighted in bold.

For the chosen PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 6$ (for the pass slice), the pull distributions for the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ signal yield and parameters of interest are shown in Fig. 4.20. The results indicate that all three parameters agree well with the expected standard Gaussian distribution. The pull mean and width are also indicated in Tab. 4.20. This indicates that there is a slight but rather insignificant negative systematic shift of the pull mean. In addition, there is a systematic increase in the pull width away from unity, suggesting that the fit tends to underestimate the statistical fluctuation of the parameters. If this shift is larger than three times the uncertainty on the width, then a correction factor is applied to the measured statistical uncertainty of the parameters of interest, equal to the difference between the measured and expected width. Table 4.20 indicates that this is only necessary for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$, with a correction factor of 2%. The results with the inflated statistical uncertainties for the nominal $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut are shown in Tab. 4.21.

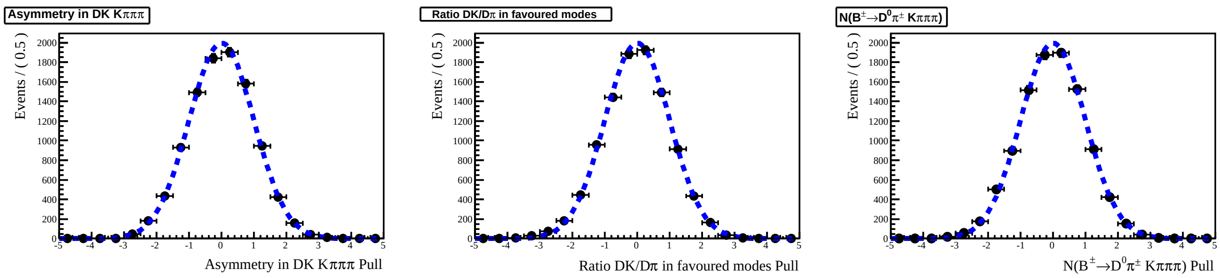


Figure 4.20: Pull distributions for the principal parameters floating in the fit. The dashed line represents the expected distribution of the pulls: a standard Gaussian distribution.

Quantity	μ	σ	Correction factor (%)
$N_{D\pi}$	-0.024 ± 0.010	1.018 ± 0.007	
$\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$	-0.022 ± 0.010	1.024 ± 0.008	2
$A_{fav}^{K3\pi}$	0.005 ± 0.010	1.007 ± 0.007	

Table 4.20: Pull mean (μ) and width (σ) for the principal parameters floating in the fit. Correction factors (if applicable) are shown in percent.

Quantity	$A_{fav}^{K3\pi}$	$\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$
Central value	0.0080	0.0575
Statistical uncertainty	0.1070	0.0065

Table 4.21: Central values and statistical uncertainties for $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A_{fav}^{K3\pi}$ measurements for the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ nominal cut. The statistical uncertainty for $A_{fav}^{K3\pi}$ has been inflated by the amount suggested by the pull studies.

4.11.2 Final results

Figure 4.21 shows the final fit to the B invariant-mass distribution in the pass and fail slices for the nominal $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut of 6, split by the charge of the B hadron.

A summary of the final results for a bachelor cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 6$, including the various sources of systematic uncertainty, are shown in Tab. 4.22.

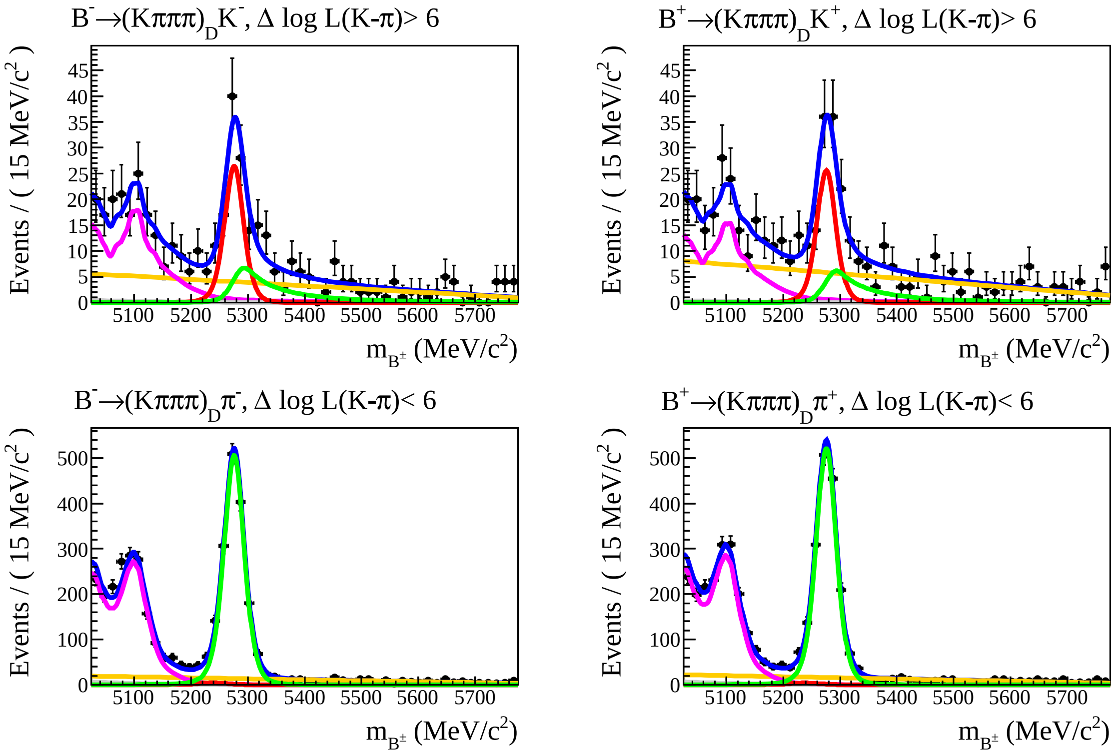


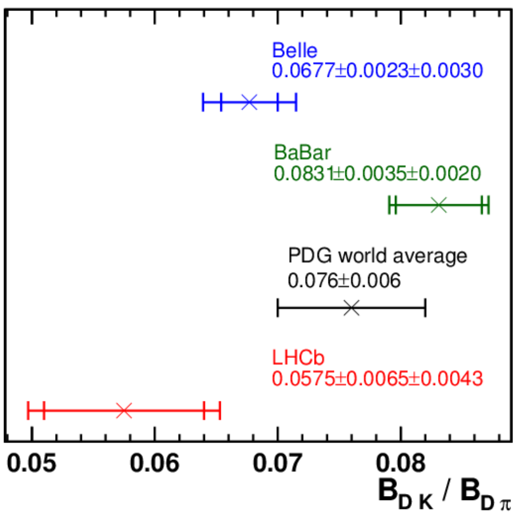
Figure 4.21: Final fit to $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ (red) and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ (green) data for $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts of > 6 and < 6 (pass and fail slice respectively). The data are further split by the charge of the B meson. Contributions from combinatoric (orange), semi-leptonic (grey) and partially reconstructed (magenta) backgrounds are also shown. The blue line indicates the sum of the signal and background fit components.

Quantity	$A_{fav}^{K3\pi}$	$\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$
Central value	0.0080	0.0575
Statistical uncertainty	0.1070	0.0065
Signal PDF systematic	0.0000	0.0005
PID systematic	0.0025	0.0021
Mis-ID PDF systematic	0.0046	0.0023
Low mass PDF systematic	0.0034	0.0001
MC scaling systematic	0.0001	0.0002
Combinatoric PDF systematic	0.0014	0.0003
A_{prod}/A_K^{det} systematic	0.0556	
Trigger choice systematic	0.0241	0.0030
$DK/D\pi$ correction systematic		0.0010
Total systematic	0.0610	0.0043
Total uncertainty	0.1232	0.0078

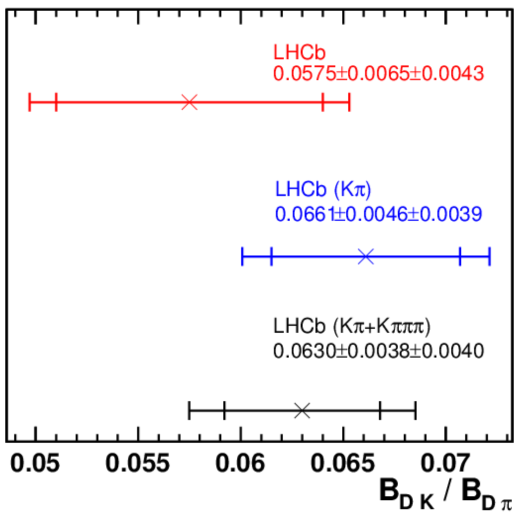
Table 4.22: Results for $A_{fav}^{K3\pi}$ and $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$, with statistical and systematic uncertainties.

4.11.3 Summary

Using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ candidates from $\mathcal{L}_{\text{int}} = 35.6$ pb^{-1} of LHCb collision data collected during 2010, the observables $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$ have been measured; the values are $\mathcal{B}_{DK}/\mathcal{B}_{D\pi} = 0.057 \pm 0.007 \pm 0.004$ and $A^{K3\pi}_{fav} = 0.01 \pm 0.11 \pm 0.06$. The latter measurement is consistent with the expected value of zero within uncertainties. The uncertainty on $A^{K3\pi}_{fav}$ is dominated by the statistical uncertainty and the uncertainty associated with the A_{prod} and A^{det}_{K} asymmetries. In a future measurement of $A^{K3\pi}_{fav}$ with more events, it should be possible to constrain one of these parameters in the fit and allow the other to float freely, significantly reducing the systematic uncertainty.



(a)



(b)

Figure 4.22: (a) comparison of the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement presented in this thesis (red) with results from BELLE (blue), BABAR (green) and the current PDG world average (black). (b) $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurements utilising $B^{\pm} \to (K^{\pm}\pi^{\mp})_D h^{\pm}$ decays (blue) and a combination of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D h^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp})_D h^{\pm}$ decays (black). In both plots, the statistical uncertainties (inner vertical lines) and the combined statistical and systematic uncertainties (outer vertical lines) are shown for each measurement.

Figure 4.22a shows a comparison of the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement presented here with measurements from the BELLE [69] and BABAR [70] experiments and the world average of $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ according to the PDG [23]. As can be seen, the result using LHCb data is consistent within uncertainties with the BELLE result, with just over 1σ deviation between the central values. There is a statistically significant deviation of over 3σ between the central values of the LHCb and BABAR results. The statistical uncertainty on the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ measurement at LHCb is competitive with the current PDG world average.

In Fig. 4.22b, the LHCb $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ result using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ decays is compared with LHCb measurements of the same observable from $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\,h^{\pm}$ decays, and from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\,h^{\pm}$ decays combined, using the same $\mathcal{L}_{\text{int}} = 35.6$ pb^{-1} data sample as the analysis presented in this thesis [3]. The measurement from $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ alone is consistent with the results from Ref. [3] within uncertainties, with less than 1σ deviation between the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ result and the result using both $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\,h^{\pm}$ decays.

Chapter 5

D^0 production cross-section using $D^0 \to K^-\pi^+\pi^-\pi^+$ decays at LHCb

The procedure to measure the D^0 production cross-section in bins of the D meson transverse momentum and rapidity over the ranges $[0, 8]$ GeV/c and $[2, 4.5]$ respectively using $D^0 \to K^-\pi^+\pi^-\pi^+$ decays from data collected by the LHCb experiment is discussed, and the results presented. The results are compared to two independent theoretical predictions and to a specific tuning of the Pythia event generator. Comparison is also made to the LHCb D^0 cross-section results using $D^0 \to K^-\pi^+$ decays.

5.1 Chapter outline

This chapter outlines the measurement of the D^0 production cross-section[1] at LHCb in bins of the D^0 p_{T} and rapidity, using Cabibbo favoured $D^0 \to K^-\pi^+\pi^-\pi^+$ decays[2].

The chapter begins with brief overviews of the analysis strategy, and the theory of the production of heavy-flavoured hadrons (*i.e.* a D or B hadron) in hadron-hadron collisions. The data samples analysed are then outlined, and the offline selection criteria are indicated. The procedure for determining the optimal selection criteria is also briefly discussed. The $D^0 \to K^-\pi^+\pi^-\pi^+$ amplitude model, which is used to re-weight the LHCb MC, is then

[1] Throughout this chapter, the average production cross-section of D^0 and $\overline{D}^0$ mesons is implied, *i.e.* $\frac{1}{2}\left(\sigma(D^0) + \sigma(\overline{D}^0)\right)$.
[2] Charge conjugate decays are also implied throughout this chapter unless otherwise specified.

outlined. The determination of the LHCb acceptance efficiencies and the reconstruction and selection efficiencies using MC are discussed, as are various sources of correction to these efficiencies. The calibration of the PID is then discussed, followed by an outline of the fit procedure and sources of correction to the yields extracted from the fit. A summary of the systematics uncertainties associated with the efficiencies, fit procedure, correction factors, integrated luminosity and branching ratio is given and the results presented. Finally, comparisons of the cross-section measurements to predictions from MC and theoretical models, and to a separate analysis using $D^0 \to K^-\pi^+$ decays are presented.

5.2 Introduction

D^0 mesons can be produced via several methods:

- *direct production* (produced directly in the pp collision);
- *feed-down* (from the instantaneous decay of an excited charm or charmonium resonance);
- *secondary production* (from the decay of a b-hadron).

The first two methods are collectively referred to as *prompt production.* This chapter describes the measurement of the prompt D^0 production cross-section using $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates from data accumulated at LHCb during the first few weeks of data-taking.

Since the D^0 production cross-section varies as a function of the p_{T} and rapidity[1] (y) of the D^0, the cross-section is determined in two-dimensional bins of these variables. To allow a meaningful comparison with theoretical expectations, p_{T} and y are computed in the centre-of-mass (CM) frame of the pp collision. LHCb's design means it is able to measure the production cross-section for charmed hadrons with low p_{T} and high y. This is a region that is inaccessible to many other experiments and is unique in terms of the LHC experiments. The bin scheme employed by this analysis comprises eight uniform bins of width $1\,\mathrm{GeV}/c$ in the range $0 < p_{\mathrm{T}} < 8\,\mathrm{GeV}/c$ and five uniform bins of width 0.5 in the range $2 < y < 4.5$.

[1] Rapidity is defined by $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$, where E is the energy of the particle, and p_z is the component of the momentum along the beam direction.

5.2.1 Analysis strategy

The D^0 production cross-section in a given D^0 (p_{T}, y) bin i, $\sigma_i\left(D^0\right)$, is determined from prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ decays by

$$\sigma_i\left(D^0\right) = \frac{N_i\left(D^0 \to K^-\pi^+\pi^-\pi^+\right)}{\varepsilon_{i,\mathrm{tot}} \cdot \mathcal{B}\left(D^0 \to K^-\pi^+\pi^-\pi^+\right) \cdot \mathcal{L}_{\mathrm{int}}} \,, \tag{5.1}$$

where $N_i\left(D^0 \to K^-\pi^+\pi^-\pi^+\right)$ is the yield of prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates in bin i, $\varepsilon_{i,\mathrm{tot}}$ is the *total efficiency* (the probability of a prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ candidate being retained after reconstruction, triggering and offline selection) in bin i, $\mathcal{B}\left(D^0 \to K^-\pi^+\pi^-\pi^+\right)$ is the branching ratio of $D^0 \to K^-\pi^+\pi^-\pi^+$ (including all intermediate resonances) and $\mathcal{L}_{\mathrm{int}}$ represents the integrated luminosity of the data sample analysed.

For each bin, $N_i\left(D^0 \to K^-\pi^+\pi^-\pi^+\right)$ is extracted using a simultaneous maximum likelihood fit to the D^0 invariant mass and $\log\left(IP\,\chi^2\right)$ distributions of the candidates. The latter distribution is used to distinguish between prompt and secondary candidates. The fit procedure is outlined in Sec. 5.8.

As with the measurements described in Chap. 4, it is necessary to apply a series of selection cuts in order to reduce the combinatoric background prior to performing the likelihood fit. The procedure used to determine the optimal cuts is discussed in Sec. 5.4.

Using $D^0 \to K^-\pi^+\pi^-\pi^+$ decays to measure the D^0 cross-section adds complexity to the analysis, since, as has been discussed in previous chapters, the D^0 can decay via various intermediate resonances. The LHCb MC models the resonance structure of the D^0 decays, as summarised in Tab. 4.1. However, this model neglects interference between the resonances. The $D^0 \to K^-\pi^+\pi^-\pi^+$ resonance structure has been studied by the MarkIII collaboration [76]. Section 5.5 describes how this $D^0 \to K^-\pi^+\pi^-\pi^+$ amplitude model is used to re-weight the LHCb MC.

The total efficiency in each (p_{T}, y) bin can be factorised into its constituent components as

$$\varepsilon_{i,\text{tot}} = \varepsilon_{i,\text{acc}} \cdot \varepsilon_{i,\text{trig}|\text{acc}} \cdot \varepsilon_{i,\text{sel}|\text{trig}} \cdot \varepsilon_{i,\text{PID}|\text{sel}} \ , \tag{5.2}$$

where:

- $\varepsilon_{i,\text{acc}}$ is the *acceptance efficiency*, indicating the probability that all the final-state tracks of the D^0 fall within the LHCb acceptance;

- $\varepsilon_{i,\text{trig}|\text{acc}}$ is the *trigger efficiency*, which corresponds to the probability that a $D^0 \to K^-\pi^+\pi^-\pi^+$ candidate with all final-state tracks within the LHCb acceptance will pass the trigger lines;

- $\varepsilon_{i,\text{sel}|\text{trig}}$ is the *selection and reconstruction efficiency*, which is the probability that a triggered $D^0 \to K^-\pi^+\pi^-\pi^+$ candidate will be fully reconstructed and pass all offline cuts (other than the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ requirements on the D^0 daughter tracks);

- $\varepsilon_{i,\text{PID}|\text{sel}}$ is the *PID efficiency*, which represents the probability of a $D^0 \to K^-\pi^+\pi^-\pi^+$ candidate that passes all other offline cuts also passing the track $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ selection criteria.

Note that the acceptance efficiency is *not* the same as the generator efficiency described in Chap. 4. The D^0 production cross-section is calculated with respect to D^0 mesons produced *within LHCb acceptance*, and so the D^0 mesons are required to have been generated within LHCb acceptance (in the lab frame) in the denominator of the efficiency calculation. The acceptance efficiency in this case corrects for D^0 mesons that are within acceptance, but are not reconstructible due to one or more D^0 daughter tracks being produced outside of the LHCb acceptance.

The choice of triggers employed within the analysis are discussed in Sec. 5.4.1. As will be explained in this section, the efficiencies of these triggers are uncorrelated with the other efficiencies and independent of the D^0 p_T and y, *i.e.* $\varepsilon_{i,\text{trig}|\text{acc}} \equiv \varepsilon_\text{trig}$. As a result, it is convenient to define an *effective integrated luminosity*, $\mathcal{L}_\text{eff} = \mathcal{L}_\text{int} \cdot \varepsilon_\text{trig}$. The value of $\mathcal{L}_\text{eff}$ is also given in Sec. 5.4.1.

The acceptance efficiencies and the reconstruction and selection efficiencies are determined using MC samples filtered using the same stripping and offline cuts as the collision data. These MC samples are summarised in Sec. 5.4.2. The methods used to determine the efficiencies are outlined in Sec. 5.6. In addition, various correction factors to these efficiencies are investigated, and these are also discussed in that section. Sources of correction factors to the yields are discussed in Secs. 5.8.2 and 5.8.3.

As with the $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$ measurements detailed in Chap. 4, the PID efficiencies are determined using a data-driven approach, since the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distributions are poorly described in the MC. However, there is a major difference for the cross-section analysis. Here, the PID efficiencies of all D^0 daughter tracks need to be evaluated, and the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts in the stripping prevent a determination of the efficiency of these cuts from the signal data samples. The PID calibration method is described in Sec. 5.7.

5.2.2 Physics motivation

As already mentioned, LHCb's design means it is able to measure the production cross-section in a region of phase space not accessible to many other experiments. Furthermore, it is expected that the charmed-hadron production cross-sections in the y and p_{T} regions covered by this analysis will have sensitivity to enhancements of the charm parton density function as a result of non-perturbative effects, referred to as *intrinsic charm* [77]. Thus, the measurement of the production cross-section of charmed hadrons at LHCb provides important experimental tests of theoretical QCD models, as discussed in Sec. 5.3. In addition, these cross-sections are used as inputs to *CP* violation and mixing measurements, and to estimate the sensitivity of LHCb to rare decays of charmed hadrons.

The simplest measurement of the D^0 production cross-section at LHCb comes from $D^0 \to K^-\pi^+$ decays. However, LHCb also allows the reconstruction of multi-body D^0 decays such as $D^0 \to K^-\pi^+\pi^-\pi^+$, as was demonstrated in Chap. 4, providing an additional independent method of calculating the D^0 production cross-section. Agreement between the two analyses would also indicate that the various sources of uncertainty are well understood, particularly those correlated between the analyses, such as the uncertainty due to the long-track reconstruction efficiency.

5.3 Heavy flavour production in hadron-hadron collisions

This section describes how the hard-scattering cross-section can be determined for heavy quarks, how this information can be used to predict the cross-section for heavy flavoured hadrons in hadron-hadron collisions, and introduces the theoretical models of heavy flavour production.

5.3.1 Heavy quark cross-section

The cross-section of a hard-scattering process from the collision of two hadrons with four-momenta P_1 and P_2 is given by

$$\sigma\left(P_1, P_2, \mu\right) = \sum_{i,j} \int dx_1 \int dx_2 \, f_i\left(x_1, \mu\right)^2 f_j\left(x_2, \mu\right)^2 \hat{\sigma}_{ij}\left(p_1, p_2, \alpha_S\left(\mu^2\right)\right) \ , \tag{5.3}$$

where x_1 and x_2 are the fraction of the four-momenta of the hadrons carried by the two *partons* (quarks or gluons) that participate in the hard-scattering interaction, $f_i\left(x, \mu\right)$ is the *parton density function* of parton i (the probability density of an initial-state parton i carrying momentum fraction x) measured at the mass scale μ, $\alpha_S\left(\mu^2\right)$ is the coupling constant of the strong force at mass scale μ and $\hat{\sigma}_{ij}$ is the short-distance cross-section for the scattering of partons i and j. The characteristic mass scale of the hard-scattering process, μ, depends on the process in question.

