CALIBRATION OF THE LHCB VELO DETECTOR AND
STUDY OF THE DECAY MODE $D^0 \rightarrow K^- \mu^+ \nu_\mu$

A THESIS
SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY (PhD)
IN THE FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

GRANT D. MCGREGOR
B.Sc. (hons), UNIVERSITY OF CANTERBURY, 2005
M.Sc., UNIVERSITY OF BRITISH COLUMBIA, 2009

PARTICLE PHYSICS GROUP
SCHOOL OF PHYSICS AND ASTRONOMY
2011
## Abstract

13

## Declaration

15

## Copyright

17

## Acknowledgements

19

### 1 Introduction

21

### 2 Theoretical Introduction

23

#### 2.1 The Standard Model

23

#### 2.2 Semileptonic Charm Decays

26

#### 2.3 Form Factor Measurements

33

#### 2.4 $D^0 - \overline{D}^0$ Mixing in Semileptonic Decays

36

### 3 The LHCb Experiment

39

#### 3.1 The Large Hadron Collider

39

##### 3.1.1 Introduction

39

##### 3.1.2 LHC Design

41

#### 3.2 The LHCb Experiment

42

##### 3.2.1 The Vertex Locator

46

##### 3.2.2 The Tracking System

53

##### 3.2.3 The Magnet

55

##### 3.2.4 Particle Identification

56

##### 3.2.5 The Trigger

59

##### 3.2.6 The LHCb Software Environment

62

### 4 VELO Gain Calibration and Monitoring

63

#### 4.1 Introduction

63

#### 4.2 Full Header Swing

64
LIST OF FIGURES

2.1 The Feynman diagram for the $W$-mediated semileptonic decay of $D^0 \rightarrow K^- \ell^+ \nu$. ......................................................... 27
2.2 A graphical representation of the Unitarity Triangle in the complex plane. 30
2.3 Borrowed from the FOCUS experiment: Background-subtracted $f_+(q^2)$ values compared to three theoretical models. ......................... 34
3.1 An underground schematic of the SPS, the LHC and the four largest LHC experiments ................................................................. 39
3.2 Overview of the CERN accelerator complex. ............................... 42
3.3 A schematic diagram of the LHC. ............................................. 43
3.4 Schematic diagram of the LHCb detector. ................................. 45
3.5 Angular distribution of $b$-quarks inside LHCb. .......................... 45
3.6 A VELO half shown with its RF box. ........................................ 47
3.7 A diagram showing how the two RF boxes mesh together when the VELO is closed. ................................................................. 48
3.8 Schematic cross-section of the VELO in the $xz$ plane at $y = 0$. ...... 49
3.9 A VELO module. ........................................................................ 49
3.10 Raw Beetle test pulse output. .................................................... 50
3.11 Summary of the horizontal and vertical misalignments for all VELO sensors. 51
3.12 Vertex resolutions in $x$ (red) and $y$ (blue). Above $N = 20$ a curve proportional to $1/\sqrt{N}$ is fitted to each. ................................. 52
3.13 The $x$ component of IP resolution for 2010 data and simulation with linear best-fit curves overlaid. ................................................ 53
3.14 Single hit resolution for the binary estimate and for low-angle and high-angle tracks from collisions. ............................................. 53
3.15 Illustration of the TT and the IT. ................................................ 54
3.16 Illustration of the $y$ component of the LHCb magnetic field and its position relative to the silicon trackers. ..................................... 55
3.17 The RICH1 detector seen in the vertical plane. ............................ 56
3.18 Cherenkov polar angle $\theta$ versus particle momentum for simulated $B^0 \rightarrow \pi^+\pi^-$ events. ................................. 57
3.19 A schematic (left) and a photograph (right) of a pixel HPD. .......... 58
3.20 A view of the ECAL installed in the LHCb pit and the three types of ECAL modules. ........................................... 59
3.21 A schematic side-view of the muon system. .............................. 60
3.22 An overview of the Level-0 trigger showing the number of channels involved. ........................................ 61
3.23 A flow-diagram of the LHCb trigger. ............................... 61

4.1 Typical high and low header distributions for a link. ............... 65
4.2 Uncalibrated distribution of FHS values. ............................... 66
4.3 The test pulse distributions for the VELO calibrated with the FHS method. .............................................. 67
4.4 A typical Landau distribution for a sample of collision data. ........ 68
4.5 The track definitions used in LHCb. .................................. 69
4.6 Summary of the number of hits in each link for the minimum bias sample. ........................................... 70
4.7 Difference between the Landau maximum bin and the fitted MPV. ....... 70
4.8 Distribution of fitted MPVs. ........................................... 71
4.9 Distribution of FWHM of fit for each link. ............................... 71
4.10 Distribution of MPVs for $R$ sensors for a VELO calibrated under the FHS method. ................................... 72
4.11 Distribution of MPVs for $\Phi$ sensors for a VELO calibrated under the FHS method. ................................... 73
4.12 FHS response to a Gain Scan. ....................................... 74
4.13 FHS distributions after a gain calibration. .......................... 76

5.1 A typical Landau distribution for VELO sensor 38, link 41 for clusters on tracks. ........................................... 79
5.2 The track definitions used in LHCb. .................................. 80
5.3 Number of clusters per track for long and upstream tracks. .......... 81
5.4 MC distributions for the four generalised means as a function of the true momentum. .................................. 82
5.5 MC distributions for mean, median and the two truncated means as a function of the true momentum. ......................... 82
5.6 The 40% truncation distribution for each particle species separately. ... 83
5.7 Profile plot of the 40% truncation distribution. ........................ 84
5.8 True momentum compared to reconstructed momentum. ............. 84
5.9 Scatter plot of the 40% truncation for MC with reconstructed momentum and long tracks only. ............................ 85
5.10 Profile plot of Figure 5.9 (long tracks only). .......................... 85
5.11 Scatter plot of the 40% truncation for MC with reconstructed momentum for long and upstream tracks. ........................................ 86
5.12 Profile plot of Figure 5.11 (long and upstream tracks). ............... 86
5.13 Distribution of the 40% truncated mean estimate in the momentum bin 0.7–0.9 GeV/c. ........................................ 87
5.14 Distribution of the 40% truncated mean estimate in the momentum bin 0.9–1.1 GeV/c. ........................................ 87
5.15 Distribution of MC for 40% truncated mean with no requirements on ID. 90
5.16 Distribution of collision data for 40% truncated mean. .................. 90
5.17 40% truncated mean for MC tracks selected using global PID. .......... 92
6.1 RICH PID performance as a function of track momentum. ............... 97
6.2 Signal and background distributions for DIRA. .......................... 98
6.3 Signal and background distributions for HLT2 cut variables (1). ......... 102
6.4 Signal and background distributions for HLT2 cut variables (2). ......... 102
6.5 Signal and background distributions for HLT2 cut variables (3). ......... 103
6.6 Signal and background distributions for HLT2 cut variables (4). ......... 103
6.7 Correlations between variables used in the HLT2 line. .................... 105
6.8 Kinematics for the semileptonic decay of a $c$ meson. .................. 108
6.9 Generator-level true $q^2$ distribution for one million $D^0 \rightarrow K^- \mu^+ \nu_\mu$ events with daughters within the LHCb acceptance. .......... 109
6.10 Generator-level and loose pre-selected MC $q^2$ distributions. .......... 111
6.11 Pre-selected MC $q^2$ distributions with lower-level triggers. .......... 112
6.12 Pre-selected MC $q^2$-bias distributions with lower-level triggers. ....... 113
6.13 HLT1 MC $q^2$ distribution with and without a cut on $D^0$ flight distance. 114