For the production of a heavy quark Q, the *leading-order* (LO) processes are $q\bar{q}$ scattering by light quarks ($q\bar{q} \to Q\bar{Q}$) and *gluon-gluon fusion* ($gg \to Q\bar{Q}$). Assuming $m_q \ll m_Q$, where m_q and m_Q are the masses of the light and heavy quarks respectively, the short-distance cross-section can be written to LO in terms of the invariant matrix element as

$$d\hat{\sigma}_{ij}\left(p_1,p_2,p_3,p_4\right) = \frac{1}{16\pi^2\hat{s}}\frac{d^3p_3}{E_3}\frac{d^3p_4}{E_4}\delta^4\left(p_1+p_2-p_3-p_4\right)\overline{\sum}\left|\mathcal{M}_{ij}\right|^2 \ , \tag{5.4}$$

where p_1 and p_2 are the four-momenta of the initial-state partons, p_3 and p_4 are the four-momentum of the Q and $\bar{Q}$, with energies E_3 and E_4 respectively, $\hat{s} = (p_1+p_2)^2$ and $\overline{\sum}$ indicates that the invariant matrix elements squared are averaged (summed) over the initial (final) colours and spins.

5.3.2 Theoretical models of heavy flavour production

The cross-section for the production of a heavy quark Q with four-momentum P being produced and hadronising into to a heavy hadron H_Q with four-momentum $p = zP$ in a hadron-hadron collision at a CM energy $\sqrt{s}$ is given by

$$\sigma_{hh\to H_Q+X}\left(s,p,\mu\right) = \sum_{i,j}\int dx_1\int dx_2\int dz\, f_i\left(x_1,\mu\right)^2 f_j\left(x_2,\mu\right)^2\hat{\sigma}_{ij}\left(s,m_Q^2,\mu^2\right)D_Q^{H_Q}\left(z,\mu^2\right) \ , \tag{5.5}$$

where $D_Q^{H_Q}\left(z,\mu^2\right)$ is the *fragmentation function* for parton Q and hadron H_Q (the probability density for Q to hadronise to H_Q, carrying a fraction z of the quark's momentum).

The short-distance cross-section, $\hat{\sigma}_{ij}\left(s,m_Q^2,\mu^2\right)$ can be written in a general form as

$$\hat{\sigma}_{ij}\left(s,m_Q^2,\mu^2\right) = \frac{\alpha_S\left(\mu^2\right)}{m_Q^2}\mathcal{F}_{ij}\left(\rho,\mu^2,m_Q^2\right) \ , \tag{5.6}$$

where $\rho = \frac{4m_Q^2}{\hat{s}}$ and, as before, $\hat{s}$ is the CM energy of the initial-state partons. The dimensionless functions $\mathcal{F}_{ij}$ are given by the perturbative expansion

$$\mathcal{F}_{ij}\left(\rho,\mu^2,m_Q^2\right) = \mathcal{F}_{ij}^{(0)}(\rho) + 4\pi\alpha_S\left(\mu^2\right)\left(\mathcal{F}_{ij}^{(1)}(\rho) + \overline{\mathcal{F}}_{ij}^{(1)}(\rho)\ln\left(\frac{\mu^2}{m_Q^2}\right)\right) + \mathcal{O}\left(\alpha_S\left(\mu^2\right)^2\right) \ . \tag{5.7}$$

where the functions $\mathcal{F}_{ij}^{(0)}(\rho)$ corresponds to the LO processes, and the functions $\mathcal{F}_{ij}^{(1)}(\rho)$ and $\overline{\mathcal{F}}_{ij}^{(1)}(\rho)$ to *next-to-leading order* (NLO). Similar functions exist at higher orders. From

Eq. (5.7), it can be seen that if $\mu = m_Q$, *i.e.* the heavy quark mass is the characteristic scale of the hard-scattering process, then the $\overline{\mathcal{F}}_{ij}$ terms cancel, and the short-distance cross-section is calculable using perturbation theory. However, there is a second scale, namely the p_{T} of the heavy quark/hadron.

If $p_{\mathrm{T}} \lesssim m_Q$, then the gluons and light quarks (*i.e.* the quarks lighter than the heavy quark Q) can be treated as the only *active* partons (partons that contribute to the short-distance cross-section) and since m_Q is the characteristic scale in this case, the logarithmic terms can be ignored. In this scheme, the heavy quark only appears in the final-state, *i.e.* in the fragmentation function.

If $p_{\mathrm{T}} \gg m_Q$, then p_{T} is the characteristic scale of the hard-scattering process. In this case the logarithmic terms, shown to NLO in Eq. (5.7), become large and spoil the convergence of the perturbative expansion. These logarithmic terms then need to be *resummed* by absorbing them into the experimentally determined parton density functions and fragmentation functions. In this approach, the heavy quark is treated as an active parton, so it needs to be considered in the short-distance cross-section calculation. Since the heavy quark is highly relativistic in this kinematic region, it can be treated as massless in the calculation of the short-distance cross-section.

There are various methods for interpolating between the two schemes described above to construct a procedure for determining the cross-section that is valid over a wide range of p_{T} values. Two such schemes are considered in this thesis, namely the *General-Mass Variable-Flavour-Number Scheme* (GM-VFNS) [77, 78, 79, 80, 81, 82] and the *Fixed-Order with Next-to-Leading Logarithms* (FONNL) approach [83, 84].

The GM-VFNS approach accounts for the NLO terms in the $p_{\mathrm{T}} \lesssim m_Q$ scheme and the *leading-logarithm* terms of the $p_{\mathrm{T}} \gg m_Q$ scheme (*i.e.* all terms proportional to $\ln\left(\mu^2/m_Q^2\right)$ in the full expansion of Eq. (5.7)), preserving the m_Q-dependence of the cross-section at all energy scales, and resumming terms with mass singularities where appropriate.

In the FONNL approach, the NLO terms in the $p_{\mathrm{T}} \lesssim m_Q$ scheme and the *next-to-*

leading-logarithm terms of the $p_{\mathrm{T}} \gg m_Q$ scheme (*i.e.* terms proportional to $\ln\left(\mu^2/m_Q^2\right)$ and $\ln^2\left(\mu^2/m_Q^2\right)$) are considered. The cross-sections of the $p_{\mathrm{T}} \lesssim m_Q$ and $p_{\mathrm{T}} \gg m_Q$ schemes are linear combined in this approach, weighting the cross-section of the latter scheme by a coefficient function of the form $p_{\mathrm{T}}^2/\left(p_{\mathrm{T}}^2 + 25m_Q^2\right)$ to suppress the $p_{\mathrm{T}} \lesssim m_Q$ scheme at low p_{T}.

5.4 Selection

The main source of background for $D^0 \to K^-\pi^+\pi^-\pi^+$ decays comes from the random combination of kaon and pion tracks to form a fake vertex. Various cuts are applied in the stripping selection to reduce this source of background.

The proportion of candidates that originate from secondary decays are estimated by performing a maximum likelihood fit to the (base-10) logarithm of the *IP* χ^2 of the D^0 meson. As such, the stripping and offline cuts have been chosen to avoid tight cuts on variables correlated with this variable.

The offline cuts are chosen to maximise the figure of merit $S/\sqrt{S+B}$, where S and B are the expected number of associated signal and background events respectively, normalised to $14\,\mathrm{nb}^{-1}$; an estimate of the total integrated luminosity of the data samples. The offline selection cuts are optimised using a *genetic algorithm*. The stripping and offline cuts are summarised in Sec. 5.4.4. The samples used to perform the optimisation are discussed in Secs. 5.4.1 and 5.4.2, whilst the optimisation procedure is outlined in Sec. 5.4.3.

5.4.1 Data samples

The pp collision data samples analysed comprise an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 14.96 \pm 0.52\ \mathrm{nb}^{-1}$. These samples correspond to a dedicated reprocessing of the data[1], utilising stripping lines that were specially configured for the open charm cross-section measurements. These data correspond to a period of LHCb data-taking in which the *pile-up* was low.

[1] Reconstruction version *Reco07*.

Pile-up is the average number of pp interactions per event. Since the L0 trigger system can only trigger on visible interactions, the pile-up is by definition ≥ 1. Pile-up is one of the parameters that affects the HLT processing time. Assuming a *minimum-bias* L0, *i.e.* the L0 global trigger decision accepts all (visible) events, then the pile-up indicates the average size of the events passed to the EFF: the larger the pile-up, the more restrictive the HLT triggers need to be to reduce the event rate to the required $2-3\,\mathrm{kHz}$ rate, as discussed in Sec. 2.8.

In general, to maintain a $1\,\mathrm{MHz}$ output rate, the L0 cannot accept all events. However, the data used by this analysis correspond to a pile-up of $1.04-1.08$[1], which is low enough to utilise a minimum-bias L0 trigger. The low pile-up also means that this analysis is able to employ *micro-bias* HLT trigger lines, which simply accept events with an observable track within:

- the **VELO R-sensors**, referred to as the *Hlt1MBMicroBiasRZVelo* trigger line;
- the **T-stations**, referred to as the *Hlt1MBMicroBiasTStation* trigger line.

This L0 and HLT trigger configuration significantly simplifies the analysis at the expense of limiting the amount of data which can be analysed. However, since the triggers were far more stringent during the rest of 2010 as a result of the high pile-up, and tended to prioritise the triggering of B decays, these data still contain a significant proportion of the total prompt D^0 decays accumulated by LHCb in this period.

The first $\mathcal{L}_{\mathrm{int}} = 1.87 \pm 0.07\ \mathrm{nb}^{-1}$ of data were taken with a Trigger Configuration Key (TCK) for which the micro-bias trigger lines accepted every event with a VELO R-sensor or T-station track segment[2], whilst the TCK used to collect the remaining $\mathcal{L}_{\mathrm{int}} = 13.09 \pm 0.46\ \mathrm{nb}^{-1}$ of data used a combination of prescaled and rate-limited lines to maintain the required trigger retention-rate[3].

The signal candidates used in the cross-section analysis are required to pass the logical

[1] The average number of *visible* pp interactions per bunch-crossing, μ, of these data lie in the range $0.074 \leq \mu \leq 0.160$ [85], from which the pile-up has been inferred.

[2] TCK `0x00051710`.

[3] TCK `0x00081710`.

OR of the micro-bias trigger lines. The micro-bias trigger algorithms are assumed to have a 100% efficiency for selecting signal candidates that decay within the LHCb acceptance, so no trigger requirements are applied to the MC samples.

The effective trigger efficiency of the combination of prescaled and rate-limited lines is $\varepsilon_{\text{trig}} = 0.2399 \pm 0.0019$. Consequently, the total effective integrated luminosity from the data spanning the two TCKs is $\mathcal{L}_{\text{eff}} = (5.011 \pm 0.177)\,\text{nb}^{-1}$. The uncertainty on this luminosity measurement is assigned as a source of systematic uncertainty on the cross-section.

The stripping cuts used to select the $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates are summarised in Sec. 5.4.4. Data collected with both magnet polarities are used in the analysis. Due to the limited number of events, no attempt has been made to separate the data by field polarity or D meson type to look for a detector left-right asymmetry or D^0-$\overline{D}^0$ production asymmetry, respectively.

5.4.2 MC samples

The reconstruction and selection efficiencies have been determined using *signal MC* (*i.e.* MC events containing a $D^0 \to K^-\pi^+\pi^-\pi^+$ decay). These MC samples were reconstructed with a slightly later version of the reconstruction software[1] than used to reconstruct the data. However, the difference between these reconstruction versions is minimal; the later version contains minor bug fixes that should not affect the event reconstruction.

The MC samples were generated with $\nu = 1$, where ν is the average number of pp collisions per bunch crossing. This quantity includes *all* pp interactions, not only those that are visible in the detector, so it is not a useful measure of the interaction rate in data. However, assuming that 69.9% of interactions are visible[2], this quantity can be used to infer the pile-up. In this case, the ν value of the MC corresponds to a pile-up value of 1.39. This pile-up is significantly larger than the pile-up in the data (see Sec. 5.4.1). Since the track reconstruction efficiency is known to vary significantly with the number of tracks in the event (and consequently the

[1] Reconstruction version *Reco08*.

[2] This assumption is based on the L0 minimum-bias accept rate of single-interaction events in 2010 MC [86].

number of interactions), it is preferable to have a similar pile-up in the data and MC samples. Thus, only MC events with precisely one generated pp interaction are considered.

Inclusive $c\bar{c}$ MC is used to investigate peaking backgrounds, as discussed in Sec. 5.8.3. These samples were produced using the same reconstruction version as the signal MC, but were generated with $\nu = 2.5$ rather than $\nu = 1$ to account for the high pile-up in data collected during the late 2010 data-taking period. Nevertheless, as with the signal MC, events are only considered if they have only one generated pp collision, so this difference is assumed to be negligible.

The offline cut optimisation is performed with signal and inclusive $c\bar{c}$ MC, produced prior to the first pp collisions. These MC samples were produced with much earlier reconstruction version[1] than the data, and are consequently not used at any other stage of the analysis. These samples were generated with $\nu = 1$; however, the requirement of only one generated pp collision was not applied during the cut optimisation.

A full list of MC samples used can be found in Tab. 5.1.

5.4.3 Offline cut optimisation

As previously mentioned, the offline cuts have been optimised using a genetic algorithm. A brief description of the optimisation procedure is given here.

First, a list of cut variables, referred to as *genes*, are defined, with the range of possible cut values chosen to minimise the impact on the D^0 p_{T} and $\log\left(IP\,\chi^2\right)$ distributions of the signal candidates. Each variable is then assigned a random value. This procedure is repeated to give a *population* of twenty samples of cut variables, each with randomly chosen values. The population is randomly split into ten pairs, each of which is made to *reproduce* twice, so that each of the *offspring* samples contains genes from the parent sample. The genes have a 50% chance of being inherited from each parent, and all offspring genes have a 10% probability of *mutating* to a new random value. Once the population has doubled in size, for each member of the population, the cuts are applied to the MC candidates, and $S/\sqrt{S+B}$ is calculated

[1] Reconstruction version *Reco03*.

Sample type	Magnet polarity	Events ($\times 10^6$)	ν	Pile-up	Acc. cuts?	Cut opt.?
$D^0 \to K^-\pi^+\pi^-\pi^+$	Down	1.8	1	1.39	Yes	Yes
$D^0 \to K^-\pi^+\pi^-\pi^+$	Up	1.7	1	1.39	Yes	Yes
Inclusive $c\bar{c}$	Down	4.2	1	1.39	Yes	Yes
Inclusive $c\bar{c}$	Up	3.7	1	1.39	Yes	Yes
$D^0 \to K^-\pi^+\pi^-\pi^+$	Down	5.1	1	1.39	Yes	No
$D^0 \to K^-\pi^+\pi^-\pi^+$	Up	5.1	1	1.39	Yes	No
Inclusive $c\bar{c}$	Down	2.5	2.5	2.72	Yes	No
Inclusive $c\bar{c}$	Up	2.5	2.5	2.72	Yes	No
$D^0 \to K^-\pi^+\pi^-\pi^+$	Down	0.4	1	1.39	No	No
$D^0 \to K^-\pi^+\pi^-\pi^+$	Up	0.4	1	1.39	No	No

Table 5.1: List of MC samples used in the D^0 production cross-section measurement. The magnet polarity, number of events, average number of pp interactions (ν) and pile-up are indicated. In addition, the penultimate column indicates whether generator-level acceptance cuts have been applied. The final column indicates whether the samples are used to perform the optimisation of the offline selection cuts. For samples with acceptance cuts applied, the "Events" column indicates the number of events retained *after* applying the generator-level cuts.

from the signal and $c\bar{c}$ MC samples. The worst performing half of the population is removed, and the process repeated until the cuts converge on a maximum value of $S/\sqrt{S+B}$.

5.4.4 Stripping and offline selection cuts

The full list of cuts applied in the stripping and offline selection can be found in Tab. 5.2. In addition to the optimised offline cuts, all "clone-like" candidates are rejected by applying clone-distance cuts, as defined in Sec. 2.3.4, to the D^0 daughter tracks. Furthermore, a cut of D^0 $\ln(VD\,\chi^2) < 10$ has been applied, since candidates in the rejected region of this cut are almost exclusively background according to the MC, and the cut significantly improves the quality of the $\log(IP\,\chi^2)$ fit. Furthermore, additional cuts are applied as a result of the PID performance study, as detailed in Sec. 5.7.

Variable name	Stripping	Offline
Mass window (MeV/c^2)	75	
D^0 $VD\,\chi^2 >$	16	
D^0 $\ln(VD\,\chi^2) <$		10
Signed VD (mm) $>$	0	
τ (ps) $>$	0	
D^0 $DIRA >$		0.99979
D^0 vertex $\chi^2/N_{\mathrm{dof}} <$	20	3
D^0 max($DoCA$) (mm) $<$	0.5	0.21
h sIP $\chi^2 >$	1	1.5, 6, 9, 16
h p_{T} (MeV/c) $>$	250	250, 300, 350, 400
h $\lvert\vec{p}\rvert$ (GeV/c) $>$	3	
h track $\chi^2/N_{\mathrm{dof}} <$	5	
h clone distance $<$		0
$K^{\pm}$ $\log(\mathcal{L}_K/\mathcal{L}_\pi) >$	0	5
$\pi^{\pm}$ $\log(\mathcal{L}_K/\mathcal{L}_\pi) <$	5	

Table 5.2: Stripping and offline selection criteria for $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates. h denotes a pion or kaon product of a $D^0 \to K^-\pi^+\pi^-\pi^+$ decay. Where multiple values are listed for selection criteria, all of the final-state particles are required to pass the least restrictive requirement, at least three of the four are required to pass the second least restrictive requirement, at least two of the four are required to pass the second most restrictive requirement and one of the four is required to pass the most restrictive requirement. The *Stripping* column indicates the selection criteria used in the stripping lines. The *Offline* column indicates the selection criteria applied after the stripping.

5.5 $D^0 \to K^-\pi^+\pi^-\pi^+$ amplitude model

The decay model used to generate the $D^0 \to K^-\pi^+\pi^-\pi^+$ decays in the MC is expected to give a poor description of the true decay kinematics, as it incoherently sums the decay amplitudes of the intermediate resonances. This could potentially lead to inaccuracies in the calculated efficiencies. Consequently, the MC is re-weighted according to the amplitude model for (Cabibbo favoured) $D^0 \to K^-\pi^+\pi^-\pi^+$ decays published by the MarkIII collaboration [76], and the re-weighted MC is used to determine the MC-derived efficiencies. A short description of the formalism of four-body amplitude models and its implementation follows.

5.5.1 Formalism of four-body amplitude models

The square of the CM energy between two particles i and j, s_{ij}, is defined by

$$s_{ij} = (p_i + p_j)^2 \equiv m_i^2 + m_j^2 + 2\, p_i \cdot p_j \ , \tag{5.8}$$

where p_i and p_j are the four-momenta of the two particles and m_i and m_j are their corresponding invariant masses.

As a result of energy and momentum conservation, the kinematics of three-body decays such as $D^+ \to K^-\pi^+\pi^+$ can be fully described by two of the three possible Lorentz-invariant s_{ij} parameters of the daughters. The scatter plot of the distribution of decay candidates in two of these variables is known as a *Dalitz plot* [87]. Dalitz plots are a useful way of investigating the resonance structure of three body decays. Since the three-body phase space is flat in the Dalitz plot variables, any variation in the distribution of events in the Dalitz plot is a result of variations in the matrix element rather than the kinematics of the event[1]. Consequently, high density regions in the Dalitz plot indicate resonances, and the distribution of the events within these high density regions can be used to determine the properties of the resonance (*i.e.* its mass, decay width and charge). Furthermore, given a particular model for the dynamics and angular distributions of the resonances, the full amplitude structure of the

[1] After correcting for the variation of the acceptance, reconstruction and selection efficiencies over the Dalitz plot.

decay can be extracted.

The Dalitz-plot technique can be extended to four-body final states such as $D^0 \to K^-\pi^+\pi^-\pi^+$. However, in this case, five Lorentz-invariant variables are needed to fully describe the decay kinematics. A common choice of variables is t_{01}, s_{12}, s_{23}, s_{34} and t_{40}, where index 0 denotes the D^0 and t_{ij} is defined by

$$t_{ij} = (p_i - p_j)^2 \equiv m_i^2 + m_j^2 - 2\, p_i \cdot p_j \ . \tag{5.9}$$

The four-body phase space density, ϕ_4, is not flat in five dimensions, but rather proportional to the inverse square-root of the *Gram determinant*:

$$\phi_4 (\mathbf{x}) \equiv \frac{\Phi_4}{dt_{01} ds_{12} ds_{23} ds_{34} dt_{40}} = \frac{\pi^2}{32 m_0^2} \left(- \begin{vmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{vmatrix} \right)^{-\frac{1}{2}} , \tag{5.10}$$

where Φ_4 is the integrated four-body phase space and x_{ij} is the inner product between the four-momenta of daughter tracks i and j.

The total decay amplitude at point $\mathbf{x}$, $A(\mathbf{x})$, can be parametrised by a complex sum over the individual decay amplitudes, *i.e.*

$$A(\mathbf{x}) = \sum_i a_i e^{i\phi_i} \mathcal{A}_i (\mathbf{x}) \ , \tag{5.11}$$

where the decay amplitude function, $\mathcal{A}_i(\mathbf{x})$, is a complex function that parametrises the decay amplitude of contribution i. The complex coefficients $a_i e^{i\phi_i}$ indicate the level of interference between the resonances (and the non-resonant contribution). The magnitudes, a_i, and phases, ϕ_i, of the coefficients are inputs to the amplitude model. The decay amplitude functions are a product of the modelled dynamical functions and *spin factors* (which represent the angular distributions of the final-state particles). Consequently, these functions depend on the model conventions, and the *fit fractions*, $\mathcal{F}_i$, are often quoted by analysts, since they are convention-independent. For four-body decays, the fit fractions are defined by:

$$\mathcal{F}_i = \frac{\int |\mathcal{A}_i(\mathbf{x})|^2 \, d\Phi_4}{\int |A(\mathbf{x})|^2 \, d\Phi_4} \ . \tag{5.12}$$

5.5.2 Implementation of the MarkIII amplitude model

The *Mint-Dalitz* software was developed by Jonas Rademacker (University of Bristol) to determine the amplitudes and phases of multi-body charm decays [88].