6.14 HLT1 MC $q^2$ distribution with and without a cut on $D^0$ corrected mass. 114
6.15 HLT1 MC $q^2$ distribution with and without a cut on the sum of daughter transverse momenta. .......................... 115
6.16 HLT1 MC $q^2$ distribution with and without a cut on kaon transverse momentum. ........................................ 115
6.17 HLT1 MC $q^2$ distribution with and without a cut on daughter DOCA. 116
6.18 HLT1 MC $q^2$ distribution with and without a cut on $D^0$ momentum. .... 116
6.19 HLT1 MC $q^2$ distribution with and without a cut on flight distance $\chi^2$. 117
6.20 HLT1 MC $q^2$ distribution with and without a cut on minimum daughter IP $\chi^2$. ........................................ 118
6.21 Generator-level and HLT2 selected MC $q^2$ distributions. .......... 118
6.22 MC True distributions of $\delta_m$ and DOCA ($D^0 - \pi_s$) for signal (blue) and background (red) MC. .......................... 121
6.23 MC True distributions of $D^*$ vertex $\chi^2$ for signal (blue) and background (red) MC. ............................................. 121
6.24 HLT1 MC $q^2$ distribution with and without a cut on kaon PID. .............. 122
6.25 The efficiency and mis-ID rates for muon PID definitions. .................... 123
6.26 Number of candidates per event in each signal MC event selected by the stripping line. .................................................. 126
6.27 Number of candidates per event for each candidate with a MC true $D^0$ decay associated to a false $D^*$. ........................................... 126
6.28 $\delta_m$ distribution for true $D^{*+}$ decays in signal MC. ......................... 127
6.29 $\delta_m$ distribution for all right-sign candidates in signal MC. ................. 127
6.30 $\delta_m$ distribution for all wrong-sign candidates in signal MC. .............. 127
6.31 $\delta_m$ distribution for non-signal, right-sign candidates in $b\bar{b}$ MC. .... 128
6.32 $\delta_m$ distribution for non-signal, wrong-sign candidates in $b\bar{b}$ MC. .... 128
6.33 $\delta_m$ distribution for non-signal, right-sign candidates in $c\bar{c}$ MC. ....... 128
6.34 $\delta_m$ distribution for non-signal, wrong-sign candidates in $c\bar{c}$ MC. ...... 128
6.35 $M(K\mu)$ distribution for true $D^{*+}$ decays in signal MC. ................ 129
6.36 $M(K\mu)$ distribution for all right-sign candidates in signal MC. ............ 129
6.37 $M(K\mu)$ distribution for all wrong-sign candidates in signal MC. .......... 129
6.38 $M(K\mu)$ distribution for non-signal, right-sign candidates in $b\bar{b}$ MC. .. 130
6.39 $M(K\mu)$ distribution for non-signal, wrong-sign candidates in $b\bar{b}$ MC. .. 130
6.40 $M(K\mu)$ distribution for non-signal, right-sign candidates in $c\bar{c}$ MC. ... 130
6.41 $M(K\mu)$ distribution for non-signal, wrong-sign candidates in $c\bar{c}$ MC. ... 130
6.42 $\delta_m$ distribution for all right-sign candidates in the sample of charm microDST data events. ................................................. 131
6.43 $\delta_m$ distribution for all wrong-sign candidates in the sample of charm microDST data events. ................................................. 131
6.44 $M(K\mu)$ distribution for all right-sign candidates in a sample of charm microDST data events. ................................................. 131
6.45 $M(K\mu)$ distribution for all wrong-sign candidates in a sample of charm microDST data events. ................................................. 131
6.46 $\delta_m$ distribution for all right-sign candidates in the sample of charm microDST data events which are TOS on the $D^0 \rightarrow K^- \mu^+ \nu_\mu$ HLT2 line. ... 132
6.47 $\delta_m$ distribution for all wrong-sign candidates in the sample of charm microDST data events which are TOS on the prescaled $D^0 \rightarrow K^+ \mu^+ \nu_\mu$
HLT2 line. ...................................................... 132
6.48 $M(K\pi)$ distribution for all right-sign candidates in the sample of charm microDST data events which are TOS on the $D^0 \rightarrow K^- \mu^+ \nu_\mu$ HLT2 line. .. 132
LIST OF FIGURES

6.49 $M(K\pi)$ distribution for all wrong-sign candidates in the sample of charm microDST data events which are TOS on the prescaled $D^0 \rightarrow K^+\mu^+\nu_\mu$ HLT2 line. ........................................... 132

7.1 Distribution of $p_{\text{rec}}/p_{\text{true}}$ plotted against the reconstructed $M(K\mu)$. ........... 136
7.2 A profile plot of Figure 7.1, with a quadratic best-fit curve forced to pass through 1 at the $D^0$ mass. ........................................... 137
7.3 The effect of a $k$-factor momentum correction compared to true $D^0$ momentum. .................. 137
7.4 $D^*$ and $D^0$ invariant mass difference for $D^0$ and $\overline{D}^0$ decays. ......................... 139
7.5 Number of primary vertices in signal MC and stripped data. ........................................... 140
7.6 Angle between reconstructed $D^0$ flight vector and the true flight vector, in radians. ............ 140
7.7 The effect of the neutrino reconstruction method for MC true quadratic solutions. ................ 141
7.8 The effect of the neutrino reconstruction method for reconstructed quadratic solutions for incorrect and correct solution choice. ......................... 142
7.9 The effect of the neutrino reconstruction method for reconstructed quadratic solutions for solutions that were chosen and not chosen. ............. 142
7.10 Comparison of the $D^0$ distributions between the $k$-factor and neutrino reconstruction methods for reconstructed MC. ...................... 143
7.11 Distribution of $\delta_m$ after neutrino reconstruction for signal and fake $D^*$ MC. ......... 145
7.12 Distribution of $\delta_m$ after neutrino reconstruction for MC of the peaking backgrounds $D^+ \rightarrow K^{*0}\mu^+\nu_\mu$ and $D^{*+} \rightarrow \pi^+_s D^0 \rightarrow \pi^+_s K^{*-}\mu^+\nu_\mu$. ....... 146
7.13 Distribution of $\delta_m$ after neutrino reconstruction for MC of the peaking background $D^+_s \rightarrow \phi\mu^+\nu_\mu$ and minimum bias. ......................... 146
7.14 Distribution of $\delta_m$ after neutrino reconstruction for $b\overline{b}$ and $c\overline{c}$ MC. ................. 147
7.15 Distribution of $\delta_m$ after neutrino reconstruction for data. ........................................... 148
7.16 Distribution of $M(K\mu)$ after neutrino reconstruction for data. ................................. 149
7.17 Signal and fake $D^*$ distributions, after trigger and stripping, for the $D^0 \pi_s$ DOCA and pion $p_T$. ........................................... 150
7.18 Signal and fake $D^*$ distributions, after trigger and stripping, for pion DLL($K-\pi$) and pion DLL($e-\pi$). ........................................... 150
7.19 The sum of two core Gaussians and a bifurcated Gaussian fitted to the $\delta_m$ MC signal peak. ................. 151
7.20 Distribution of RS $\delta_m$ for the two largest peaking backgrounds, with a fitted background PDF superimposed. ......................... 153
7.21 Distribution of $\delta_m$ for the fake $D^*$ MC, with a fitted background PDF superimposed. ................. 154
LIST OF FIGURES