A large number of different conventions are used in the implementation of amplitude models. Furthermore, the conventions that have been used are often undocumented, making it very difficult for other analysts to recreate the amplitude model. This is the case for the model from the MarkIII collaboration. A reasonable approximation to the MarkIII model has been implemented by generating a series of toy amplitude models using the Mint-Dalitz software in which the phases of the decay amplitudes, ϕ_i, are fixed to those given in the MarkIII paper [76], and the magnitudes, a_i, are randomly varied. The fit fractions are then determined, and a fit is performed to determine the amplitude model that minimises the difference between the fit fractions according to Mint-Dalitz and the MarkIII paper. The final $D^0 \to K^-\pi^+\pi^-\pi^+$ amplitude model according to Mint-Dalitz is shown in Tab. 5.3. For each resonance, the lowest possible orbital angular momentum state that conserves angular momentum has been considered by the model. The exception is $D^0 \to \overline{K}^{*0}(892)\rho^0(770)$, for which the D-wave state has also been included in the model.

The differences between the fit fractions according to the Mint-Dalitz model and those quoted by the MarkIII collaboration are generally smaller than the statistical uncertainties on the fractions quoted in the MarkIII paper, indicating that the Mint-Dalitz model provides a reasonable approximation to the MarkIII model. Nevertheless, conservative systematic uncertainties arising from the model uncertainty have been assigned, and the method for determining these uncertainties is discussed in Sec. 5.9.2.

In order to use the Mint-Dalitz model to re-weight the MC, it is also necessary to

Decay chain	$\mathcal{F}_i$	a_i	ϕ_i
non-resonant	0.22	0.97	2.07
$D^0 \to (a^+(1260) \to \rho^0(770)\pi^+)\, K^-$	0.47	1.0	0.0
$D^0 \to \left(K_1^-(1270) \to \rho^0(770)K^-\right)\pi^+$	0.032	0.13	0.71
$D^0 \to \left(K_1^-(1270) \to \overline{K}^{*0}(1430)\pi^-\right)\pi^+$	0.021	0.57	3.85
$D^0 \to \left(K_1^-(1270) \to \overline{K}^{*0}(892)\pi^-\right)\pi^+$	0.013	0.040	0.1
$D^0 \to \overline{K}^{*0}(892)\,\{\pi^+\pi^-\}_S$	0.13	0.93	6.21
$D^0 \to \rho^0(770)\,\{K^-\pi^+\}_S$	0.088	0.5464	2.84
$D^0 \to \overline{K}^{*0}(892)\rho^0(770)$	0.29	0.1232	1.69
$D^0 \to \overline{K}^{*0}(892)\rho^0(770)$ (D-wave)	0.13	0.1963	1.96

Table 5.3: Amplitude model of $D^0 \to K^-\pi^+\pi^-\pi^+$ according to the Mint-Dalitz software's implementation of the MarkIII model. The fit fractions ($\mathcal{F}_i$), magnitudes (a_i) and phases (ϕ_i) of the decay amplitudes are shown. $\overline{K}^{*0}(892)$ and $\rho^0(770)$ are assumed to exclusively decay to $K^-\pi^+$ and $\pi^+\pi^-$ respectively. The phases have been calculated with respect to $D^0 \to \left(K_1^-(1270) \to \rho^0(770)K^-\right)\pi^+$.

implement the amplitude model of the un-weighted MC in order to reproduce the four-momentum distributions of the D^0 and its daughters. The results of a Mint-Dalitz fit to the LHCb MC can be found in Tab. 5.4.

Whilst the LHCb MC simulation treats the $D^0 \to K^-\pi^+\pi^-\pi^+$ decay as a incoherent sum of decay amplitudes, some of these amplitudes are in fact coherent sums over different angular-momentum states. In these cases, the orbital angular momentum state is indicated in parentheses in Tab. 5.4 and (if applicable) the state of the two-body non-resonant contribution is indicated by the subscript after the curly brackets. In addition, the phases for these amplitudes are indicated, which are calculated with respect to the spin-0 (S-wave) contribution.

Some of the amplitudes included in the fit are unphysical, in that they ignore the angular momentum of the particles. In this case, the spin factors are set to unity in the decay amplitude function, and are identified in Tab. 5.4 by "(no spin)". To improve the quality of the fit, some decays have been included in both the coherent sum *and* as unphysical spin-independent amplitudes; hence, these amplitudes appear twice in Tab. 5.4.

Decay chain	$\mathcal{F}_i$	a_i	ϕ_i
non-resonant	0.40	1.53	
$D^0 \to (a^+(1260) \to \rho^0(770)\pi^+)\, K^-$	0.33	1.0	
$D^0 \to \left(K_1^-(1270) \to \rho^0(770)K^-\right)\pi^+$	$5.1e^{-3}$	0.063	
$D^0 \to \left(K_1^-(1270) \to \overline{K}^{*0}(892)\pi^-\right)\pi^+$	$7.1e^{-3}$	0.035	
$D^0 \to \left(K_1^-(1270) \to \{\pi^+\pi^-\}_P\, K^-\right)\pi^+$	$3.4e^{-3}$	0.048	
$D^0 \to \overline{K}^{*0}(892)\,\{\pi^+\pi^-\}_S$ (no-spin)	0.063	0.23	
$D^0 \to \rho^0(770)\,\{K^-\pi^+\}_S$ (no-spin)	$4.3e^{-3}$	0.052	
$D^0 \to \omega^0(782)\,\{K^-\pi^+\}_S$ (no-spin)	0.012	0.019	
$D^0 \to \overline{K}^{*0}(892)\rho^0(770)$ (S-wave)	0.17	0.11	0.0
$D^0 \to \overline{K}^{*0}(892)\rho^0(770)$ (P-wave)	$4.9e^{-4}$	$6.5e^{-3}$	-2.6
$D^0 \to \overline{K}^{*0}(892)\rho^0(770)$ (D-wave)	0.21	0.29	-2.8
$D^0 \to \overline{K}^{*0}(892)\omega^0(782)$ (S-wave)	$4.3e^{-3}$	$4.0e^{-3}$	0.0
$D^0 \to \overline{K}^{*0}(892)\omega^0(782)$ (P-wave)	$1.6e^{-4}$	$7.8e^{-4}$	-9.5
$D^0 \to \overline{K}^{*0}(892)\omega^0(782)$ (D-wave)	0.010	0.015	3.1
$D^0 \to \rho^0(770)\,\{K^-\pi^+\}_S$	$6.2e^{-3}$	0.17	0.0
$D^0 \to \rho^0(770)\,\{K^-\pi^+\}_P$ (S-wave)	0.010	0.25	1.9
$D^0 \to \rho^0(770)\,\{K^-\pi^+\}_P$ (P-wave)	0.011	0.49	2.9
$D^0 \to \rho^0(770)\,\{K^-\pi^+\}_P$ (D-wave)	$5.1e^{-3}$	0.63	-0.23
$D^0 \to \overline{K}^{*0}(892)\,\{\pi^+\pi^-\}_S$	0.013	0.35	0.0
$D^0 \to \overline{K}^{*0}(892)\,\{\pi^+\pi^-\}_P$ (S-wave)	0.012	0.16	-0.34
$D^0 \to \overline{K}^{*0}(892)\,\{\pi^+\pi^-\}_P$ (P-wave)	$4.2e^{-3}$	0.27	-0.47
$D^0 \to \overline{K}^{*0}(892)\,\{\pi^+\pi^-\}_P$ (D-wave)	$6.1e^{-3}$	0.54	2.94

Table 5.4: Amplitude model of $D^0 \to K^-\pi^+\pi^-\pi^+$ according to a Mint-Dalitz fit to the LHCb MC. The fit fractions ($\mathcal{F}_i$), magnitudes (a_i) and phases (ϕ_i) of the decay amplitudes are shown. Where no phase is given, this indicates that the amplitudes are incoherently summed. $\overline{K}^{*0}(892)$ and $\rho^0(770)$ are assumed to exclusively decay to $K^-\pi^+$ and $\pi^+\pi^-$ respectively.

5.5.3 Amplitude model weights

Each MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ decay in the LHCb MC has been assigned a weight, $w(\mathbf{x})$, defined by

$$w(\mathbf{x}) = \frac{\left|A^{\mathrm{MarkIII}}(\mathbf{x})\right|^2}{\left|A^{\mathrm{LHCb}}(\mathbf{x})\right|^2}, \tag{5.13}$$

where the superscripts "MarkIII" and "LHCb" indicate the decay amplitude according to the Mint-Dalitz implementation of the MarkIII and LHCb amplitude models respectively. The Dalitz-plot variables $\mathbf{x}$ are determined for each signal decay using the using the *generated* four-momenta of the D^0 and its daughters (as opposed to the reconstructed four-momenta). The following section describes how these model weights are used to determine the acceptance efficiencies and selection and reconstruction efficiencies.

5.6 MC-derived efficiencies

The acceptance efficiencies, and the reconstruction and selection efficiencies in bins of p_{T} and y of the D^0 meson have been determined using MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. In each case, the model weights defined in Sec. 5.5 are used. For a given (p_{T}, y) bin i, the acceptance/selection and reconstruction efficiency, ϵ_i, is given by

$$\epsilon_i = \frac{\sum_{j=0}^{r_i} w(\mathbf{x}_j)}{\sum_{k=0}^{n_i} w(\mathbf{x}_k)}, \tag{5.14}$$

where $w(\mathbf{x}_j)$ indicates the model weight of the j^{th} truth-matched signal candidate. The denominator is summed over all n_i MC signal candidates in (p_{T}, y) bin i, whilst the numerator is summed over all r_i signal candidates in this bin that pass the efficiency requirements.

The statistical uncertainty on the weighted efficiency ϵ_i, σ_i, is determined using the Gaussian approximation to the binomial distribution

$$\sigma_i = \frac{\sqrt{(1-2\epsilon_i)\sum_{j=0}^{r_i} w(\mathbf{x}_j)^2 + \epsilon_i \sum_{k=0}^{n_i} w(\mathbf{x}_k)}}{\sum_{j=0}^{n_i} w(\mathbf{x}_j)}. \tag{5.15}$$

These statistical uncertainties are assigned as sources of systematic uncertainty on the cross-section measurements.

Since the particle identification ($\log(\mathcal{L}_K/\mathcal{L}_\pi)$) distributions are poorly described in the MC, cuts on the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ of the D^0 daughter tracks are excluded when determining the selection and reconstruction efficiencies. The efficiencies of these cuts are instead determined using a data-driven approach discussed in Sec. 5.7.

A more detailed explanation of how the MC-derived efficiencies are determined is given below, and the efficiencies are also presented. In addition, the following sources of corrections to these efficiencies are discussed:

- corrections due to the failure of the MC truth-matching algorithms;
- corrections due to differences between the cut-variable distributions in data and MC.

5.6.1 Acceptance efficiency

Requirements that the D^0 daughters are produced within the LHCb acceptance have been applied in the generation of the MC used to determine the selection and reconstruction efficiencies. Thus, the acceptance efficiency is determined using *generator-level* MC (*i.e.* the decays have not been reconstructed, nor propagated through the simulated detector) produced without imposing that the D^0 decay products are generated within the LHCb acceptance.

Only generated events for which the event status according to HepMC [89] returns a value consistent with a signal decay are considered. Secondary decays are rejected by requiring that the MC particle associated with the D meson does not have any ancestors with a mean lifetime (according to the MC-truth information) greater than 0.1 fs. This requirement excludes all D mesons which originate from b-hadrons.

The acceptance efficiencies in bins of D^0 p_{T} and y are shown in Tab. 5.5. Here and subsequently, bins with an insufficient number of events to determine the prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ yield in data have been excluded.

p_{T} (GeV/c)	y (2.0, 2.5)	(2.5, 3.0)	(3.0, 3.5)	(3.5, 4.0)	(4.0, 4.5)
(0, 1)			0.915 ± 0.006	0.787 ± 0.008	
(1, 2)		0.955 ± 0.004	0.940 ± 0.005	0.872 ± 0.007	
(2, 3)	0.833 ± 0.008	0.988 ± 0.003	0.974 ± 0.004	0.945 ± 0.007	0.796 ± 0.012
(3, 4)	0.886 ± 0.009	0.996 ± 0.002	0.984 ± 0.005	0.971 ± 0.007	0.893 ± 0.015
(4, 5)	0.954 ± 0.008	0.999 ± 0.002	0.983 ± 0.006	0.987 ± 0.007	0.952 ± 0.015
(5, 6)	0.971 ± 0.009	0.997 ± 0.003	1.000 ± 0.001	0.992 ± 0.008	0.999 ± 0.003
(6, 7)	0.993 ± 0.006	1.000 ± 0.000	0.984 ± 0.011	0.993 ± 0.009	0.976 ± 0.021
(7, 8)	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000

Table 5.5: Acceptance efficiencies for MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ decays.

5.6.2 Selection and reconstruction efficiency

As previously mentioned, the selection and reconstruction efficiency is determined from a sample of MC events for which generator-level MC cuts have been applied. The denominator is determined from the MC events, and is evaluated prior to reconstruction, whilst the numerator is determined by reconstructing the MC events used in the denominator and requiring that the reconstructed $D^0 \to K^-\pi^+\pi^-\pi^+$ decay candidates be associated to a genuine $D^0 \to K^-\pi^+\pi^-\pi^+$ decay according to the MC-truth information. The reconstructed MC sample is filtered using the same stripping and offline selection requirements applied to the data (apart from the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts). The selection and reconstruction efficiencies in bins of D^0 p_{T} and y are shown in Tab. 5.6.

p_{T} (GeV/c)	y (2.0, 2.5)	(2.5, 3.0)	(3.0, 3.5)	(3.5, 4.0)	(4.0, 4.5)
(0, 1)			0.13 ± 0.00	0.15 ± 0.00	
(1, 2)		0.12 ± 0.00	0.53 ± 0.00	0.51 ± 0.00	
(2, 3)	0.07 ± 0.00	0.98 ± 0.00	2.19 ± 0.01	2.09 ± 0.01	0.60 ± 0.00
(3, 4)	0.39 ± 0.00	3.08 ± 0.01	5.28 ± 0.03	4.49 ± 0.03	1.60 ± 0.01
(4, 5)	1.26 ± 0.01	5.68 ± 0.04	8.47 ± 0.06	7.56 ± 0.06	2.98 ± 0.03
(5, 6)	2.12 ± 0.02	8.28 ± 0.07	11.18 ± 0.11	9.61 ± 0.11	3.41 ± 0.05
(6, 7)	3.53 ± 0.04	10.17 ± 0.12	12.55 ± 0.16	11.38 ± 0.18	3.73 ± 0.08
(7, 8)	4.21 ± 0.06	11.41 ± 0.17	14.16 ± 0.25	12.92 ± 0.27	2.46 ± 0.07

Table 5.6: Selection and reconstruction efficiencies for MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ decays (%).

5.6.3 Corrections for failure of MC truth-matching

Failures in the algorithms that match MC particles to reconstructed particles leads to an underestimation of the number of reconstructed and selected $D^0 \to K^-\pi^+\pi^-\pi^+$ decays, and subsequently an underestimation of the selection and reconstruction efficiency. The failure rate of the MC-association algorithms, F, is determined by

$$F = \frac{N_{\text{failed}}}{N_{\text{failed}} + N_{\text{matched}}} \,, \tag{5.16}$$

where N_{matched} is the number of reconstructed, selected and successfully MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates, whilst N_{failed} is the number of reconstructed and selected "true" $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates for which one or more particles in the decay chain failed to be associated to an MC particle. The quantity N_{failed} has been estimated by fitting the D^0 invariant-mass distribution of candidates which fail the association. Due to limited event numbers, it is not possible to fit both the D^0 invariant mass and D^0 $IP\,\chi^2$ distributions, so the resulting samples contain a mixture of prompt and secondary $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. In addition, it is not possible to assign model weights to the $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates for which the truth-matching failed, since the Dalitz-plot variables used to determine the weights are obtained from the MC-truth information. It is assumed that the effect of these two issues on the measured correction factor will be small compared to the statistical uncertainty on the correction factor, which is assigned as a systematic uncertainty on the cross-section results.

The mass distribution of $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates (prompt or secondary) which fail MC association are modelled by a Gaussian PDF, whilst the combinatoric background is modelled by a first-order Chebychev polynomial. The fit results in the full p_{T} and y range are shown in Fig. 5.1. From this, a correction factor to the selection and reconstruction efficiency of $1 - F = (96.65 \pm 0.24)\%$ is found. The statistical uncertainty on this result is assigned as a source of systematic uncertainty on the cross-section.

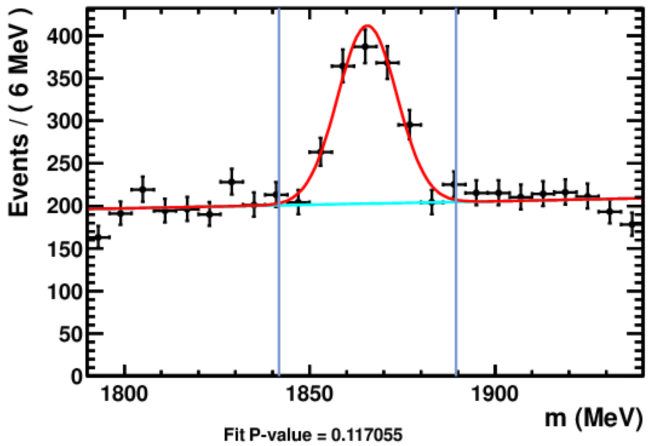


Figure 5.1: D^0 invariant mass unbinned maximum likelihood fit to $D^0 \to K^-\pi^+\pi^-\pi^+$ MC candidates which are not associated to a true $D^0 \to K^-\pi^+\pi^-\pi^+$ decay according to MC-truth information. The p-value according to the maximum likelihood fit is 0.117.

5.6.4 Corrections due to data/MC disagreement

The MC does not always perfectly reproduce the distribution of variables used in the stripping or offline selections. This means that the selection and reconstruction efficiency determined from MC may differ from the efficiency in the data. For cut-variable distributions that show significant differences between data and MC, a correction factor of $\frac{\varepsilon_{\rm data}}{\varepsilon_{\rm MC}}$ is applied, where $\varepsilon_{\rm MC}$ is the efficiency of the offline selection cut according to MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates and $\varepsilon_{\rm data}$ is the equivalent efficiency according to background-subtracted data candidates. The efficiencies are determined after all other offline cuts have been applied. The correction factors and associated systematic uncertainties are shown in Tab. 5.7. Since there is an insufficient number of events to determine whether the data-MC disagreement varies with the $p_{\rm T}$ or y of the D^0, conservative systematic uncertainties are assigned throughout.

For variables which are not cut on in the stripping or are cut on loosely (D^0 vertex $\chi^2/N_{\rm dof}$, D^0 *DoCA* and D^0 *DIRA*), the efficiencies are determined directly from the cut-variable distributions. Systematic uncertainties are assigned to these correction factors equal to half the absolute difference between the data and MC efficiencies.

For variables with moderate stripping cuts ($\ln(IP\,\chi^2)$ and track $\chi^2/N_{\rm dof}$ of D^0 daughters) the efficiencies are estimated by binning the data and MC distributions of the variables close to the stripping cut and linearly extrapolating into the region rejected by the stripping cut. This procedure is performed over three different ranges, as indicated by the second column in Tab. 5.7. The median value of $\frac{\varepsilon_{\text{data}}}{\varepsilon_{\text{MC}}}$ over these three ranges is assigned as the correction factor, with the systematic uncertainty determined as the absolute difference between the other two efficiency ratios.

For the remaining variables considered (p_{T} and $|\vec{p}|$ of the D^0 daughters), the stripping cuts were necessarily tight in order to reduce the level of combinatoric background. In each of these cases, the data and MC cut-variable distributions are binned, and the ratio of the data and MC distributions calculated. This ratio distribution is then fitted with a first-order Chebychev polynomial. The cut-variable distribution in data is simulated by using the fitted functions to re-weight the cut-variable distributions of the MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ decays without the associated stripping cut applied. Finally, the efficiencies of the cuts according to the re-weighted data and MC are calculated, and the correction factor determined. The fitted functions are shown in the second column of Tab. 5.7. The systematic uncertainties associated with these correction factors are determined by the absolute difference between the efficiency ratios determined when the gradients of the reweighting functions are varied by $\pm 1\sigma$, where σ is the fit error.

The final row of Tab. 5.7 indicates the overall correction factor, determined by the product of the individual correction factors. Since several of the cut-variable distributions are highly correlated (*e.g.* the p_{T} and $|\vec{p}|$ of the D^0 daughters), applying all offline cuts other than the cut associated with the variable under investigation is likely to distort the cut-variable distribution, and this may lead to an underestimation of the difference between the cut efficiencies in data and MC. Consequently, a conservative systematic uncertainty is assigned to the combined correction factor by including correlation terms when propagating the individual uncertainties. The correlation between uncertainties are determined by the Pearson correlation between the variables according to the MC-associated candidates.

Variable	Fit range / function	$\frac{\varepsilon_{\text{data}}}{\varepsilon_{\text{MC}}}$	$\sigma\left(\frac{\varepsilon_{\text{data}}}{\varepsilon_{\text{MC}}}\right)$
D^0 vertex χ^2/N_{dof}	—	**0.8828**	0.0586
D^0 *DoCA*	—	**0.9532**	0.0234
D^0 *DIRA*	—	**0.9684**	0.0158
D^0 daughter with lowest $\ln(IP\,\chi^2)$	$0 < \ln(IP\,\chi^2) < 1.0$	0.9952	0.0222
	$0 < \ln(IP\,\chi^2) < 1.5$	**1.0113**	
	$0 < \ln(IP\,\chi^2) < 2.0$	1.0175	
D^0 daughter with 2$^{\text{nd}}$ lowest $\ln(IP\,\chi^2)$	$0.4 < \ln(IP\,\chi^2) < 1.8$	0.9847	0.0160
	$0.4 < \ln(IP\,\chi^2) < 2.2$	**0.9772**	
	$0.4 < \ln(IP\,\chi^2) < 2.5$	0.9686	
D^0 daughter with 2$^{\text{nd}}$ highest $\ln(IP\,\chi^2)$	$1.7 < \ln(IP\,\chi^2) < 2.5$	0.9943	0.0055
	$1.7 < \ln(IP\,\chi^2) < 3.0$	0.9888	
	$1.7 < \ln(IP\,\chi^2) < 3.5$	**0.9915**	
D^0 daughter with highest $\ln(IP\,\chi^2)$	$2.1 < \ln(IP\,\chi^2) < 3.0$	0.9942	0.0019
	$2.1 < \ln(IP\,\chi^2) < 3.5$	**0.9944**	
	$2.1 < \ln(IP\,\chi^2) < 4.0$	0.9962	
D^0 daughter with largest χ^2/N_{dof}	$2.8 < \chi^2/N_{\text{dof}} < 5.0$	0.9921	0.0107
	$3.5 < \chi^2/N_{\text{dof}} < 5.0$	**0.9906**	
	$4.2 < \chi^2/N_{\text{dof}} < 5.0$	0.9813	
D^0 daughter with lowest p_{T}	$-0.448 \times p_{\text{T}} + 1.23$	0.9823	0.0051
	$-0.393 \times p_{\text{T}} + 1.23$	**0.9849**	
	$-0.338 \times p_{\text{T}} + 1.23$	0.9874	
D^0 daughter with lowest $\lvert\vec{p}\rvert$	$-0.019 \times \lvert\vec{p}\rvert + 1.12$	0.9891	0.0041
	$-0.016 \times \lvert\vec{p}\rvert + 1.12$	**0.9912**	
	$-0.012 \times \lvert\vec{p}\rvert + 1.12$	0.9932	
Total		0.768	0.088

Table 5.7: Correction factors and systematic uncertainties arising from differences between the distributions of cut variables in data and MC. The correction factors highlighted in bold are those that are used to determine the total correction factor in the bottom row of the table.