7.22 Distribution of $\delta_m$ for data, with signal and background PDFs superimposed. 155
7.23 Comparison of $q^2$ distributions in nine bins for generated and reconstructed signal MC. 156
7.24 Reconstructed $q^2$ distributions in nine bins for reconstructed peaking background candidates. 156
7.25 Reconstructed $q^2$ distributions in nine bins for reconstructed fake $D^*$ candidates. 158
7.26 Reconstructed $q^2$ distributions in nine bins for reconstructed data candidates. 158
7.27 Reconstructed $q^2$ distributions in nine bins for variations in the background cocktail. 159
7.28 Reconstructed $q^2$ distributions in nine bins for wrong-sign subtracted data, background cocktail and estimated signal in data. 160
7.29 Corrected signal $q^2$ distribution in data with estimated uncertainties. 160
7.30 Corrected signal $q^2$ distribution in data with best-fit single-pole form factor model. 162
7.31 Variation of $\chi_r^2$ as a function of $m_{\text{pole}}$. 162
7.32 Distribution of $\delta_m$ after neutrino reconstruction for signal and wrong-sign $\pi_s$ candidates in data. 164
## LIST OF TABLES

2.1 Some observed properties of the quarks, from [1]. ........................................ 24
2.2 Some observed properties of the leptons, from [1]. The uncertainty on the last few decimal places of mass is indicated in parentheses where appropriate. Upper limits with a 95% confidence level (CL) are given for the neutrino masses. ................................................................. 25
2.3 Some observed properties of the gauge bosons, from [1]. All bosons in the table have spin of 1. The Higgs boson has not been observed so is not included here, but it is described in the text. ......................................................... 25
2.4 Measurements of $r_{\text{mix}}$ in $D^0$ semileptonic decays, from [1]. ................. 38
3.1 LHC beam parameters relevant for the peak luminosity. ................................. 44
3.2 Some important requirements of the three calorimeter sub-detectors. ............... 58
5.1 Relative numbers of each species passing selection from minimum bias sample. ...................................................................................................................... 83
5.2 Numbers of each species in the momentum bin 0.7–0.9 GeV/$c$, with a minimum 40% truncated mean value. .......................................................... 88
5.3 Numbers of each species in the momentum bin 0.9–1.1 GeV/$c$, with a minimum 40% truncated mean value. .......................................................... 88
5.4 Numbers of each species in the momentum bin 0.7–0.9 GeV/$c$, with a minimum $k = -6$ generalised mean value. ................................................. 89
5.5 Identification of particles using global PID. ....................................................... 91
6.1 List of loose preselections for $D^0 \to K^- \mu^+ \nu_{\mu}$ applied to signal and minimum bias MC for optimising an HLT2 line. ....................................................... 96
6.2 List of loose preselections for $D^{*+} \to \pi^+ D^0$ applied to signal and minimum bias MC for optimising an HLT2 line. ....................................................... 96
6.3 Requirements of hits in muon stations for the ISMUON flag (momentum measured in GeV/$c$). ................................................................. 97
6.4 Efficiencies for lower-level triggers ............................................................... 100
LIST OF TABLES

6.5 List of very loose pre-selection for $D^{*+} \rightarrow \pi^+ D^0$ applied to signal MC for $q^2$ optimisation. ................................................................. 100
6.6 Cuts used in the $D^0 \rightarrow K^- \mu^+ \nu_\mu$ Hlt2 line. .................. 105
6.7 List of cuts used for $D^0$ selection in the $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ stripping line. ......................................................... 120
6.8 List of cuts used for $D^*$ selection in the $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ stripping line. ......................................................... 120
6.9 A list of the most frequently-firing triggers in stripped, right-sign $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ data. ......................................................... 133