5.7 PID calibration

Since the PID variable $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ is poorly described in the LHCb simulation, MC signal samples cannot be used to reliably determine the efficiency of cuts on this variable. Instead, the following *calibration* data samples are used for this purpose:

- $\Lambda \to p^+\pi^-$, $\overline{\Lambda} \to p^-\pi^+$ and $K^0_\mathrm{S} \to \pi^+\pi^-$, commonly referred to as V^0 decays;
- $\phi \to K^+K^-$.

To ensure that the kinematic distributions of the calibration track samples are indicative of the data used elsewhere in the analysis, the decays used for the PID calibration have been extracted from the same $\mathcal{L}_\mathrm{int} = 14.96\ \mathrm{nb}^{-1}$ sample containing the $D^0 \to K^-\pi^+\pi^-\pi^+$ decays discussed in Sec. 5.8. In addition, the events are required to pass at least one of the micro-bias HLT triggers.

Section 5.7.2 describes how the calibration samples are used to determine the PID efficiencies of kaon and pion tracks. The resulting PID efficiencies are then combined to determine the PID efficiency for each truth-matched $D^0 \to K^-\pi^+\pi^-\pi^+$ MC candidate, discussed in Sec. 5.7.3. Ideally, these signal-candidate PID efficiencies would be also calculated using data; however, the PID cuts applied in the stripping prevent this. The uncertainty due to differences in the track kinematics in data and MC is incorporated into the systematic uncertainties determined in Sec. 5.7.5. The "final" PID efficiency in each D^0 (p_T, y) bin is determined by performing a weighted average over the PID efficiencies of the truth-matched signal candidates, explained in Sec. 5.7.4.

5.7.1 Calibration samples

The weakly decaying V^0 hadrons have been selected using a series of stripping cuts shown in Tab. 5.8. To ensure the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distributions are unbiased, no cut on this variable for either daughter track has been applied in the stripping. A total of 1.21 (1.10) million K^0_S

Variable name	$\Lambda \to p^+\pi^-$	$K^0_S \to \pi^+\pi^-$	$\phi \to K^+K^-$
Mass window (MeV/c^2)	25	50	25
max(Z) (cm) $<$	220	220	
Wrong mass (MeV/c^2)	20	9	
$c\tau$ (mm) $<$	5	1	
Lifetime fit $\chi^2 >$	36	36	
Vertex $\chi^2 <$	25	25	10
h track $\chi^2/N_{\rm dof} <$	5	5	
h IP $\chi^2 >$	25	25	
h $\lvert\vec{p}\rvert$ (GeV/c) $>$	2	2	2
Kaon RICH $\log(\mathcal{L}_K/\mathcal{L}_\pi) >$			15

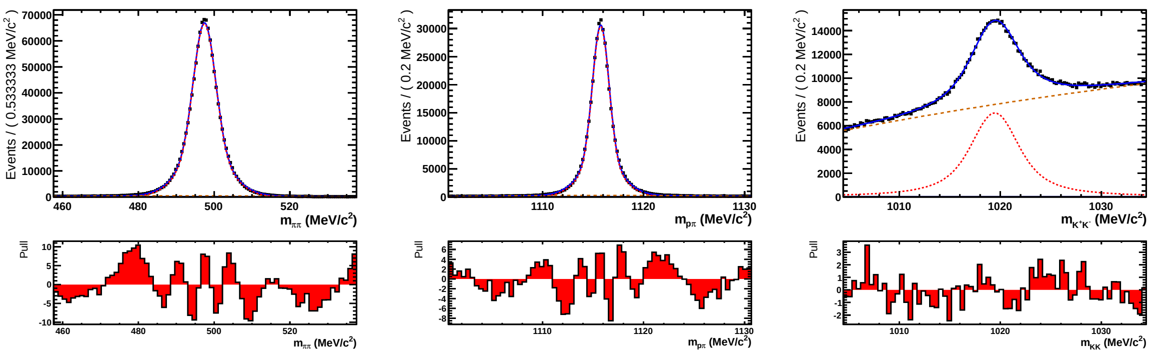
Table 5.8: Selection criteria for $\Lambda \to p^+\pi^-$, $K^0_S \to \pi^+\pi^-$ and $\phi \to K^+K^-$ PID calibration candidates. *max(Z)* represents the maximum allowed position of the reconstructed V^0 vertex along the beam axis. The quantity $c\tau$ is the proper lifetime of the candidate multiplied by the speed of light in vacuum. *Wrong mass* refers to the absolute difference between the "correct" and "wrong" mass hypotheses. For $K^0_S \to \pi^+\pi^-$, the wrong mass hypothesis means that the candidate can be reconstructed as $\Lambda \to p^+\pi^-$ (and charge conjugate), whilst the $\Lambda \to p^+\pi^-$ candidates can be reconstructed as $K^0_S \to \pi^+\pi^-$. Kaon $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ represents the log-likelihood difference between the kaon and pion mass hypothesis using only information obtained from the RICH detectors.

and 0.43 (0.38) million Λ[1] magnet-up (-down) data candidates have been reconstructed with these cuts, with a purity greater than 90%. The invariant-mass distributions of the K^0_S and Λ distributions for magnet up data are shown in Figs. 5.2a and 5.2b respectively.

The strongly decaying $\phi \to K^+K^-$ resonance suffers from large backgrounds, so achieving a high purity from the stripping selection is not possible without exploiting PID information. In this case, a *tag-and-probe* method has been employed. The stripping requirements have been chosen such that ϕ candidates are selected if *at least* one track satisfies the RICH PID requirement $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 15$. The ϕ candidates are filtered offline so the track being considered has no PID requirements. Again, the full list of cuts can be found in Tab. 5.8. From these selection requirements, 5.76 (5.30) million magnet up (down) ϕ candidates are obtained. As indicated by Fig. 5.2c, even with the RICH PID requirement on the single kaon, the ϕ data samples are dominated by combinatoric background. The ${}_s\mathcal{P}lot$ method, discussed in Sec. 4.8, is used to extract the kaon $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ distribution of the correctly reconstructed ϕ candidates using the invariant mass of the ϕ as the discriminating variable.

[1] Hereby referring to either Λ or $\overline{\Lambda}$.

Although the background contamination of the V^0 samples is small in comparison to the ϕ sample, the same ${}_s\mathcal{P}lot$ technique is used for all track samples. Fig. 5.2 shows the results of an unbinned maximum likelihood fit to the mass distributions. The two V^0 signal peaks are parametrised by the sum of two Gaussian PDFs with a common mean parameter, whilst the ϕ resonance is parametrised by a *Voigt* PDF (the convolution of a Gaussian function and a Breit-Wigner function). In all three cases, the background shape is modelled by a third-order Chebychev polynomial. The calibration data are split into six distinct datasets, categorised by particle type (kaon or pion), magnet polarity and track charge. To maximise the statistical power of the analysis, the pion tracks from the two V^0 samples are combined into a single dataset.



(a) $K^0_S \to \pi^+\pi^-$ (b) $\Lambda \to p^+\pi^-$ (c) $\phi \to K^+K^-$

Figure 5.2: Invariant mass distributions of $K^0_S \to \pi^+\pi^-$ (a), $\Lambda \to p^+\pi^-$ (b) and $\phi \to K^+K^-$ (c) candidates from magnet up data. The fitted signal (magenta) and background (brown) PDFs are shown on each plot, along with the combined signal and background PDF (blue). The pull values of the combined PDFs in uniform bins are shown below the mass plots.

As discussed in Sec. 5.7.2, due to the large variation in background level as a function of track momentum, the ϕ samples are initially split by the momentum of the unbiased track in order to determine the ${}_s\mathcal{P}lot$ weights, then these sub-samples are recombined to determine the PID efficiencies. The PIDCalib [90] software packages, developed by Andrew Powell (University of Oxford) and the author, have been used to obtain the calibration samples, perform the maximum likelihood fits, calculate and apply the ${}_s\mathcal{P}lot$ weights and determine the PID efficiencies of the calibration tracks. An explanation of how the PID efficiencies of the calibration tracks are determined now follows.

5.7.2 Track PID efficiencies

Once the $_s\mathcal{P}lot$ weights have been determined, the PID efficiency of the kaon and pion tracks can be calculated. However, determining the overall PID efficiency for a given signal candidate is complicated by the fact that the PID performance depends on the track kinematics. This variation is dealt with by binning the PID efficiency in momentum, $|\vec{p}|$ and pseudo-rapidity, η of the track. Since the calibration tracks tend to be *softer* (have lower momentum) than those originating from signal decays, there are regions of kinematic space which are populated by signal tracks but not calibration tracks. To ensure that these differences do not bias the PID efficiencies determined from this study, the following additional kinematic cuts are applied to all signal tracks to veto the regions with no events in the calibration samples:

- $3.2\,\mathrm{GeV}/c < |\vec{p}| < 100\,\mathrm{GeV}/c$;
- $2 < \eta < 5$;
- $p_{\mathrm{T}} < 5\,\mathrm{GeV}/c$.

The following bin scheme is employed for kaon tracks, as a compromise between minimising bin area and maximising the number of calibration events in each bin:

- $\boldsymbol{|\vec{p}|}$ **(GeV/$\boldsymbol{c}$)**: [3.2,6.0], [6.0,9.3], [9.3,15.6], [15.6,25.0], [25.0,40.0], [40.0,100.0];
- $\boldsymbol{\eta}$: [2.0,2.5], [2.5,3.0], [3.0,3.5], [3.5,4.0], [4.0,5.0].

As the number of V^0 signal candidates available is significantly larger than the size of the ϕ sample, a finer binning scheme is employed for the pion tracks:

- $\boldsymbol{|\vec{p}|}$ **(GeV/$\boldsymbol{c}$)**: [3.2,6.0], [6.0,9.3], [9.3,12.0], [12.0,15.6], [15.6,25.0], [25.0,40.0], [40.0,60.0], [60.0,100.0];
- $\boldsymbol{\eta}$: [2.0,2.5], [2.5,3.0], [3.0,3.5], [3.5,3.75], [3.75,4.0], [4.0,4.25], [4.25,4.5], [4.5,5.0].

The PID efficiency in each $(|\vec{p}|\,, \eta)$ bin is determined using the same technique used to determine the MC-derived efficiencies introduced in Sec. 5.6, with the model weights replaced by the ${}_s\mathcal{P}lot$ weights. However, the weighted PID efficiency determined may not yield reliable results if the parametrisation of the various species (signal and background in this case) differs between the kinematic bin considered and the full calibration data sample from which the weights are constructed. In the case of the ϕ samples, both the background level compared to the signal peak and the shape of the background PDF are found to vary significantly with track momentum, as can be seen in Fig. 5.3. As such, the weights for the ϕ data samples are constructed using the same $|\vec{p}|$ bin scheme used to determine the PID efficiencies.

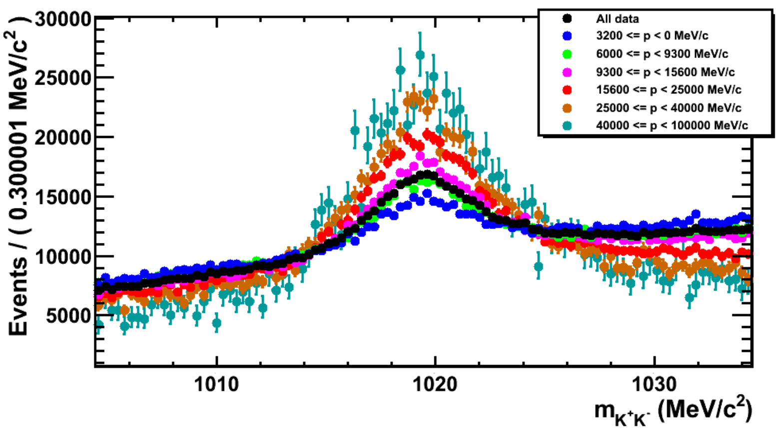


Figure 5.3: Invariant mass distribution of $\phi \to K^+K^-$ candidates reconstructed from $\sqrt{s} = 7\,\mathrm{TeV}$ magnet down data in bins of $|\vec{p}|$ of the unbiased K^+ track. The black points show the mass distribution integrated over the full $|\vec{p}|$ range. The data in each bin share a common normalisation.

Another issue with the determination of the PID efficiencies of the kaon calibration tracks is unreliable PID efficiencies in some of the high η, low momentum bins, as a result of the majority of ${}_s\mathcal{P}lot$ weights in this region being negative. The η versus $|\vec{p}|$ distribution of the unbiased K^+ track from magnet up data is shown in Fig. 5.4. As can be seen, the top left region of the plot contains very few tracks. This low acceptance region roughly corresponds to the $(|\vec{p}|\,, \eta)$ range $([3.2, 8], [4.4, 5.0])$. In this corner of kinematic space, the tracks produced are soft, and almost collinear to the beam pipe. As a result, kaon tracks in this region are vetoed in both the calibration and signal samples.

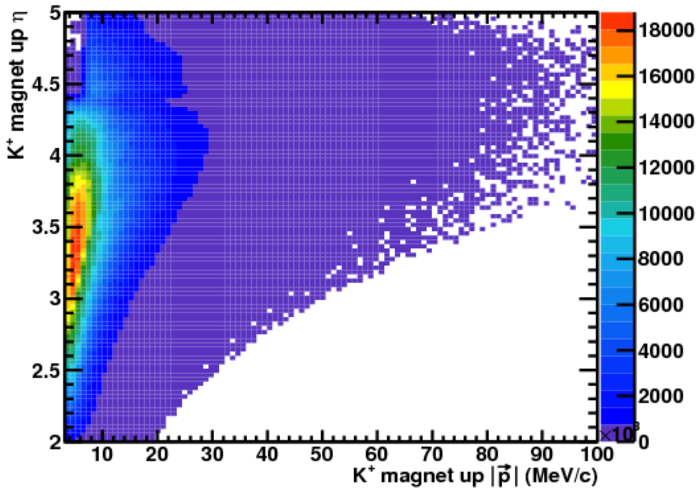


Figure 5.4: Plot of η against $|\vec{p}|$ of the unbiased K^+ track from $\phi \to K^+K^-$ candidates, reconstructed from $\sqrt{s} = 7\,\text{TeV}$ magnet-up data. The data correspond to the "raw" number of candidates, *i.e.* before unfolding the signal using the ${}_s\mathcal{P}lot$ technique.

5.7.3 Signal candidate PID efficiencies

For each MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ signal candidate, i, the PID efficiency of the candidate, $\epsilon(\mathbf{z}_i)$, is determined by the product of the PID efficiencies of its final-state tracks,

$$\epsilon(\mathbf{z}_i) = \prod_{j=1}^{4} \epsilon\left(\mathbf{x}_j^i\right) \quad , \tag{5.17}$$

where $\epsilon\left(\mathbf{x}_j^i\right)$ is the efficiency of the j^{th} final-state track of the i^{th} $D^0 \to K^-\pi^+\pi^-\pi^+$ signal decay i, $\mathbf{x}_j^i$ indicates the $|\vec{p}|$ and η of the track, along with the track type (kaon/pion), magnet polarity and charge and the quantity $\mathbf{z}_i \equiv (\mathbf{x}_1^i, \mathbf{x}_2^i, \mathbf{x}_3^i, \mathbf{x}_4^i)$ corresponds to a particular combination of signal-track PID efficiencies. This information is then used to assign a PID efficiency to the signal track according to the calibration-track PID efficiencies based on which calibration sample it corresponds to (*i.e.* $K^\pm/\pi^\pm$, magnet up/down) and which of the calibration kinematic bins the tracks reside in.

The statistical uncertainty associated with the PID efficiency of $D^0 \to K^-\pi^+\pi^-\pi^+$ candidate i is given by

$$\sigma(\mathbf{z}_i) = \sqrt{\sum_{j=1}^{4}\sum_{k=1}^{4} \text{cov}_{j,k}\left(\mathbf{x}_j^i, \mathbf{x}_k^i\right)} \ , \tag{5.18}$$

where $\mathrm{cov}^i_{j,k}\left(\mathbf{x}^i_j, \mathbf{x}^i_k\right)$ is the covariance between tracks j and k of candidate i, equal to the square of the statistical uncertainty in the case that j and k are equal. The two same-charge pions are treated as 100% statistically correlated, since the same calibration sample is used to determine the PID efficiencies of these tracks. All other tracks in the signal decay are treated as uncorrelated.

5.7.4 Final PID efficiencies

There are only a discrete number of the $\mathbf{z}_i$ combinations (Eq. (5.17)) as a result of the binning the calibration tracks to determine the signal-track PID efficiencies. Consequently, the PID efficiency of the signal candidates in each (p_T, y) bin j, $\langle\epsilon_j\rangle$ is determined by performing a weighted average over all "unique" signal-candidate PID efficiencies, *i.e.*

$$\langle\epsilon_j\rangle = \frac{\sum_{i=1}^{N} w_i^j \epsilon\left(\mathbf{z}_i\right)}{\sum_{i=1}^{N} w_i^j} \ , \tag{5.19}$$

where w_i^j corresponds to the sum of the Dalitz-model weights of all signal candidates that have been assigned a PID efficiency $\epsilon\left(\mathbf{z}_i\right)$. The superscript indicates that only candidates in (p_T, y) bin j are considered when determining the sum of weights. The numerator and denominator of Eq. (5.19) are summed over all N unique efficiencies $\epsilon\left(\mathbf{z}_i\right)$.

The statistical uncertainty of the weighted average PID efficiency, $\sigma_{\langle\epsilon_j\rangle}$, is determined by

$$\sigma_{\langle\epsilon_j\rangle} = \frac{\sqrt{\sum_{i=1}^{N} w_i^j \sigma\left(\mathbf{z}_i\right)^2}}{\sum_{i=1}^{N} w_i^j} \ . \tag{5.20}$$

The average PID efficiencies for $D^0 \to K^-\pi^+\pi^-\pi^+$ for a kaon $\log\left(\mathcal{L}_K/\mathcal{L}_\pi\right)$ cut of > 5, and a $\log\left(\mathcal{L}_K/\mathcal{L}_\pi\right)$ cut of < 5 for all three pions are shown in Tab. 5.9.

The above method of calculating the average PID efficiency by binning in $|\vec{p}|$ and η of the tracks makes the following assumptions:

- the PID performance of the tracks can be entirely parametrised by the distributions of $|\vec{p}|$ and η, *i.e.* the PID performance does not depend on any other variables;

p_{T} (GeV/c)	y (2.0, 2.5)	(2.5, 3.0)	(3.0, 3.5)	(3.5, 4.0)	(4.0, 4.5)
(0, 1)			78.68 ± 0.11	76.14 ± 0.10	
(1, 2)		82.38 ± 0.09	81.03 ± 0.05	79.10 ± 0.05	
(2, 3)	85.76 ± 0.14	84.66 ± 0.03	84.15 ± 0.02	81.99 ± 0.03	76.21 ± 0.07
(3, 4)	87.30 ± 0.07	86.82 ± 0.02	86.62 ± 0.02	83.57 ± 0.03	75.18 ± 0.07
(4, 5)	88.86 ± 0.06	88.58 ± 0.03	88.32 ± 0.03	84.50 ± 0.04	74.35 ± 0.08
(5, 6)	90.17 ± 0.07	89.79 ± 0.04	89.09 ± 0.04	83.91 ± 0.05	72.75 ± 0.13
(6, 7)	91.60 ± 0.08	90.82 ± 0.05	89.04 ± 0.07	83.60 ± 0.10	74.84 ± 0.20
(7, 8)	92.09 ± 0.12	91.65 ± 0.08	89.51 ± 0.13	83.48 ± 0.12	73.23 ± 0.30

Table 5.9: PID efficiencies (%) for a kaon PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) > 5$ and a pion PID cut of $\log(\mathcal{L}_K/\mathcal{L}_\pi) < 5$. The uncertainties shown are statistical only. Bins with an insufficient number of events to determine the prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ yield have been excluded.

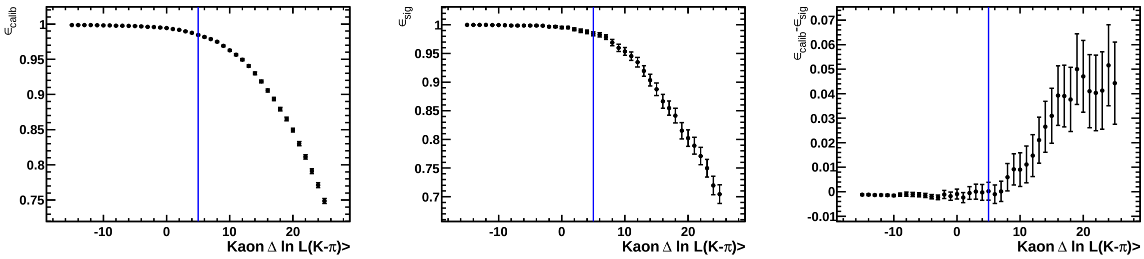
- the LHCb signal MC accurately models the distribution of these variables;
- the variation of PID performance within a specific bin of $|\vec{p}|$ and η is minimal;
- the use of the same weights returned by the $_s\mathcal{P}lot$ method to weight the track distributions before and after applying the PID cut is a valid approach (as previously discussed).

The systematic uncertainty due to the imperfection of these assumptions is discussed in Sec. 5.7.5.

5.7.5 PID efficiency systematic uncertainty

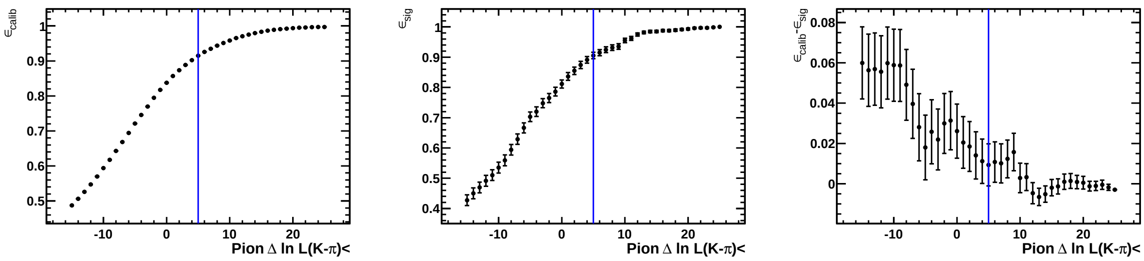
To assess the systematic uncertainty associated with the assumptions made in the implementation of the PID efficiency calibration procedure, the calibration method is used to determine the PID efficiencies using MC-associated calibration samples. This allows the predictions of the PID efficiencies from the calibration method (determined from MC rather than data), hereby referred to as the *calibration efficiency* ($\varepsilon_{\mathrm{calib}}$) to be compared to the “true” PID efficiencies determined by applying the $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts to the final-state particles of the truth-matched signal candidates, hereby referred to as the *signal efficiency* ($\varepsilon_{\mathrm{sig}}$).

Figures 5.5 and 5.6 show the signal and calibration efficiencies over a range of $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts, as well as the efficiency *residuals* (difference between calibration and signal efficiencies) for a particular (p_{T}, y) bin. Figure 5.5 shows the efficiencies determined for a $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cut applied only to the D^0 daughter kaon track, whilst Fig. 5.6 shows the efficiencies for $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts applied to the three D^0 daughter pion tracks. The central values and statistical uncertainties of the calibration efficiencies are determined using the (non-continuity corrected) Wilson score interval [91].



(a) (b) (c)

Figure 5.5: PID efficiency for $D^0 \to K^-\pi^+\pi^-\pi^+$ MC according to (a) MC calibration tracks ($\varepsilon_{\mathrm{calib}}$), (b) MC-associated signal tracks ($\varepsilon_{\mathrm{sig}}$), and (c) the difference between (a) and (b) ($\varepsilon_{\mathrm{calib}}$-$\varepsilon_{\mathrm{sig}}$) for a range of kaon $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts. The results shown correspond to the analysis bin $(p_{\mathrm{T}}, y) = ([3.0, 4.0]\,\mathrm{GeV}/c, [3.0, 3.5])$.



(a) (b) (c)

Figure 5.6: PID efficiency for $D^0 \to K^-\pi^+\pi^-\pi^+$ MC according to (a) MC calibration tracks ($\varepsilon_{\mathrm{calib}}$), (b) MC-associated signal tracks ($\varepsilon_{\mathrm{sig}}$), and (c) the difference between (a) and (b) ($\varepsilon_{\mathrm{calib}}$-$\varepsilon_{\mathrm{sig}}$) for a range of pion $\log(\mathcal{L}_K/\mathcal{L}_\pi)$ cuts. The results shown correspond to the analysis bin $(p_{\mathrm{T}}, y) = ([3.0, 4.0]\,\mathrm{GeV}/c, [3.0, 3.5])$.

To account for possible lateral shifts of the residuals in data compared to MC, the systematic uncertainty is calculated by considering residuals in the range ± 5 of the nominal PID cut, and the largest central value of the residual in this range is taken as the systematic uncertainty. For each (p_{T}, η) bin, a systematic is determined for the kaon track and separately for the three pion tracks, and the two systematics are summed in quadrature to determine the total systematic uncertainty due to the combined efficiencies of the four D^0 daughter tracks.

The technique of estimating the systematic uncertainty works well for the majority of bins. However, in some cases, the systematic uncertainties returned by the method are unrealistically large. These cases correspond to bins which have very large statistical uncertainties in the MC calibration samples according to the Wilson score interval. Since the Wilson score interval shifts the central value to the centre of the confidence interval, the naive method of taking the largest central value of the residual can lead to exaggerated systematics if the statistical uncertainty of the calibration efficiency is large compared to the signal efficiency. In these cases, a conservative estimate of the systematic uncertainty has been determined by assigning the largest systematic uncertainty over all the bins with sufficient MC events (defined as bins with $\leq 3\%$ statistical uncertainty on the efficiency residual measurement). A summary of the systematic uncertainties assigned to the PID efficiency measurement in each (p_{T}, y) bin is given in Tab. 5.10.

p_{T}			y		
(GeV/c)	$(2.0, 2.5)$	$(2.5, 3.0)$	$(3.0, 3.5)$	$(3.5, 4.0)$	$(4.0, 4.5)$
$(0, 1)$	8.72	8.72	7.29	8.72	8.72
$(1, 2)$	8.72	6.36	4.94	4.89	8.72
$(2, 3)$	8.72	6.22	4.90	3.43	5.97
$(3, 4)$	8.72	5.01	2.77	2.88	5.68
$(4, 5)$	7.43	2.60	1.65	2.51	7.48
$(5, 6)$	5.44	2.62	2.53	3.51	8.72
$(6, 7)$	3.88	2.73	3.17	5.84	8.72
$(7, 8)$	3.98	2.14	5.48	6.63	8.72

Table 5.10: Combined kaon and pion PID systematic uncertainties (%).

5.8 Yield extraction

The prompt yields are extracted from an unbinned maximum likelihood fit in the (D^0 mass, D^0 $\log(IP\,\chi^2)$) plane, implemented using the RooFit framework. As previously mentioned, the former distribution is used to distinguish between combinatoric background and $D^0 \to K^-\pi^+\pi^-\pi^+$ decays (prompt or secondary), whilst the latter distribution is used to distinguish between prompt and secondary D^0 decays. The components of the fit, the fit results, and the sources of corrections to the yields follow in this section.

5.8.1 Yield extraction method

The D^0 invariant mass projection of the PDF components are fit with:

- a common Crystal Ball function for both the prompt signal and secondary background;
- a first-order Chebychev polynomial for the combinatoric background.

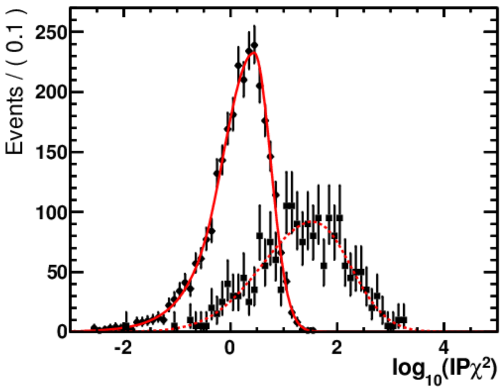
The D^0 $\log\left(IP\,\chi^2\right)$ projection of the PDF component of each species (*i.e.* prompt signal, secondary and combinatoric backgrounds) is fitted with the function, $\mathcal{A}(x)$, which comprises an asymmetric Gaussian core with asymmetric exponential tails, and is defined by

$$\mathcal{A}(x) \equiv \begin{cases} \exp\left(\frac{\rho}{2} - \frac{x-\mu}{\alpha}\rho\right) & \text{if } |x| \geq \mu + \alpha\rho \\ \exp\left(-\frac{(x-\mu))^2}{2\alpha^2}\right) & \text{if } |x| < \mu + \alpha\rho \end{cases}, \tag{5.21}$$

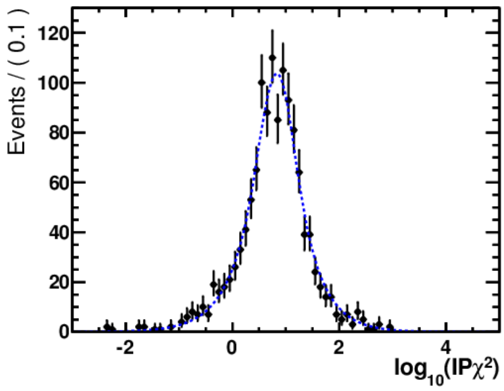
where μ and σ are the mean and width of the Gaussian core respectively. ρ and α are equal to $-\rho_{\mathrm{L}}$ and $\sigma\,(1-\epsilon)$ respectively if $x \leq \mu$, or ρ_{R} and $\sigma\,(1+\epsilon)$ respectively otherwise, where ϵ quantifies the degree of asymmetry of the Gaussian core and ρ_{L} and ρ_{R} are *tail parameters* (left and right respectively).

Due to the limited number of events, it is not possible to simultaneously fit all of the D^0 mass and $\log\left(IP\,\chi^2\right)$ parameters in every D^0 (p_{T}, y) bin. Instead, the parameters of the $\log\left(IP\,\chi^2\right)$ distribution are initially determined by fits to data using coarser bins, and fixed prior to fitting in individual bins. The exception is the mean of the Gaussian core (μ) of the $\log\left(IP\,\chi^2\right)$ projection of the prompt signal PDF which, alongside the parameters of the D^0 invariant mass projections of the PDFs, is allowed to float in the fits to the individual (p_{T}, y) bins. The prompt and secondary D^0 $\log\left(IP\,\chi^2\right)$ parameters are determined from MC, whilst the combinatoric background parameters are determined from the D^0 invariant mass side-bands in data. The *secondary fraction*, $\frac{\mathcal{N}_s}{\mathcal{N}_s+\mathcal{N}_p}$, where $\mathcal{N}_p$ and $\mathcal{N}_s$ are the number of prompt and secondary candidates respectively, is also allowed to float in the fit, but again, due to limited event numbers, this fraction is required to be the same within each group of bins used to fit the $\log(IP\,\chi^2)$ parameters.

Figure 5.7 shows the D^0 $\log(IP\,\chi^2)$ fits to the MC and data side-bands for one of the coarse bins. Figure 5.8 shows the the D^0 invariant mass and $\log(IP\,\chi^2)$ projections of the yield extraction fit to the data in a central bin, where the aforementioned D^0 $\log(IP\,\chi^2)$ parameters have been fixed from the fits shown in Fig. 5.7.



(a)



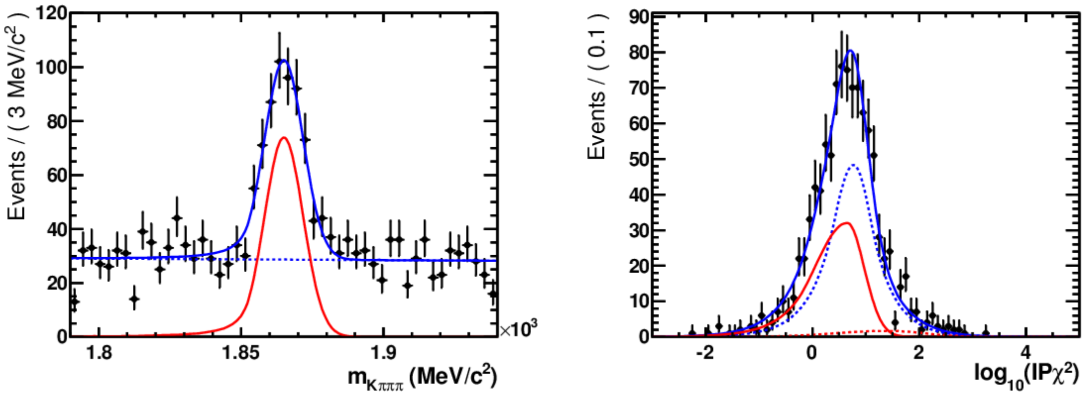
(b)

Figure 5.7: Fits to the D^0 $\log(IP\,\chi^2)$ distributions for $D^0 \to K^-\pi^+\pi^-\pi^+$ (a) MC and (b) data side-bands (*i.e.* combinatoric background) for the "coarse" kinematic bin D^0 $(p_{\mathrm{T}}, y) = (2-4\,\mathrm{MeV}/c,\ 3.0-3.5)$. In (a), the solid line shows the fit to the MC-associated prompt signal, whilst the dashed line indicates the fit to the MC-associated secondary background.

The resulting $D^0 \to K^-\pi^+\pi^-\pi^+$ prompt yields are shown in Tab. 5.11, whilst the secondary fractions are shown in Tab. 5.12; only the prompt yields are used to determine the cross-section.

p_{T} (GeV/c)	$(2.0, 2.5)$	$(2.5, 3.0)$	y $(3.0, 3.5)$	$(3.5, 4.0)$	$(4.0, 4.5)$
$(0, 1)$			46.8 ± 13.0	44.8 ± 14.7	
$(1, 2)$		60.9 ± 15.5	167.9 ± 28.1	168.6 ± 33.9	
$(2, 3)$	22.5 ± 4.8	204.6 ± 18.7	441.1 ± 33.8	291.3 ± 57.8	89.8 ± 27.7
$(3, 4)$	37.3 ± 6.4	344.7 ± 21.5	485.8 ± 30.0	366.9 ± 50.5	39.4 ± 13.7
$(4, 5)$	61.3 ± 8.2	287.3 ± 18.9	350.1 ± 23.8	213.6 ± 18.5	35.2 ± 7.8
$(5, 6)$	45.1 ± 7.7	213.3 ± 15.8	237.2 ± 18.4	107.1 ± 13.4	19.3 ± 5.5
$(6, 7)$	38.1 ± 6.4	116.8 ± 11.5	121.2 ± 13.0	61.0 ± 10.0	12.9 ± 3.8
$(7, 8)$	36.6 ± 6.1	81.6 ± 9.8	80.9 ± 10.1	33.1 ± 7.8	2.9 ± 1.7

Table 5.11: Prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ yields from fits to $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data.



(a) (b)

Figure 5.8: D^0 (a) invariant mass and (b) $\log(IP\,\chi^2)$ projections of the two-dimensional fit to the $D^0 \to K^-\pi^+\pi^-\pi^+$ data for the kinematic bin D^0 $(p_{\mathrm{T}}, y) = (2-3\,\mathrm{MeV}/c,\ 3.0-3.5)$. The solid red line in (a) represents the contribution from $D^0 \to K^-\pi^+\pi^-\pi^+$ decays (either prompt or secondary). The combinatoric background is indicated by the dashed blue line in both (a) and (b). The the prompt and secondary $D^0 \to K^-\pi^+\pi^-\pi^+$ contributions to the D^0 $\log(IP\,\chi^2)$ distribution are represented in (b) by the red solid and red dashed lines respectively. The fraction of secondary $D^0 \to K^-\pi^+\pi^-\pi^+$ decays is constrained in the fit to be the same in both this bin and the bin D^0 $(p_{\mathrm{T}}, y) = (3-4\,\mathrm{MeV}/c,\ 3.0-3.5)$. In both projections, the solid blue line indicates the sum of the signal and background contributions.

p_{T} (GeV/c)	y (2.0, 2.5)	(2.5, 3.0)	(3.0, 3.5)	(3.5, 4.0)	(4.0, 4.5)
(0, 1)			2.67 ± 5.60	2.67 ± 5.60	
(1, 2)		2.91 ± 6.79	2.67 ± 5.60	2.67 ± 5.60	
(2, 3)	6.50 ± 3.24	6.50 ± 3.24	8.14 ± 2.30	12.47 ± 3.37	12.47 ± 3.37
(3, 4)	6.50 ± 3.24	6.50 ± 3.24	8.14 ± 2.30	12.47 ± 3.37	12.47 ± 3.37
(4, 5)	3.75 ± 1.26	3.75 ± 1.26	9.79 ± 2.27	6.40 ± 2.31	6.40 ± 2.31
(5, 6)	3.75 ± 1.26	3.75 ± 1.26	9.79 ± 2.27	6.40 ± 2.31	6.40 ± 2.31
(6, 7)	3.75 ± 1.26	3.75 ± 1.26	9.79 ± 2.27	6.40 ± 2.31	6.40 ± 2.31
(7, 8)	3.75 ± 1.26	3.75 ± 1.26	9.79 ± 2.27	6.40 ± 2.31	6.40 ± 2.31

Table 5.12: $D^0 \to K^-\pi^+\pi^-\pi^+$ secondary fractions from a fit to $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data (%). The results are identical across several bins due to the coarser binning scheme used to measure the secondary fractions compared to the prompt yields.

5.8.2 Corrections for multiple candidates

As will be shown in this section, the number of events containing multiple $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates is expected to be negligible. Assuming the clone distance cuts applied offline and clone-killer algorithms applied in the reconstruction remove all clones, we would therefore expect the average number of signal candidates per event to be very close to unity. If the average number of candidates per event in the data significantly deviates from unity, then this indicates that not all clones are being rejected, and a correction factor needs to be applied to the prompt yields.

Assuming that both of the quarks from the $c\bar{c}$ pair are within LHCb acceptance, then the probability of a second $D^0 \to K^-\pi^+\pi^-\pi^+$ decay being retained after reconstruction and offline selection, $P_{c\to(D^0\to K^-\pi^+\pi^-\pi^+)}$, can be estimated by

$$P_{c\to(D^0\to K^-\pi^+\pi^-\pi^+)} = \varepsilon_{\mathrm{sel|acc}}\,\varepsilon_{\mathrm{PID|sel}}\, f\left(c \to D^0\right) \mathcal{B}\left(D^0 \to K^-\pi^+\pi^-\pi^+\right) \;, \tag{5.22}$$

where $\varepsilon_{\mathrm{sel|acc}}$ is the selection and reconstruction efficiency, $\varepsilon_{\mathrm{PID|sel}}$ is the PID efficiency and $f\left(c \to D^0\right)$ is the fraction of produced c (or $\bar{c}$) quarks that hadronise to a D^0 (or $\overline{D}{}^0$). At energies near the $\Upsilon(4S)$ resonance, $f\left(c \to D^0\right) = 0.565 \pm 0.032$ [23]. For the purposes of this calculation, it is assumed that this fraction is the same at $\sqrt{s} = 7\,\mathrm{TeV}$ (hence giving only an approximate estimate). Integrated over the D^0 p_{T} and y ranges used to determine the prompt yields, the central value of $\varepsilon_{\mathrm{sel|acc}}$ is determined from MC to be 0.010, whilst $\varepsilon_{\mathrm{PID|sel}}$ is determined from the data-driven method described in Sec. 5.7 to be 0.857. Thus, the probability of reconstructing and selecting a prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ decay is estimated to be 0.039%. Neglecting events with more than two offline-selected $D^0 \to K^-\pi^+\pi^-\pi^+$ decays [1], the average number of signal candidates per event (given at least one signal candidate is selected) is thus estimated to be $1 + P_{c\to(D^0\to K^-\pi^+\pi^-\pi^+)} = 1.00039$. The average number of truth-matched prompt $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates per event is determined to

[1] More than two $D^0 \to K^-\pi^+\pi^-\pi^+$ decays can be produced in an event due to pile-up; however, the probability of reconstructing these additional decays is much lower due to the limited LHCb acceptance. In addition, the pile-up in the data analysed is close to unity. Thus, the probability of reconstructing additional $D^0 \to K^-\pi^+\pi^-\pi^+$ decays is assumed to be negligible.

be 1.00065 ± 1.00017, which is close to the rough estimate given above.

The signal-candidate multiplicity in the data, M, is determined by

$$M = \frac{c_s - c_b}{e_s - e_b} \ , \tag{5.23}$$

where c_s and c_b are the number of reconstructed $D^0 \to K^-\pi^+\pi^-\pi^+$ decays in the data inside the signal mass window and side-band regions of the D^0 invariant-mass distribution respectively, and e_s and e_b are the number of events in the same regions. The mass window and side-bands are chosen such that their total ranges are equal and the signal yield in the side-bands is expected to be negligible. Since the combinatoric background shape is assumed to be linear in D^0 invariant mass, this procedure is effectively a linear background subtraction.

The statistical uncertainty of M is determined by

$$\sigma\left(M\right) = \frac{\sqrt{c_s + c_b}}{e_s - e_b} \ . \tag{5.24}$$

Using Eq. (5.23) and Eq. (5.24), the signal-candidate multiplicity over the full (p_{T}, y) range is determined to be 1.040 ± 0.041. The value of M is in addition determined in each (p_{T}, y) bin to investigate any possible variation with p_{T} and/or y. The results in the individual bins are compared with the above global result in Fig. 5.9. These results are consistent with the global multiplicity within statistical uncertainties.

The central value of the global result for the signal-candidate multiplicity is much larger than expected, although the result is consistent with expectations within statistical uncertainties. Nevertheless, we have decided to apply this result as a correction factor to each of the binned prompt yields, and assign the statistical uncertainty on the multiplicity as a source of systematic uncertainty.

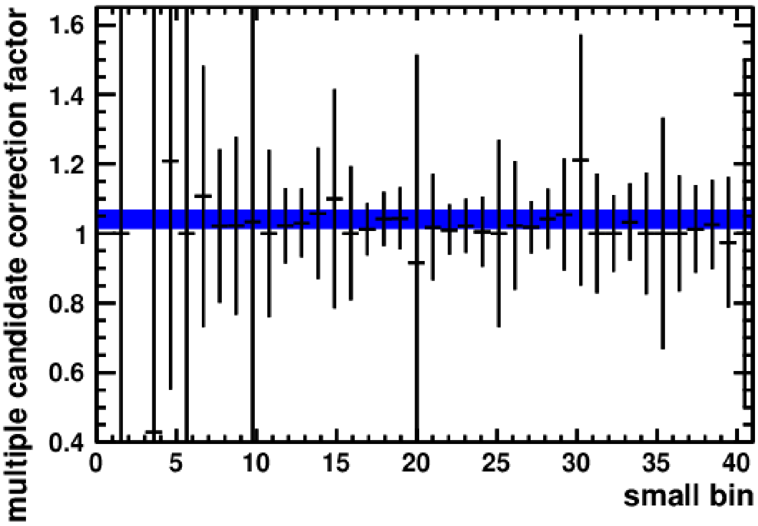


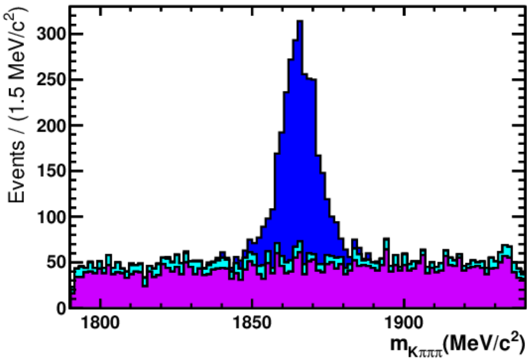
Figure 5.9: Signal-candidate multiplicity versus (p_{T}, y) bin number. The (p_{T}, y) bin number (denoted *small bin*) is determined by $i + n \cdot j$, where i is the p_{T} bin number (0 for the bin with the lowest p_{T} range, 1 for the second lowest and so on), j is the y bin number (using similar indexing to the p_{T} bins) and n is the number of p_{T} bins. The blue band indicates $\pm 1\sigma$ from the central value of the signal-candidate multiplicity determined using data from the full (p_{T}, y) range.