7.1 Efficiencies in MC after stripping and neutrino reconstruction, with no requirements on trigger. ................................................................. 145
7.2 List of cuts used for offline $\pi_s$ selection. ................................. 151
7.3 Fitted parameters for the RooDstD0BG PDF for the three largest backgrounds. ................................................................. 154
7.4 Efficiency matrix $\epsilon_{ij}$ for signal MC candidates, for true $q^2$ values up to 1.0 GeV$^2$/c$^2$. The column gives the true $q^2$ bin $j$, and the row gives the reconstructed $q^2$ value $i$. The statistical uncertainty in the least significant digits is given in parentheses. ................................................................. 157
7.5 Efficiency matrix $\epsilon_{ij}$ for signal MC candidates, for true $q^2$ values above 1.0 GeV$^2$/c$^2$. The column gives the true $q^2$ bin $j$, and the row gives the reconstructed $q^2$ value $i$. The statistical uncertainty in the least significant digits is given in parentheses. ................................................................. 157
7.6 Luminosity estimates of the three main background MC samples. .... 159
The LHCb experiment, based at the Large Hadron Collider at CERN, is primarily designed to make precision measurements of the decays of heavy flavour hadrons, such as $B$ and $D$ mesons. This thesis is composed of two parts: the first consists of two studies of LHCb’s vertex locator (VELO) and the second describes the development of methods for recording the decay $D^0 \rightarrow K^- \mu^+ \nu_\mu$.

The first VELO study involves calibration and monitoring of the gain (i.e. the detector response to input charge from particles). We propose a robust method to measure the gain response of each silicon sensor using calibration bits output by the sensors, and a method to recalibrate the gain simple enough to be followed by non-expert VELO users. This is followed by an investigation into the prospects of using the VELO to perform particle identification using the characteristic energy deposition of each particle species ($dE/dx$).

Finally, studies into the development of a trigger and a so-called ‘stripping line’ for recording $D^0 \rightarrow K^- \mu^+ \nu_\mu$ decays is presented. The relatively high cross-section for charm decays in LHCb mean this decay (with a branching fraction of 3%) occurs frequently, and the challenge of a trigger is to reduce this to a rate acceptable to write to disk. Finally, based on a sample of data from July and August 2011, the measured $q^2$ distribution for this decay is compared to the simple single-pole theoretical model, and the pole mass is measured to be $m_{pole} = 2.35^{+0.81}_{-0.35} \text{ GeV}/c^2$.

Institution The University of Manchester
Candidate Grant D. McGregor
Degree Title Doctor of Philosophy
Thesis Title Calibration of the LHCb VELO Detector and Study of the Decay Mode $D^0 \rightarrow K^- \mu^+ \nu_\mu$
Date September 2011
DECLARATION

This work represents the combined efforts of the LHCb collaboration. Some of the content has been published elsewhere and/or presented to several audiences, as detailed later in this thesis. I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed: Date: November 8, 2011

Grant D. McGregor
School of Physics and Astronomy
University of Manchester
Oxford Road
Manchester
M13 9PL
United Kingdom
The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy¹, in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations² and in The University’s policy on Presentation of Theses.

¹http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487
²http://www.manchester.ac.uk/library/aboutus/regulations
ACKNOWLEDGEMENTS

Fred Loebinger played a significant role in giving me the chance to carry out this research, so thanks to him and the entire Manchester particle physics group for their support. I had many useful discussions with Marius and Adam, my Manchester LHCb colleagues at CERN, so thanks to them and of course, to my supervisor George Lafferty, in particular for his excellent editing and proof-reading. This work has been funded by the Overseas Research Students Award Scheme and a fellowship from the University of Manchester.

Thanks to the members of the LHCb charm working group; to Marco Gersabeck, Vladimir Gligorov and the group co-ordinators Walter Bonivento, Matt Charles and Patrick Spradlin.

Thanks also to the members of the LHCb Velo group for creating an excellent working environment and for their ideas and contributions to the work presented in this thesis. In particular, thanks to Kazu Akiba, Martin van Beuzekom, Themis Bowcock, Paula Collins, Karol Hennessy, Chris Parkes, Kurt Rinnert and Tomasz Szumlak.

I was fortunate to live with many interesting flatmates during my stay in France, but particularly worthy of mention are Kenny Wraight and Sam Whitehead who were always available to talk about things unrelated to physics. Finally, thanks to my parents who have been highly supportive during my ten years at university.