5.8.3 Corrections for physics background

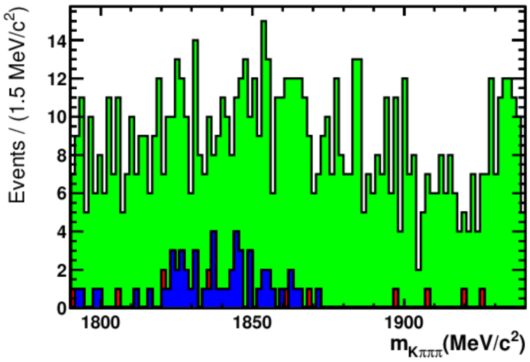
The yield extraction method described in Sec. 5.8.1 assumes that the only sources of background are combinatoric and secondary $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. However, it is also possible to reconstruct "physics backgrounds", *i.e.* genuine decays (as opposed to the random combination of tracks) other than $D^0 \to K^-\pi^+\pi^-\pi^+$. Some of these backgrounds may peak at the D^0 invariant mass, or have tails within the signal mass window, leading to an overestimation of the prompt yield. To determine if any such backgrounds are retained after applying the offline selection cuts, samples of magnet up and down inclusive $c\bar{c}$ MC are analysed. These samples were listed in Sec. 5.4.2.

The inclusive $c\bar{c}$ decays in MC retained after offline selection are shown in Fig. 5.10a. According to the LHCb background classification tool [92], the physics backgrounds can be further split into *reflections*, *low mass* and *partially reconstructed* backgrounds.

As can be seen in Fig. 5.10b, the level of reflection background is very low, and the partially reconstructed backgrounds are distributed relatively uniformly within the signal



(a)



(b)

Figure 5.10: Histograms showing the composition of inclusive $c\bar{c}$ MC decays reconstructed as $D^0 \to K^-\pi^+\pi^-\pi^+$ passing the offline selection according to the LHCb background classification tool. In (a), the MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ decays (blue), combinatoric background (magenta) and physics backgrounds (cyan) are shown. In (b), the physics background decays are sub-divided into partially reconstructed (green), low mass (blue) and reflection backgrounds (red). No attempt has been made to distinguish between prompt and secondary $D^0 \to K^-\pi^+\pi^-\pi^+$ decays in (a).

mass window. However, the low mass backgrounds do peak in D^0 invariant mass and impinge on the lower edge of the signal peak. According to the MC truth information, only two of the 55 decays classified as low mass are in fact correctly classified. These correspond to the decays $D^0 \to \overline{K}^{*0}\eta$, where $\overline{K}^{*0} \to K^-\pi^+$ and $\eta \to \pi^+\pi^-\gamma$, and $D^0 \to K^-\pi^+\eta'$, where $\eta' \to \rho^0\gamma$. The remaining 53 decays correspond to fully-reconstructed $D^0 \to K^-\pi^+\pi^-\pi^+$ decays where one of the final-state particles has radiated one or more photons. The LHCb background classification tool includes a tunable maximum threshold on the photon momenta (which defaults to 300 MeV/c) in order to avoid associating photons from π^0 decays as signal. The photons in these cases exceed the threshold, so are they are wrongly associated as background.

Since the level of peaking physics background is negligible, no correction factor is assigned to the prompt yields. However, the fact that some signal candidates have been wrongly assigned as background means that a correction factor needs to be assigned to the reconstruction and selection efficiency. The correction factor due to signal candidates in MC mis-associated as low-mass background, P, is given by

$$P = \frac{N_s + N_b}{N_s} \;, \tag{5.25}$$

where N_s is the number of MC-associated $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates in the inclusive $c\bar{c}$ sample and N_b is the number of $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates in the sample that are wrongly associated as low mass background. The statistical uncertainty on P is given by

$$\sigma(P) = \sqrt{\frac{N_b(N_s + N_b)}{N_s^3}} \, . \tag{5.26}$$

From Eqs. (5.25) and (5.26) the correction factor is determined to be 1.022 ± 0.003. The statistical uncertainty on this correction factor is assigned as a source of systematic uncertainty on the cross-section.

5.9 Systematic uncertainties

Including all correction factors, the D^0 production cross-section in a given (p_{T}, y) bin i is determined by

$$\sigma_i\left(D^0\right) = \frac{N_i\left(D^0 \to K^-\pi^+\pi^-\pi^+\right) \cdot F \cdot M \cdot P \cdot C}{\varepsilon_{i,\mathrm{acc}} \cdot \varepsilon_{i,\mathrm{sel|acc}} \cdot \varepsilon_{i,\mathrm{PID|sel}} \cdot \mathcal{B}\left(D^0 \to K^-\pi^+\pi^-\pi^+\right) \cdot \mathcal{L}_{\mathrm{eff}}} \, , \tag{5.27}$$

where F, M, P and C are the correction factors for MC-association failure, multiple candidates, peaking backgrounds and data/MC disagreement respectively, expressed as multiplicative correction factors.

The statistical uncertainties on the cross-section measurements are determined by the (propagated) statistical uncertainties on the prompt yields $N_i\left(D^0 \to K^-\pi^+\pi^-\pi^+\right)$. In addition, each of the parameters in Eq. (5.27) has an associated systematic uncertainty. The methods for determining many of these systematic uncertainties have already been discussed in previous sections. The methods for determining the remaining uncertainties are discussed in this section. A summary of all sources of systematic uncertainty follows. Where *relative* uncertainties are quoted, this means that the systematic uncertainty associated with a quantity has been divided by the (central) value of the quantity itself.

5.9.1 Yield extraction

The systematic uncertainty from the yield extraction is determined by performing two alternative fits in each (p_{T}, y) bin.

The first alternative yield method is identical to the yield extraction method described in Sec. 5.8.1, but using different PDFs to model the signal and backgrounds. In the alternative fit, the D^0 invariant-mass distributions of the prompt and secondary candidates are parametrised by a Gaussian in place of a Crystal Ball, and the $\log(IP\,\chi^2)$ distributions are parametrised by *Bukin* PDFs, defined in Appendix A.4.

In the second method, the data are fit using an alternative yield extraction. In this fit method, a binned D^0 invariant mass side-band subtraction is performed, and the $\log(IP\,\chi^2)$ distribution of the side-band-subtracted data is used distinguish between prompt and secondary $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. The $\log(IP\,\chi^2)$ distributions of prompt and secondary decays are parametrised by *Bukin* PDFs, and the data are *adaptively binned*, meaning that bin widths are automatically adjusted so that at least two $D^0 \to K^-\pi^+\pi^-\pi^+$ candidates lie in each $\log(IP\,\chi^2)$ bin.

The largest deviation between the main fits and the two alternative fits is assigned as the systematic uncertainty in the bin. These systematic uncertainties are shown in Tab. 5.13.

p_{T} (GeV/c)			y		
	$(2.0, 2.5)$	$(2.5, 3.0)$	$(3.0, 3.5)$	$(3.5, 4.0)$	$(4.0, 4.5)$
$(0, 1)$			11.63	35.49	
$(1, 2)$		19.66	2.60	8.24	
$(2, 3)$	6.38	2.51	8.71	11.90	27.50
$(3, 4)$	6.90	2.81	6.02	9.53	19.95
$(4, 5)$	16.37	8.90	8.16	2.40	4.06
$(5, 6)$	11.34	7.33	8.92	1.04	25.08
$(6, 7)$	9.09	12.13	6.66	1.04	26.07
$(7, 8)$	6.98	3.99	14.16	5.48	19.13

Table 5.13: Relative systematic uncertainties associated with the yield extraction methods (%).

5.9.2 Dalitz model

To assess the systematic uncertainty associated with the $D^0 \to K^-\pi^+\pi^-\pi^+$ amplitude model in the efficiency determination, various alternative amplitude models have been generated where the phases are randomly varied between $\pm\pi$, and the magnitudes are varied between a factor 0.5 and 2 of the nominal values shown in Tab. 5.3. The weights resulting from these alternative models are then used to re-calculate the acceptance, reconstruction and selection, and PID efficiencies. These efficiencies are then used to re-determine the cross-section, and the standard deviation from the alternative models in a given (p_T, y) bin is assigned as the systematic uncertainty associated with the Dalitz model in that bin, shown in Tab. 5.14.

p_T			y		
(GeV/c)	$(2.0, 2.5)$	$(2.5, 3.0)$	$(3.0, 3.5)$	$(3.5, 4.0)$	$(4.0, 4.5)$
$(0, 1)$			8.77	10.98	
$(1, 2)$		1.79	3.35	2.61	
$(2, 3)$	6.99	3.40	2.11	2.44	2.82
$(3, 4)$	3.61	2.72	2.06	1.63	1.28
$(4, 5)$	3.93	1.76	2.17	3.34	2.84
$(5, 6)$	3.21	2.02	2.67	2.25	3.44
$(6, 7)$	3.50	1.17	1.50	1.69	8.17
$(7, 8)$	4.49	2.27	2.72	3.35	3.88

Table 5.14: Relative systematic uncertainties associated with the $D^0 \to K^-\pi^+\pi^-\pi^+$ amplitude model (%).

5.9.3 Long track reconstruction efficiency

The systematic uncertainty associated with the long-track reconstruction efficiency is 3% according to a study of the ratio of the rates of $D^0 \to K^-\pi^+\pi^-\pi^+$ and $D^0 \to K^-\pi^+$ decays in MC and data [33]. This uncertainty is treated as uncorrelated between the D^0 final-state particles. Thus, a relative tracking-efficiency uncertainty of 12% is assigned, which is applied uniformly across all bins.

5.9.4 Branching fraction

The global average of the branching fraction for $D^0 \to K^-\pi^+\pi^-\pi^+$ determined by the PDG is $\left(8.07^{+0.21}_{-0.19}\right)$ % [23]. The average of the asymmetric uncertainties is assigned as a symmetric systematic uncertainty.

5.9.5 Summary

A summary of the correction factors and systematic uncertainties is given in Tab. 5.15.

Measurement	Correction factor	Systematic uncertainty
Yield extraction		1.0 – 35.5
$\varepsilon_{\mathrm{acc}}$		0.0 – 3.2
$\varepsilon_{\mathrm{sel}\|\mathrm{acc}}$		0.2 – 2.9
$\varepsilon_{\mathrm{PID}\|\mathrm{sel}}$		1.9 – 12.0
Dalitz model		1.2 – 11.0
Multiple candidates	0.961	2.74
MC association	0.966	0.003
Peaking background	0.978	0.03
Cut-variable distributions	1.302	11.46
Tracking efficiency		12.00
Integrated luminosity		3.50
Branching fraction		2.47

Table 5.15: Summary of correction factors and relative systematic uncertainties for the D^0 production cross-section measurements. Correction factors are expressed as fractions, whilst systematic uncertainties are expressed in %. The top section of the table shows the range of systematic uncertainties determined in (p_{T}, y) bins. The middle section shows the correction factors to $\varepsilon_{\mathrm{sel|acc}}$ or the raw yields and their associated uncertainties, which are applied uniformly across the (p_{T}, y) bins. The final section shows the systematic uncertainties corresponding to parameters which are determined from external measurements, which are also applied uniformly across all bins.

5.10 Results

The central values, statistical and systematic uncertainties of the D^0 production cross-section measurements in D^0 (p_{T}, y) bins are given in Tab. 5.16.

p_{T}			y		
(GeV/c)	$(2.0, 2.5)$	$(2.5, 3.0)$	$(3.0, 3.5)$	$(3.5, 4.0)$	$(4.0, 4.5)$
$(0, 1)$			$240.3 \pm 63.5 \pm 58.9$	$278.9 \pm 97.0 \pm 118.8$	
$(1, 2)$		$192.3 \pm 54.3 \pm 52.7$	$119.2 \pm 18.8 \pm 22.5$	$147.2 \pm 24.9 \pm 30.0$	
$(2, 3)$	$145.2 \pm 32.1 \pm 32.3$	$70.2 \pm 6.9 \pm 13.6$	$68.7 \pm 5.3 \pm 14.0$	$47.4 \pm 5.2 \pm 10.3$	$68.6 \pm 13.6 \pm 23.1$
$(3, 4)$	$38.1 \pm 6.6 \pm 8.2$	$37.9 \pm 2.7 \pm 7.1$	$30.1 \pm 1.8 \pm 5.7$	$24.9 \pm 2.0 \pm 5.0$	$8.7 \pm 2.0 \pm 2.4$
$(4, 5)$	$17.6 \pm 2.3 \pm 4.5$	$17.9 \pm 1.2 \pm 3.5$	$13.2 \pm 0.9 \pm 2.6$	$10.2 \pm 0.8 \pm 1.8$	$5.1 \pm 1.1 \pm 1.1$
$(5, 6)$	$9.5 \pm 1.6 \pm 2.1$	$9.5 \pm 0.7 \pm 1.8$	$6.3 \pm 0.5 \pm 1.2$	$3.8 \pm 0.4 \pm 0.7$	$2.2 \pm 0.6 \pm 0.7$
$(6, 7)$	$3.8 \pm 0.6 \pm 0.8$	$3.9 \pm 0.4 \pm 0.8$	$2.9 \pm 0.3 \pm 0.6$	$2.0 \pm 0.3 \pm 0.4$	$1.7 \pm 0.5 \pm 0.6$
$(7, 8)$	$2.8 \pm 0.5 \pm 0.6$	$2.4 \pm 0.3 \pm 0.4$	$1.8 \pm 0.2 \pm 0.4$	$1.0 \pm 0.2 \pm 0.2$	$0.5 \pm 0.3 \pm 0.1$

Table 5.16: D^0 direct production cross-section results measured in D^0 (p_{T}, y) bins, determined from $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. The cross-sections are expressed in units of μb.

5.10.1 Comparison with results from $D^0 \to K^-\pi^+$

By way of cross-check, the D^0 production cross-section has also been calculated in a separate analysis [93] using the more numerous $D^0 \to K^-\pi^+$ mode, and utilising the same methods for determining the raw yields and PID efficiencies. The differences between the two cross-section results are determined in each (p_{T}, y) bin for which both analyses are able to measure the cross-section. These differences are then divided by the quadrature sum of the statistical uncertainties and systematic uncertainties (systematic uncertainties that are correlated between the two analyses such as luminosity and track efficiency are *not* considered) to determine the pull distribution. If the results are consistent, then the pulls will be distributed according to a unit Gaussian. The pulls are shown in Fig. 5.11, in which a χ^2 fit to a Gaussian PDF has been performed. The mean and width parameters of the fit function correspond to -0.007 ± 0.162 and 0.929 ± 0.114 respectively with a p-value of 0.992. Hence, the results agree well with those determined from the $D^0 \to K^-\pi^+$ analysis. The central value of the pull width is less than unity, which may indicate that the systematic uncertainties assigned are overly conservative; however, the width is consistent with unity within uncertainties.

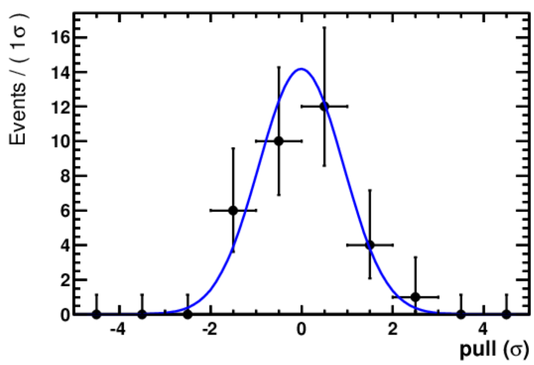


Figure 5.11: The pull distribution between the D^0 production cross-section results determined from the decay modes $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^-\pi^+$. The blue line shows the result of a Gaussian fit.

5.10.2 Comparison with theoretical predictions

Theoretical predictions for the D^0 production cross-section have been produced by two independent theory groups using the GM-VFNS [94] and FONNL [95] approaches. As discussed in Sec. 5.3, these methods are able to predict the production cross-section over a wide range of p_{T} values. The theoretical predictions from both methods have been provided by the theory groups as differential cross-sections calculated as a function of p_{T} and integrated over y, using the same bins of rapidity as Tab. 5.16. To compare the experimental results with theory, the theoretical predictions have been integrated over the p_{T} bin ranges and divided by the bin width (1 GeV/c). The integration is performed numerically using a cubic spline interpolation of $\frac{1}{p_{\mathrm{T}}}\frac{d\sigma}{dp_{\mathrm{T}}}$, where the weight factor $1/p_{\mathrm{T}}$ is included to stabilise the interpolation in the low p_{T} region, where the cross-section varies most rapidly. The theoretical predictions for the two methods are shown in Tabs. 5.17 and 5.18.

p_{T} (GeV/c)	y (2.0, 2.5)	(2.5, 3.0)	(3.0, 3.5)	(3.5, 4.0)	(4.0, 4.5)
(3,4)	64^{+43}_{-40}	57^{+38}_{-35}	48^{+31}_{-29}	36^{+34}_{-20}	
(4,5)	24^{+12}_{-10}	21^{+10}_{-8}	17^{+8}_{-6}	12^{+6}_{-4}	$8.9^{+5.0}_{-3.2}$
(5,6)	11^{+5}_{-4}	$9.3^{+3.8}_{-3.1}$	$7.3^{+3.0}_{-2.4}$	$5.4^{+2.3}_{-1.7}$	$3.4^{+1.0}_{-1.0}$
(6,7)	$5.5^{+1.7}_{-1.2}$	$4.6^{+1.4}_{-1.0}$	$3.5^{+1.1}_{-0.7}$	$2.5^{+0.7}_{-0.5}$	$1.5^{+0.6}_{-0.2}$
(7,8)	$3.0^{+0.8}_{-0.6}$	$2.4^{+0.7}_{-0.5}$	$1.8^{+0.5}_{-0.3}$	$1.2^{+0.3}_{-0.2}$	$0.71^{+0.18}_{-0.10}$

Table 5.17: GM-VNFS theoretical predictions for the D^0 production differential cross-section, $d\sigma/dp_{\mathrm{T}}$. The cross-sections are expressed in units of μb/ GeV/c.

p_{T} (GeV/c)	y (2.0, 2.5)	(2.5, 3.0)	(3.0, 3.5)	(3.5, 4.0)	(4.0, 4.5)
(0,1)	62^{+119}_{-41}	61^{+118}_{-39}	59^{+118}_{-36}	56^{+115}_{-33}	50^{+107}_{-29}
(1,2)	102^{+142}_{-84}	96^{+134}_{-78}	88^{+124}_{-70}	78^{+110}_{-60}	65^{+94}_{-50}
(2,3)	59^{+63}_{-41}	53^{+57}_{-37}	47^{+50}_{-32}	39^{+42}_{-26}	30^{+33}_{-20}
(3,4)	27^{+23}_{-15}	24^{+20}_{-13}	20^{+17}_{-10}	16^{+14}_{-8}	12^{+10}_{-6}
(4,5)	13^{+9}_{-5}	11^{+8}_{-5}	$8.9^{+6.4}_{-3.6}$	$6.8^{+4.9}_{-2.7}$	$4.7^{+3.5}_{-1.8}$
(5,6)	$6.2^{+3.8}_{-2.2}$	$5.3^{+3.2}_{-1.8}$	$4.2^{+2.6}_{-1.4}$	$3.1^{+1.9}_{-1.0}$	$2.0^{+1.3}_{-0.7}$
(6,7)	$3.3^{+1.8}_{-1.0}$	$2.7^{+1.5}_{-0.8}$	$2.1^{+1.2}_{-0.7}$	$1.5^{+0.8}_{-0.5}$	$0.94^{+0.53}_{-0.30}$
(7,8)	$1.9^{+0.9}_{-0.5}$	$1.5^{+0.7}_{-0.4}$	$1.1^{+0.6}_{-0.3}$	$0.78^{+0.39}_{-0.23}$	$0.47^{+0.24}_{-0.14}$

Table 5.18: FONNL theoretical predictions for the D^0 production differential cross-section, $d\sigma/dp_{\mathrm{T}}$. The cross-sections are expressed in units of μb/ GeV/c.

The FONNL calculations include theoretical uncertainties due to the charm quark mass, m_c, and the *renormalisation* and *factorisation scales*, μ_R and μ_F respectively. Both scales are varied independently in the range $\frac{1}{2}m_T < \mu_{R,F} < 2m_T$ where $m_T = \sqrt{m_c^2 + p_{\mathrm{T}}^2}$, with the constraint $0.5 < \frac{\mu_R}{\mu_F} < 2$. Uncertainties due to the choice of parton density functions[1], which are expected to be small, are neglected. Also neglected are theoretical uncertainties due to higher-order QCD effects, which can be estimated from comparing the results with predictions from parton shower MC simulation. The theoretical uncertainties are expected to be small, except in the soft-collinear region (low p_{T}, high y), so the uncertainties may be underestimated in this region. Since we have low event numbers and high backgrounds in this region and are therefore unable to measure the cross-section, this has no bearing on the comparison of theory and experimental results. Note that the FONNL predictions provided represent the total charm cross-section and have been converted into the D^0 production cross-section by multiplying their values by the fraction of charm quarks that hadronise to a D^0, $f\left(c \to D^0\right) = 0.565 \pm 0.032$. The uncertainty on this fraction is small compared to other sources of theoretical uncertainty, so is not propagated to the final result.

The GM-VFNS predictions represent the D^0 production cross-section, which were determined by convolution of the parton density functions[2] and the short-distance cross-section with the D^0 charm fragmentation function, $D_c^{D^0}(x, p_{\mathrm{T}}^2)$, as defined in Sec. 5.3.2. The fragmentation function was determined from fits to LEP experimental data [98].