This old guitar has caught some breaks
But it never searched for gold.
It can’t be blamed for my mistakes;
It only does what it’s told.

Neil Young, This Old Guitar. Prairie Wind. Reprise, 2005.
LHCb is one of four large particle detectors located at the Large Hadron Collider (LHC) along with ATLAS, CMS and Alice. LHCb has recorded physics data since late 2009. As suggested in its name, LHCb (where ‘b’ represents beauty) is designed to make measurements of decays of hadrons containing $b$ quarks. However, in addition to being a source of $b$ decays, the LHC is also a high-rate source of mesons containing $c$ quarks.

After an introduction to the theory of semileptonic charm decays in Chapter 2, an introduction to the LHCb experiment and the various subdetectors is given in Chapter 3.

Chapters 4 and 5 describe two studies of LHCb’s vertex locator. The first on the gain and methods for calibrating and monitoring this quantity. Chapter 5 summarises an investigation into the prospects of using the VELO to perform $dE/dx$ particle identification. The final conclusion is that this method is limited by the geometry of LHCb, which causes the lowest momenta tracks to be swept out of the acceptance region before sufficiently many measurements can be made of the track.

Measurements of the decay $D^0 \rightarrow K^- \mu^+ \nu_\mu$ have several potential applications. Like other semileptonic decays of heavy mesons, the Standard Model provides a relatively simple theoretical model for this decay, because the leptonic and hadronic current can be untangled. This provides a relatively simple access for measurements of the coupling of quarks.

Chapter 6 describes the development of a trigger for this mode and also of the similar Cabibbo-suppressed decay $D^0 \rightarrow \pi^- \mu^+ \nu_\mu$. To assess combinatoric backgrounds same-sign decays of the form $D^0 \rightarrow h^+ \mu^+ \nu_\mu$ (where $h = K$ or $\pi$) are also kept at a reduced rate.

In Chapter 7 two methods to account for the energy carried by the invisible neutrino are assessed. The expectations of $q^2$ (defined as the invariant mass of the lepton system) in MC and in a simple theoretical model are compared to a sample of data from July and August 2011 which was collected using the dedicated trigger.
The majority of the original research presented in this thesis is focussed on charm physics, i.e. the physics of the third-heaviest (charm or c) quark. To introduce charm physics it is first necessary to describe our current understanding of the interactions of sub-atomic particles (Section 2.1). The framework which describes this is commonly referred to as the Standard Model (SM) of particle physics.

The sections following this introduction cover the aspects most relevant to this thesis: the details of semileptonic decays in the charm sector (Section 2.2), a discussion of the methods typically used to measure the so-called form factors in this type of decay (Section 2.3) and a discussion of using this type of decay to observe $CP$ violation through $D^0 - \bar{D}^0$ mixing (Section 2.4).

### 2.1 The Standard Model

The Standard Model was developed during the twentieth century and has been extensively tested experimentally. It is a gauge theory which describes the interactions of all known particles and forces, with the exception of gravity. This section briefly describes the most relevant aspects of the SM, but more complete descriptions are available elsewhere: for example [2–4].

By ‘gauge theory’ we mean a theory which is invariant under certain local transformations. The group of symmetries resulting from these ‘gauge’ transformations is the gauge group and each gauge symmetry necessitates the introduction of a particle (the ‘gauge boson’) which can be said to mediate the gauge field\(^1\). The interactions between fundamental particles in a particular gauge theory can be described through the exchange of these (usually virtual) gauge bosons.