The binned differential cross-section results as a function of p_{T} for each of the rapidity bins are shown in Fig. 5.12. The theoretical predictions in Tabs. 5.17 and 5.18 are also shown, with error bands corresponding to the theoretical uncertainties. The binned predictions are shown as smooth curves by interpolating the values in the corresponding table, with the value of the curves at the centre of each p_{T} bin corresponding to the tabulated result in the table. Results are shown for GM-VFNS only for $p_{\mathrm{T}} > 3\,\mathrm{GeV}/c$ ($p_{\mathrm{T}} > 4\,\mathrm{GeV}/c$ in the largest rapidity bin), as the theoretical uncertainties increase significantly at lower p_{T} values. In addition to GM-VFNS predictions using standard parton density function parametrisations,

[1] The parton density functions used correspond to the CTEQ 6.6 parametrisation [96].
[2] CTEQ 6.5 parametrisation [97].

predictions have been calculated using parton density functions that include modelling of intrinsic charm[1], and the (interpolated) central values are also shown in Fig. 5.12. As well as theoretical models, Fig. 5.12 shows MC predictions for the differential cross-sections according to the LHCb default tuning of the PYTHIA event generator.

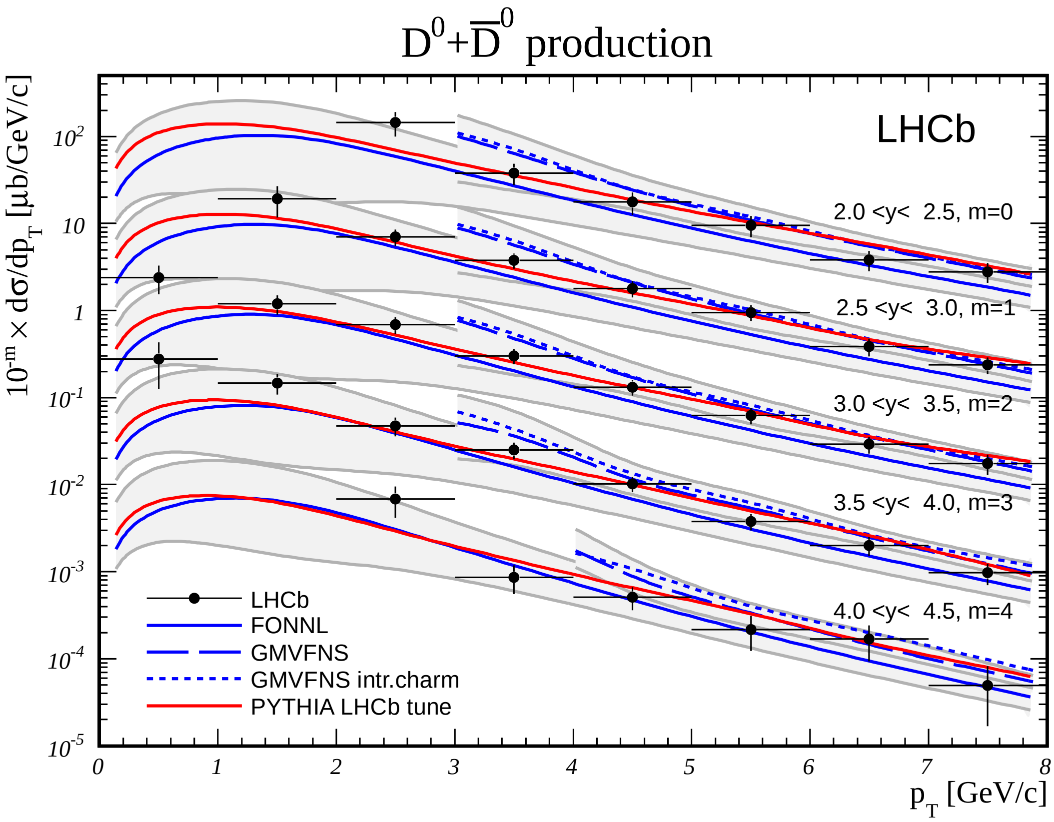


Figure 5.12: Plot showing the D^0 production cross-section, along with predictions from theoretical and MC models. The theoretical uncertainties for the GM-VFNS and FONNL predictions are shown as solid error bands, with the dark grey lines corresponding to the upper and lower limits of the (68%) confidence intervals of the two predictions. The plot was produced by the author, using code provided by Michael Schmelling (Max-Planck-Institut für Kernphysik).

5.10.3 Summary and discussion

Using $D^0 \to K^-\pi^+\pi^-\pi^+$ decays from $\mathcal{L}_{\rm int} = 14.96\ \mathrm{nb}^{-1}$ of LHCb pp collision data collected during 2010, the D^0 production cross-section has been measured in bins of $p_{\rm T}$ and y of the D^0 meson. Excellent agreement is found between these results and those from the more copious $D^0 \to K^-\pi^+$ decay mode. The results have also been compared to predictions from

[1] CTEQ 6.5c2 parametrisation [99].

the GM-VFNS and FONNL theoretical models, and the LHCb default tuning of Pythia. The measured cross-section results are consistent with both theoretical models within their uncertainties, and show very good agreement with the MC predictions in most bins (though there is some deviation at small p_{T}, where theoretical uncertainties are large). In particular, the trend of the experimental results to favour the FONNL model within the mid-range of p_{T} values considered, and to favour the GM-VFNS at high p_{T} supports the validity of the Pythia simulation.

A point of theoretical interest is the effect of intrinsic charm on the production cross-section of charmed hadrons. Several parton density function models that incorporate intrinsic charm show large enhancements of the production cross-section from the standard GM-VFNS predictions at high rapidity [77]. The predictions from one such model have been considered here, and within the phase-space region accessible to LHCb, the effect of intrinsic charm is small.

Chapter 6

Conclusions

Proton-proton collision data at $\sqrt{s} = 7\,\text{TeV}$ has been collected by the LHCb experiment and has been used to determine the ratio of branching fractions, $\Gamma(B \to DK)/\Gamma(B \to D\pi)$, using $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ and $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D \pi^{\pm}$ decays. This ratio was previously measured by the BELLE [69] and BABAR [70] experiments, with the results differing by around 2σ. The analysis described in this thesis determines the ratio of branching fractions to be $0.057 \pm 0.007 \pm 0.004$, a result which is consistent within uncertainties with that measured in the $D \to K^{\pm}\pi^{\mp}$ channel [3]. Whilst the $D \to K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}$ channel has approximately a factor two larger branching fraction than $D \to K^{\pm}\pi^{\mp}$, the reconstruction efficiency of the former is significantly lower due to the additional tracks. The excellent vertex resolution and tracking capabilities of LHCb gives a good indication that LHCb will be capable of measuring γ from suppressed multi-body D decay channels in the future.

The rate asymmetry between the charged states $B^{-} \to (K^{-}\pi^{+}\pi^{-}\pi^{+})_D K^{-}$ and $B^{+} \to (K^{+}\pi^{-}\pi^{+}\pi^{-})_D K^{+}$ due to CP violation, $A^{K3\pi}_{fav} = \frac{\Gamma(B^{-})-\Gamma(B^{+})}{\Gamma(B^{-})+\Gamma(B^{+})}$ has also been measured. The result of $-0.01 \pm 0.11 \pm 0.07$ is consistent with zero within uncertainties, which is as expected since the favoured $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$ channel contains very little sensitivity to γ compared to the suppressed $B^{\pm} \to (K^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_D K^{\pm}$ channel. This measurement is an important step towards a measurement of the CKM angle γ, since the method of determining the rate asymmetry will be similar to that employed for the suppressed decays. Furthermore, the two types of decays share many of the same backgrounds.

The rate asymmetry measurement is currently statistics dominated. However, there is also a large systematic uncertainty due to the fixing of the production and kaon-detection asymmetries in the fit. With more events, it will be possible for one or both of these parameters to be measured in the fit, reducing the overall systematic uncertainty.

The D^0 production cross-section has been measured at $\sqrt{s} = 7\,\mathrm{TeV}$ in bins of p_T and y using Cabibbo favoured $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. The results are shown to be consistent with predictions from Monte Carlo (MC), the results from an independent measurement using $D^0 \to K^-\pi^+$ decays, and several of the theoretical models.

The cross-section measurements are systematics-dominated in the majority of bins. This is as expected since the cross-section is determined from absolute yields rather than a ratio; hence, many sources of systematics do not cancel. Two sources of uncertainty that are significant in all bins are those due to the uncertainty on the track reconstruction efficiency and differences between the distributions of cut variables in data and MC. It is likely that the level of systematic uncertainty due to data/MC differences can be reduced in future by relaxing or removing cuts in the stripping and offline selections. A future LHCb analysis of the D^0 production cross-section will also benefit from a much lower tracking-efficiency systematic; improvements have already been made after the data presented in this thesis were reconstructed and analysed [100].

Essential to all measurements presented in this thesis is the excellent particle identification (PID) performance provided by the RICH detectors, principally the ability to distinguish between kaons and pions with high efficiency and with a low mis-identification fraction. The branching fraction for $B \to D\pi$ is an order of magnitude larger than $B \to DK$, and without PID information it would be practically impossible to resolve $B \to DK$ decays from the $B \to D\pi$ background. An important part of the operation of the RICH detectors is the time alignment of the front-end Level-0 (L0) boards. These boards control the read-out and calibration of the HPD photon detectors, and have programmable timing delays that must be synchronised across the RICH system. The time alignment which was performed shortly before the first *pp* collisions has been described, using a pulsed laser system installed in the LHCb cavern. The L0 boards were aligned to *pp* collision data by adjusting the delays of the L0 boards, determined by the laser scans, in order to determine the working point corresponding to the maximum number of detector hits. The final results of the timing scan indicate that alignment has been successfully achieved to within around ± 1 ns across all L0 boards, hence maximising the RICH PID efficiencies.

Appendix A

List of Probability Density Functions

Various Probability Density Functions (PDFs) are used throughout this thesis. The less common, and non-standard functions are outlined below.

A.1 Crystal Ball

The *Crystal Ball* function ($\mathcal{C}(x)$) comprises a Gaussian core with a power-law tail. The PDF has four free parameters: α, n, μ and σ. Quantities μ and σ are the Gaussian mean and width, α represents the threshold at which the power-law tail comes into effect, and n is the power of the tail portion of the PDF. The power is often fixed prior to fitting. If α is positive, then the radiative tail will be to the left of the distribution, whereas if α is negative, it will lie to the right.

The Crystal Ball function is defined by

$$\mathcal{C}(x) \equiv \begin{cases} N \exp\left(-\frac{1}{2}t^2\right) & \text{if } t \geq -|\alpha| \\ \dfrac{A}{(B-t)^2} & \text{if } t < -|\alpha| \end{cases}, \tag{A.1}$$

where t, A and B are defined by

$$\begin{aligned} t &\equiv \operatorname{sgn}(\alpha) \cdot \tfrac{x-\mu}{\sigma} \\ A &\equiv \left(\tfrac{n}{|\alpha|}\right)^{n} \exp\left(-\tfrac{1}{2}\alpha^{2}\right) \\ B &\equiv \tfrac{n}{|\alpha|} - |\alpha| \end{aligned} \quad ,$$

and N is a normalisation factor, which ensures that the integral of the PDF over the full data range is equal to unity. Such a normalisation factor is implied in all subsequently defined PDFs.

A.2 Cruijff

The *Cruijff* function $\mathcal{J}(x)$ is a modified Gaussian function with tail parameters on both sides. As with the Crystal Ball function, the PDF has four free parameters: μ, σ, α_L and α_R. Again, μ and σ are the mean and width of the Gaussian core, and α_L and α_R are the left and right tail parameters respectively. The Cruijff PDF is given by

$$\mathcal{J}(x) \equiv \exp\left(-\frac{(x-\mu)^{2}}{2\sigma^{2} + \alpha (x-\mu)^{2}}\right) , \tag{A.2}$$

where α is defined by

$$\alpha \equiv \begin{cases} \alpha_L & \text{if } x \leq \mu \\ \alpha_R & \text{if } x > \mu \end{cases} .$$

A.3 Novosibirsk

The *Novosibirsk* function, $\mathcal{N}(x)$ has a peaking distribution with an asymmetric tail. It has three free parameters: μ, σ and α. Parameter μ is the peak position, σ represents the width of the peak, and α is a tail parameter. The Novosibirsk function is defined by

$$\mathcal{N}(x) \equiv \exp\left(-\frac{1}{2}\left(\frac{\ln^{2}\Lambda}{\alpha^{2}} + \alpha^{2}\right)\right) , \tag{A.3}$$

where Λ is defined by

$$\Lambda = 1 + \frac{\alpha\left(x-\mu\right)}{\sigma}\frac{\sinh\left(\alpha\sqrt{\ln 4}\right)}{\alpha\sqrt{\ln 4}} \; .$$

A.4 Bukin

The *Bukin* function, $\mathcal{B}\left(x\right)$ is a modified version of the Novosibirsk function with an asymmetric peak. It has five free parameters: μ, σ, ξ, ρ_{L} and ρ_{R}. As with the Novosibirsk function, μ and σ represent the value at the peak summit and the peak width respectively, with ξ representing the asymmetry of the peak. Quantities ρ_{L} and ρ_{R} are the left and right tail parameters. The Bukin function is defined by

$$\mathcal{B}\left(x\right) \equiv \begin{cases} \exp\left(\dfrac{\xi\sqrt{\xi^2+1}\left(x-x_{\mathrm{L}}\right)\sqrt{2\ln 2}}{\sigma\left(\sqrt{\xi^2+1}-\xi\right)^2\ln\left(\sqrt{\xi^2+1}+\xi\right)} + \rho_{\mathrm{L}}\left(\frac{x-x_{\mathrm{L}}}{\mu-x_{\mathrm{L}}}\right)^2 - \ln 2\right) & \text{if } x < x_{\mathrm{L}} \\ \exp\left(\dfrac{\xi\sqrt{\xi^2+1}\left(x-x_{\mathrm{R}}\right)\sqrt{2\ln 2}}{\sigma\left(\sqrt{\xi^2+1}-\xi\right)^2\ln\left(\sqrt{\xi^2+1}+\xi\right)} + \rho_{\mathrm{R}}\left(\frac{x-x_{\mathrm{R}}}{\mu-x_{\mathrm{R}}}\right)^2 - \ln 2\right) & \text{if } x > x_{\mathrm{R}} \\ \exp\left(-\ln 2\left(\dfrac{\ln\left(1+2\xi\sqrt{\xi^2+1}\frac{x-\mu}{\sigma\sqrt{2\ln 2}}\right)}{\ln\left(1+2\xi^2-2\xi\sqrt{\xi^2+1}\right)}\right)^2\right) & \text{otherwise} \end{cases} , \tag{A.4}$$

where x_{L} and x_{R} are defined by

$$x_{\mathrm{L,R}} = \mu + \sigma\sqrt{2\ln 2}\left(\frac{\xi}{\sqrt{\xi^2+1}} \mp 1\right) \; .$$

A.5 Chebychev polynomials

The Chebychev polynomials are a sequence of orthogonal polynomials. Technically, Chebychev polynomials can refer to two related sets of orthogonal polynomials: *Chebychev polynomials of the first kind* ($T_n\left(x\right)$), and *Chebychev polynomials of the second kind* ($T_n\left(x\right)$).

The fits in this thesis were performed using the RooFit software [73], which uses Chebychev polynomials of the first kind, which for brevity are referred to simply as *Chebychev polynomials*.

The Chebychev polynomials are defined by the following recurrence relation:

$$\begin{aligned} T_0(x) &\equiv 1 \\ T_1(x) &\equiv x \\ T_{n+1}(x) &\equiv 2\,x\,T_n(x) - T_{n-1}(x) \end{aligned} \quad , \tag{A.5}$$

where the index represents the order of the sequence. The reason for using Chebychev polynomials over simple polynomials is that they tend to lead to more stable fits.

A.6 Adaptive Kernel Density Estimation

Kernel density estimation is a model-independent method of determining the estimator of a PDF drawn from a set of data points. The kernel density estimate (KDE) for a univariate distribution is given by

$$\hat{f}_0(x) \equiv \frac{1}{nh}\sum_{i=0}^{n}\mathcal{K}\left(\frac{x-t_i}{h}\right) \quad , \tag{A.6}$$

where t_i is the value of the observable x for data point i, and h is the *bandwidth* (a smoothing parameter). The kernel function, $\mathcal{K}(x)$ employed is arbitrary, but since its role is to smooth out the contribution of each data point in the PDF estimator, a sensible choice is a *standard* Gaussian (a Gaussian with a mean of zero, and a width of 1). This kernel function is advantageous for many physics uses, since it is positive definite, and ensures that $\hat{f}_0(x)$ is smooth and well-behaved in the tails of the distribution [101].

The choice of the bandwidth is crucial to obtain a good quality estimate of the PDF. If the bandwidth is too small, then the estimate will be too sensitive to statistical fluctuations, whilst if the bandwidth is too large, then the resulting distribution can end up smoothing over much of the structure of the PDF.

One obvious possibility for the bandwidth is a value which is proportional to the standard deviation, σ, of the data. In fact, it can be shown [102] that for Gaussian distributed data in the limit $n \to \infty$, the optimal bandwidth is given by

$$h = \left(\frac{4}{3n}\right)^{1/5} \sigma \ . \tag{A.7}$$

The problem with this procedure is that it assumes that the density is relatively flat over the full data range. In reality, the structure of the data is often such that some regions are sparsely populated, whilst others are very dense. An improvement over this so-called *fixed* KDE is an *adaptive* KDE. The difference between the fixed and adaptive estimation techniques is that in the latter case, the bandwidth is no longer a constant, but instead is a function of the data. Abramson [103] proposed the adaptive bandwidth parameter given by

$$h_i = \frac{h}{\sqrt{f\left(t_i\right)}} \ , \tag{A.8}$$

where h_i is the bandwidth parameter and f_i is the density function for data point i. This treatment means that the bandwidth is allowed to vary based on the local density of the data. The adaptive KDE is then given by

$$\hat{f}_1\left(x\right) \equiv \frac{1}{n}\sum_{i=0}^{n}\frac{1}{h_i}\mathcal{K}\left(\frac{x-t_i}{h_i}\right) \ . \tag{A.9}$$

Of course, the optimal local density function $f\left(t_i\right)$, and the optimal choice of h still need to be determined. Assuming $\hat{f}_0\left(t_i\right)$ is the optimal choice for $f\left(t_i\right)$, it can be shown [101] that the optimal bandwidth is given by

$$h_i = \left(\frac{4}{3n}\right)^{1/5}\sqrt{\frac{\sigma}{\hat{f}_0\left(t_i\right)}} \ . \tag{A.10}$$

Appendix B

$A_{fav}^{K3\pi}$ systematic uncertainty associated with A_{prod} and A_{Kdet}

The physics, production and kaon detection asymmetries for $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D K^{\pm}$, a_f, a_p and a_d respectively, are related to the ratio of the yields of negative and positive B decays, $a \equiv N_-/N_+$ by

$$a = a_f a_p a_d^2 \; . \tag{B.1}$$

These asymmetries are related to the observable asymmetries, $A_{fav}^{K3\pi}$, A_{prod} and A_K^{det} by

$$a_f = \frac{1 + A_{fav}^{K3\pi}}{1 - A_{fav}^{K3\pi}} \; , \quad a_p = \frac{1 + A_{prod}}{1 - A_{prod}} \text{ and } a_d = \frac{1 + A_{Kdet}}{1 - A_{Kdet}} \; . \tag{B.2}$$

The uncertainties on the measurements of A_{prod} and A_{Kdet}, $\sigma_{A_{prod}}$ and $\sigma_{A_{Kdet}}$ respectively, can be propagated to an uncertainty on the $A_{fav}^{K3\pi}$ measurement, $\sigma_{A_{fav}^{K3\pi}}$ by

$$\sigma^2_{A_{fav}^{K3\pi}} = \left| \frac{\partial A_{fav}^{K3\pi}}{\partial A_{prod}} \right|^2 \sigma^2_{A_{prod}} + \left| \frac{\partial A_{fav}^{K3\pi}}{\partial A_{Kdet}} \right|^2 \sigma^2_{A_{Kdet}} + 2 \frac{\partial A_{fav}^{K3\pi}}{\partial A_{prod}} \frac{\partial A_{fav}^{K3\pi}}{\partial A_{Kdet}} \operatorname{cov}(A_{prod}, A_{Kdet}) \; , \tag{B.3}$$

where $\operatorname{cov}(A_{prod}, A_{Kdet})$ is the covariance of A_{prod} and A_{Kdet}. Assuming that A_{prod} and A_{Kdet} are 100% correlated, then $\operatorname{cov}(A_{prod}, A_{Kdet}) = \sigma_{A_{prod}} \sigma_{A_{Kdet}}$. Since the relation between the asymmetries is more convenient to differentiate in ratio form (as defined by Eq. (B.1)), Eq. (B.3) can be written in terms of partial derivatives with respect to a_f, a_p and a_d:

$$\begin{aligned}\sigma^2_{A_{fav}^{K3\pi}} &= \left|\frac{\partial A_{fav}^{K3\pi}}{\partial a_f}\frac{\partial a_f}{\partial a_p}\frac{\partial a_p}{\partial A_{prod}}\right|^2 \sigma^2_{A_{prod}} + \left|\frac{\partial A_{fav}^{K3\pi}}{\partial a_f}\frac{\partial a_f}{\partial a_d}\frac{\partial a_d}{\partial A_{Kdet}}\right|^2 \sigma^2_{A_{Kdet}} \\ &+ 2\frac{\partial A_{fav}^{K3\pi}}{\partial a_f}\frac{\partial a_f}{\partial a_p}\frac{\partial a_p}{\partial A_{prod}}\frac{\partial A_{fav}^{K3\pi}}{\partial a_f}\frac{\partial a_f}{\partial a_d}\frac{\partial a_d}{\partial A_{Kdet}}\sigma_{A_{prod}}\sigma_{A_{Kdet}}\end{aligned} . \tag{B.4}$$

It can be seen from Eq. (B.1) and Eq. (B.2) that $\frac{\partial a_f}{\partial a_p}$, $\frac{\partial a_p}{\partial A_{prod}}$, $\frac{\partial a_f}{\partial a_d}$ and $\frac{\partial a_d}{\partial A_{Kdet}}$ are all negative (neglecting the signs of the asymmetries), and so the third term in Eq. (B.4) is positive. Thus, Eq. (B.4) has the form $f(x, y) = x^2 + y^2 + 2xy \equiv (x + y)^2$, so $\sigma_{A_{fav}^{K3\pi}}$ simplifies to

$$\sigma_{A_{fav}^{K3\pi}} = \frac{\partial A_{fav}^{K3\pi}}{\partial a_f}\left(\frac{\partial a_f}{\partial a_p}\frac{\partial a_p}{\partial A_{prod}}\sigma_{A_{prod}} + \frac{\partial a_f}{\partial a_d}\frac{\partial a_d}{\partial A_{Kdet}}\sigma_{A_{Kdet}}\right) . \tag{B.5}$$

Evaluating the partial derivatives in Eq. (B.5) gives

$$\sigma_{A_{fav}^{K3\pi}} = \frac{4a_f}{(a_f+1)^2}\left(\frac{\sigma_{A_{prod}}}{a_p\,(A_{prod}-1)^2} + \frac{2\sigma_{A_{Kdet}}}{a_d\,(A_{Kdet}-1)^2}\right) . \tag{B.6}$$

Finally, substituting Eq. (B.2) into Eq. (B.6) gives the systematic uncertainty for $A_{fav}^{K3\pi}$, $\sigma_{A_{fav}^{K3\pi}}$, in terms of A_{prod} and A_K^{det}, and their uncertainties:

$$\sigma_{A_{fav}^{K3\pi}} = \left(1 - \left(A_{fav}^{K3\pi}\right)^2\right)\left(\frac{\sigma_{A_{prod}}}{1 - A_{prod}^2} + 2\frac{\sigma_{A_{Kdet}}}{1 - A_{Kdet}^2}\right) . \tag{B.7}$$

Appendix C

Cross-check of $B \to Dh$ low mass parametrisation

As discussed in Sec. 4.6.2, the PDF describing the partially reconstructed backgrounds used in the final fit to measure $\mathcal{B}_{DK}/\mathcal{B}_{D\pi}$ and $A^{K3\pi}_{fav}$ is constructed from dedicated MC samples of the backgrounds that are expected to contaminate the signal region. These decays generally involve a D^* or ρ resonance and a missing pion or photon, and a D^0 or $\overline{D}^0$ which decays to the correct final state. Unfortunately, decay models involving a D decaying to $K^\pm\pi^\mp\pi^\pm\pi^\mp$ were not available when the MC samples were generated, so the low mass PDF has instead been determined from decay modes involving a D decaying to $K^\pm\pi^\mp$. As with the B invariant-mass distribution of decays in which the bachelor has been assigned the wrong mass hypothesis, discussed in Sec. 4.7, the B invariant-mass distributions of these backgrounds can be considered a convolution of the "natural" distribution, resulting from the combination of the true four-momentum of the tracks minus the missing particle (and if applicable, the wrong mass assignment of one of the final-state kaons or pions) and a function that describes the effects of the limited detector resolution. The resolution of the D decaying to a four-body mode will generally be better than a D decaying to a two-body mode due to the increased number of degrees of freedom in the vertex fit. However, as was shown in Sec. 4.7, there is also a significant difference between the detector resolution simulated in the

MC and that of the data, with the MC tending to underestimate the amount by which the B invariant-mass distribution is "smeared" due to the detector resolution.

Table 4.12 indicates that the Gaussian width of the $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ signal MC (neglecting the tails of the Cruijff PDF fit) is $13.34 \pm 0.22\,\mathrm{MeV}/c^2$. The corresponding width for $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ signal MC is $14.44 \pm 0.06\,\mathrm{MeV}/c^2$ [3]. This indicates that the resolution of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ is smaller than $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ by around $1\,\mathrm{MeV}/c$. However, the corresponding width in $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ (with no PID cut applied) is 17.15 ± 0.58. So, in fact, the difference between the resolution in data and that of MC appears to be more significant than the difference in resolution as a result of using $D \to K^{\pm}\pi^{\mp}$ decays in the low mass parametrisation instead of $D \to K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp}$.

A cross-check of the effect of the poorer detector resolution in data on the low-mass PDF has been performed by the numerical convolution of the adaptive KDE PDFs constructed from the MC data shown in Fig. 4.13 with a Gaussian kernel with width equal to the difference between the width of the Gaussian core fitted to the B mass signal peak of $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ data candidates and the width according to the fit to $B^{\pm} \to (K^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ signal MC candidates.

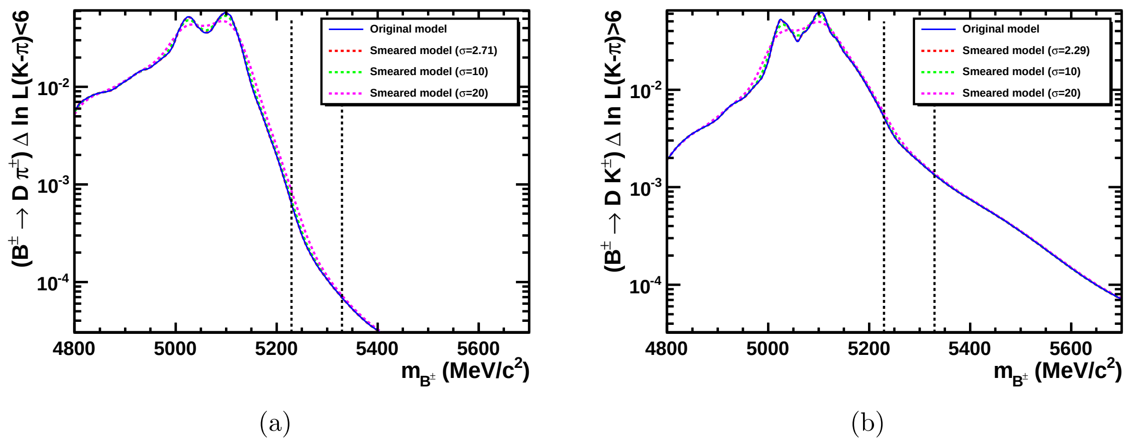


Figure C.1: Plots showing the effect of reduced resolution on the B invariant-mass distribution of low mass backgrounds for (a) $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,\pi^{\pm}$ and (b) $B^{\pm} \to (K^{\pm}\pi^{\mp}\pi^{\pm}\pi^{\mp})_D\,K^{\pm}$. The solid lines represent the PDFs constructed from the low mass background MC samples, whilst the dashed lines show the low mass PDFs convoluted with a Gaussian kernel. In each case, the width of the kernel is indicated in the legend. The vertical lines correspond to $\pm 50\,\mathrm{MeV}/c^2$ from the nominal $B^{\pm}$ mass.

The smeared PDFs are shown in Fig. C.1, along with the original low mass PDFS, plus two addition PDFs with Gaussian kernel widths of $10\,\mathrm{MeV}/c^2$ and $20\,\mathrm{MeV}/c^2$. The width of the $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,K^\pm$ signal PDF in data has been determined by scaling the width according to $B^\pm \to (K^\pm\pi^\mp\pi^\pm\pi^\mp)_D\,\pi^\pm$ MC, as described in Sec. 4.7.1. The results indicate that the poorer detector resolution in data compared to MC has a negligible effect on the B invariant-mass distribution of the low mass backgrounds; this is true for a Gaussian kernel width of $10\,\mathrm{MeV}/c^2$, around five times larger than this naive study suggests. With a Gaussian kernel width of $20\,\mathrm{MeV}/c^2$, there is a small difference before and after smearing in the tails of the low mass PDFs. However, this level of smearing is unrealistically large, since also smooths out the peaks in the distribution.

Based on the above cross-check, the effect of the reduced resolution in data compared to MC is assumed to have a negligible impact on the parametrisation of the low mass backgrounds.

Bibliography

[1] S. Brisbane, J. Libby, and N. Harnew, "CLEO-c $D \to K^0_{S/L}\pi^+\pi^-$ Binned Dalitz-Plot Analyses Optimised for the CKM Angle γ Measurement and the Commissioning of the LHCb RICH Front-End Electronics", PhD thesis, University of Oxford, Oxford, 2010. Presented 15 Dec 2010. `http://cdsweb.cern.ch/record/1335669`

[2] N. Styles and P. J. Clark, "RICH detector time alignment and studies of CP violation in the decay $B^0_s \to \phi\phi$ at the LHCb experiment", PhD thesis, University of Edinburgh, Edinburgh, 2009. Presented on 21 Oct 2009. `http://cdsweb.cern.ch/record/1302714`

[3] **LHCb** Collaboration, *A measurement of the ratio of branching fractions:* $\frac{\mathcal{B}(B^\pm\to DK^\pm)}{\mathcal{B}(B^\pm\to D\pi^\pm)}$ *for* $D \to K\pi$*,* KK*,* $K\pi\pi\pi$ *and* $K^0_S\pi\pi$. Unpublished work, Jul 2011. LHCb-ANA-2010-013; LHCb-ANA-2011-053.

[4] E. Noether, *Gott. Nachr.* **1918** (1918) 235–257, `arXiv:physics/0503066`.

[5] C. S. Wu *et al.*, *Phys. Rev.* **105** (1957) 1413–1414.

[6] G. Luders, *Annals Phys.* **2** (1957) 1–15.

[7] W. Pauli, L. Rosenfeld, and V. Weisskopf, "Niels Bohr and the Development of Physics", 1955.

[8] O. W. Greenberg, *Found. Phys.* **36** (2006) 1535–1553, `arXiv:hep-ph/0309309`.

[9] G. Luders and B. Zumino, *Phys. Rev.* **106** (1957) 385–386.

[10] E. C. G. Sudarshan and R. e. Marshak, *Phys. Rev.* **109** (1958) 1860–1860.

[11] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Phys. Rev. Lett.* **13** (1964) 138–140.

[12] **BABAR** Collaboration, B. Aubert *et al.*, *Phys. Rev. Lett.* **87** (2001) 091801, `arXiv:hep-ex/0107013`.

[13] **Belle** Collaboration, K. Abe *et al.*, *Phys. Rev. Lett.* **87** (2001) 091802, `arXiv:hep-ex/0107061`.

[14] **LHCb** Collaboration, R. Aaij *et al.*, *Phys. Rev. Lett.* **108** (2012) 111602, `arXiv:1112.0938`.

[15] V. Weisskopf and E. P. Wigner, *Z. Phys.* **63** (1930) 54–73.

[16] V. Weisskopf and E. P. Wigner, *Z. Phys.* **65** (1930) 18–29.

[17] N. Cabibbo, *Phys. Rev. Lett.* **10** (1963) 531–533.

[18] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Phys. Rev.* **D2** (1970) 1285–1292.

[19] M. Kobayashi and T. Maskawa, *Prog.Theor.Phys.* **49** (1973) 652–657.

[20] L.-L. Chau and W.-Y. Keung, *Phys. Rev. Lett.* **53** (1984) 1802.

[21] L. Wolfenstein, *Phys. Rev. Lett.* **51** (1983) 1945.

[22] **CKMfitter Group** Collaboration, J. Charles *et al.*, *Eur. Phys. J.* **C41** (2005) 1–131, `arXiv:hep-ph/0406184`.

[23] **Particle Data Group** Collaboration, K. Nakamura *et al.*, *J.Phys.G* **G37** (2010) 075021.

[24] M. Gronau and D. London, *Phys. Lett.* **B253** (1991) 483–488.

[25] M. Gronau and D. Wyler, *Phys.Lett.* **B265** (1991) 172–176.

[26] D. Atwood, I. Dunietz, and A. Soni, *Phys.Rev.Lett.* **78** (1997) 3257–3260, `arXiv:hep-ph/9612433`.

[27] **Heavy Flavor Averaging Group** Collaboration, D. Asner *et al.* `arXiv:1010.1589`.

[28] D. Atwood and A. Soni, *Phys.Rev.* **D68** (2003) 033003, `arXiv:hep-ph/0304085`.

[29] **CLEO** Collaboration, N. Lowrey *et al.*, *Phys.Rev.* **D80** (2009) 031105, `arXiv:0903.4853`.

[30] **Belle** Collaboration, X. C. Tian *et al.*, *Phys. Rev. Lett.* **95** (2005) 231801, `arXiv:hep-ex/0507071`.

[31] L. Evans (ed.) and P. Bryant (ed.), *JINST* **3** (2008) S08001.

[32] **LHCb** Collaboration, A. A. Alves Jr. *et al.*, *JINST* **3** (2008) S08005.

[33] **LHCb** Collaboration, R. Aaij *et al.*, *Phys. Lett.* **B694** (2010) 209–216, `arXiv:1009.2731`.

[34] P. Bjornstad, *JINST* **6** (2011) C12024.

[35] P. R. Barbosa-Marinho *et al.*, “LHCb inner tracker: Technical Design Report”, tech. rep., CERN, Geneva, 2002.

[36] R. Hierk, M. Merk, M. Needham, and R. Van der Eijk, “Performance of the LHCb 00 track fitting software”, Tech. Rep. LHCb-2000-086, CERN, Geneva, Aug 2000.

[37] E. Rodrigues, "Dealing with clones in the tracking", Tech. Rep. LHCb-2006-057. CERN-LHCb-2006-057, CERN, Geneva, Nov 2006.

[38] M. Needham, "Clone Track Identification using the Kullback-Liebler Distance", Tech. Rep. LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002, CERN, Geneva, Jan 2008.

[39] O. Callot, "Downstream pattern recognition", Tech. Rep. LHCb-2007-026. CERN-LHCb-2007-026, CERN, Geneva, Mar 2007.

[40] A. Arefev *et al.*, "Beam Test Results of the LHCb Electromagnetic Calorimeter", Tech. Rep. LHCb-2007-149. CERN-LHCb-2007-149, CERN, Geneva, May 2008.

[41] C. Coca *et al.*, "The hadron calorimeter prototype beam-test results", Tech. Rep. LHCb-2000-036, CERN, Geneva, Oct 2000.

[42] O. Ullaland, *Nucl. Instrum. Meth.* **A553** (2005) 107–113.

[43] **ALICE** Collaboration, G. Dellacasa *et al.*, CERN-LHCC-99-12.

[44] **LHCb** Collaboration, R. Aaij *et al.*, *Phys. Lett.* **B708** (2012) 55–67, `arXiv:1112.1600`.

[45] C. G. Lester, *Nucl. Instrum. Meth.* **A560** (2006) 621–632, `arXiv:physics/0508034`.

[46] A. Schöning, "Particle identification in the RICH detectors and study of impact parameters", Tech. Rep. LHCb-97-018, CERN, Geneva, Nov 1997.

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

[48] V. Gligorov, "A single track HLT1 trigger", Tech. Rep. LHCb-PUB-2011-003. CERN-LHCb-PUB-2011-003. LHCb-INT-2010-053, CERN, Geneva, Jan 2011.

[49] M. Williams, V. Gligorov, C. Thomas, H. Dijkstra, J. Nardulli, and P. Spradlin, "The HLT2 Topological Lines", Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan 2011.

[50] J. Christiansen, "Requirements to the L0 front-end electronics; 2001 ed.", Tech. Rep. LHCb-2001-014, CERN, Geneva, Jul 2001. Supersedes LHCb-NOTE-FE-99-29.

[51] P. Moreira, G. Cervelli, J. Christiansen, F. Faccio, A. Kluge, *et al.*, 2001, Prepared for 7th Workshop on Electronics for LHC Experiments, Stockholm, Sweden, 10-14 Sep 2001. `http://cdsweb.cern.ch/record/588665`

[52] G. Haefeli, A. Bay, A. Gong, H. Gong, M. Muecke, N. Neufeld, and O. Schneider, *Nucl. Instrum. Meth.*, 2006 494–502.

[53] W. Wooton, S. Katvars, and C. Barham, *CERN EDMS document*, 2006. `https://edms.cern.ch/file/768873/1/UKL1.pdf`

[54] B. G. Taylor, *IEEE Trans. Nucl. Sci.* **45** (1998) 821–828.

[55] D. Breton and D. Charlet, "SPECS: the Serial Protocol for the Experiment Control System of LHCb", Tech. Rep. LHCb-2003-004, CERN, Geneva, Jan 2003.

[56] M. Gonzalez-Berges, *ArXiv Physics e-prints*, 2003 `arXiv:physics/0305128`.

[57] G. Barrand *et al.*, *Comput. Phys. Commun.* **140** (2001) 45–55.

[58] M. Clemencic *et al.*, *J. Phys. Conf. Ser.* **219** (2010) 042006.

[59] T. Sjostrand, S. Mrenna, and P. Z. Skands, *JHEP* **05** (2006) 026, `arXiv:hep-ph/0603175`.

[60] D. J. Lange, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[61] **GEANT4** Collaboration, S. Agostinelli *et al.*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[62] R. Brun and F. Rademakers, *Nucl. Instrum. Meth.* **A389** (1997) 81–86.

[63] P. Moreira and A. Marchioro, 2003. `http://cdsweb.cern.ch/record/722067`

[64] J. Bibby, "LHCb RICH L-0 Interface Module", Tech. Rep. LHCb/RICH 2006-30, Geneva, May 2006. `https://edms.cern.ch/file/745431/1/LHCb-2006-30.pdf`

[65] B. Franek and C. Gaspar, *J. Phys. Conf. Ser.* **219** (2010) 022031.

[66] C. Gaspar and M. Donszelmann, Prepared for 8th Real-Time Computer Applications in Nuclear, Particle and Plasma Physics (RT 93), Vancouver, British Columbia, Canada, 8-11 Jun 1993. `http://cdsweb.cern.ch/record/254799`

[67] A. Powell, N. Harnew, and G. Wilkinson, "Measuring the CKM Angle γ with $B^{\pm} \to DK^{\pm}$ Decays at LHCb and a Determination of the $D \to K\pi\pi\pi$ Coherence Factor at CLEO-c", PhD thesis, University of Oxford, Oxford, 2009. Presented on 22 May 2009. `http://cdsweb.cern.ch/record/1233169`

[68] C. Fitzpatrick, "SVN web repository of the LHCb RichTimingScanAnalysis Panoptes package". `http://svn.cern.ch/guest/lhcb/Panoptes/trunk/Rich/RichTimingScanAnalysis`

[69] **Belle** Collaboration, Y. Horii *et al.*, *Phys.Rev.* **D78** (2008) 071901, `arXiv:0804.2063`.

[70] **BABAR** Collaboration, B. Aubert *et al.*, *Phys.Rev.Lett.* **92** (2004) 202002, `arXiv:hep-ex/0311032`.

[71] C. Fitzpatrick, "simpletools: Handy command line tools for ntuple manipulation and analysis", Tech. Rep. LHCb-INT-2009-029. CERN-LHCb-INT-2009-029, CERN, Geneva, Nov 2009.

[72] P. Hunt and M. J. J. John, "Selection Study for $B^{\pm} \to (K^{\mp}\pi^{\pm}\pi^{\mp}\pi^{\pm})_D K^{\pm}$ and $B^{\pm} \to (K^{+}K^{-}\pi^{+}\pi^{-})_D K^{\pm}$", Tech. Rep. LHCb-PUB-2009-002. CERN-LHCb-PUB-2009-002, CERN, Geneva, Sep 2009.

[73] W. Verkerke and D. Kirkby `arXiv:physics/0306116`.

[74] M. Pivk and F. R. Le Diberder, *Nucl. Instrum. Meth.* **555** (2005) 356–369, `arXiv:physics/0402083`.

[75] **LHCb** Collaboration, *Measurement of direct $\mathcal{CP}$ violation in charmless charged two-body B decays at LHCb.* Unpublished work, Apr 2011. LHCb-ANA-2011-023.

[76] **Mark-III Collaboration** Collaboration, J. Adler *et al.*, *Phys.Rev.Lett.* **64** (1990) 2615–2618.

[77] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger `arXiv:1202.0439`.

[78] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Phys.Rev.* **D71** (2005) 014018, `arXiv:hep-ph/0410289`.

[79] B. A. Kniehl and G. Kramer, *Phys.Rev.* **D71** (2005) 094013, `arXiv:hep-ph/0504058`.

[80] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Phys.Rev.Lett.* **96** (2006) 012001, `arXiv:hep-ph/0508129`.

[81] T. Kneesch, B. Kniehl, G. Kramer, and I. Schienbein, *Nucl.Phys.* **B799** (2008) 34–59, `arXiv:0712.0481`.

[82] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Phys.Rev.* **D79** (2009) 094009, `arXiv:0901.4130`.

[83] M. Cacciari and P. Nason, *JHEP* **0309** (2003) 006, `arXiv:hep-ph/0306212`.

[84] M. Cacciari, P. Nason, and C. Oleari, *JHEP* **0604** (2006) 006, `arXiv:hep-ph/0510032`.

[85] R. Le Gac, "Statistics of LHCb data-taking", tech. rep. `http://marwww.in2p3.fr/~legac/LHCb`

[86] **LHCb** Collaboration, "Nu, Mu and Pile-Up. The LHCb definitions of what we see and what we don't see.", tech. rep. `https://twiki.cern.ch/twiki/bin/view/LHCb/NuMuPileUp`

[87] R. Dalitz, *Phil.Mag.* **44** (1953) 1068–1080.

[88] J. Rademacker. Private communication.

[89] M. Dobbs and J. B. Hansen, *Comput. Phys. Commun.* **134** (2001) 41–46.

[90] **LHCb** Collaboration, "PIDCalib software packages and tools", tech. rep. `https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage`

[91] E. B. Wilson, *Journal of the American Statistical Association* **22** (1927), no. 158, 209–212. `http://dx.doi.org/10.2307/2276774`

[92] V. Gligorov, "LHCb IBackgroundCategory Doxygen page". `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/releases/latest/doxygen/df/d88/class_i_background_category.html`

[93] A. Kozlinskiy. Private communication.

[94] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, 2010. Private communication.

[95] M. Cacciari, S. Frixone, M. Mangano, M. Nason, and G. Ridolfi, 2010. Private communication.

[96] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*, *Phys.Rev.* **D78** (2008) 013004, `arXiv:0802.0007`.

[97] W. Tung, H. Lai, A. Belyaev, J. Pumplin, D. Stump, *et al.*, *JHEP* **0702** (2007) 053, `arXiv:hep-ph/0611254`.

[98] B. A. Kniehl and G. Kramer, *Phys.Rev.* **D74** (2006) 037502, `arXiv:hep-ph/0607306`.

[99] J. Pumplin, H. Lai, and W. Tung, *Phys.Rev.* **D75** (2007) 054029, `arXiv:hep-ph/0701220`.

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

[101] K. S. Cranmer, *Comput.Phys.Commun.* **136** (2001) 198–207, `arXiv:hep-ex/0011057`.

[102] D. W. Scott, "Multivariate Density Estimation: Theory, Practice, and Visualization (Wiley Series in Probability and Statistics)", Wiley, 1992.

[103] I. S. Abramson, *Ann. Stat.* **10** (1982), no. 4, 1217–1223.