Also contained within the Standard Model is the Higgs mechanism, which describes the generation of mass via interactions with the Higgs field. This mechanism is an essen-

\(^1\)The name ‘boson’ indicates that these particles have integer spin.
CHAPTER 2. THEORETICAL INTRODUCTION

The SM describes the interactions between three types of particles: leptons, quarks and the field mediators associated with the electromagnetic, strong and weak forces. Each particle has an associated anti-particle which has the same mass and spin but with certain opposite quantum numbers (for instance charge). Leptons and quarks are fermions, i.e. particles of spin $\frac{1}{2}$. Six of each type (plus the associated anti-particles) have been observed and these are commonly divided into three ‘families’ or ‘generations’ based on the particle masses.

Each quark generation contains quarks with charges $+\frac{2}{3}e$ and $-\frac{1}{3}e$ (where $e$ represents the elementary charge\(^1\)). In addition to electric charge, the quarks carry a ‘colour’ charge which can take one of three states (known as red, green and blue). Thus, there are 18 different quark states, and the same is true of the anti-quarks. Some important quark properties are summarised in Table 2.1.

A ‘free’ quark has never been observed. With the exception of the top quark (which is too short-lived to form bound states) quarks have only been observed confined together in colourless pairs (the mesons) or triplets (the baryons) which together are referred to as ‘hadrons’. Since quarks are not observed as physical particles, their masses can only be measured indirectly from the properties of hadrons and depend on the theoretical model used to calculate them. Details of calculations of the quark masses summarised in Table 2.1 are given in [1].

<table>
<thead>
<tr>
<th>Generation</th>
<th>Name</th>
<th>Mass (MeV/$c^2$)</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$u$</td>
<td>$2.49^{+0.81}_{-0.79}$</td>
<td>$+\frac{2}{3}e$</td>
</tr>
<tr>
<td></td>
<td>$d$</td>
<td>$5.05^{+0.75}_{-0.95}$</td>
<td>$-\frac{1}{3}e$</td>
</tr>
<tr>
<td>2</td>
<td>$c$</td>
<td>$1270^{+70}_{-90}$</td>
<td>$+\frac{2}{3}e$</td>
</tr>
<tr>
<td></td>
<td>$s$</td>
<td>$101^{+29}_{-21}$</td>
<td>$-\frac{1}{3}e$</td>
</tr>
<tr>
<td>3</td>
<td>$t$</td>
<td>$172000 \pm 2200$</td>
<td>$+\frac{2}{3}e$</td>
</tr>
<tr>
<td></td>
<td>$b$</td>
<td>$4190^{+180}_{-60}$</td>
<td>$-\frac{1}{3}e$</td>
</tr>
</tbody>
</table>

The leptons consist of three charged particles: the electron ($e$), the muon ($\mu$) and the tau lepton ($\tau$); and three neutral particles: the neutrinos ($\nu_e$, $\nu_\mu$ and $\nu_\tau$). Leptons do not carry colour charge and for this reason do not interact with the strong force. Some important properties of the leptons are summarised in Table 2.2.

Each of the three forces described by the standard model has at least one associated force mediator. Photons ($\gamma$) carry the electromagnetic interaction, the $W^\pm$ and $Z^0$ bosons

---

\(^1\)The sign of electric charge is defined so that the electron has charge $-e$. 

24
2.1. THE STANDARD MODEL

Table 2.2: Some observed properties of the leptons, from [1]. The uncertainty on the last few decimal places of mass is indicated in parentheses where appropriate. Upper limits with a 95% confidence level (CL) are given for the neutrino masses.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Name</th>
<th>Mass (MeV/c²)</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>e⁻</td>
<td>0.510998910(13)</td>
<td>−e</td>
</tr>
<tr>
<td></td>
<td>νₑ</td>
<td>&lt; 0.002 (CL=95%)</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>μ⁻</td>
<td>105.6583668(38)</td>
<td>−e</td>
</tr>
<tr>
<td></td>
<td>νμ</td>
<td>&lt; 0.19 (CL=95%)</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>τ⁻</td>
<td>1776.82(16)</td>
<td>−e</td>
</tr>
<tr>
<td></td>
<td>ντ</td>
<td>&lt; 18.2 (CL=95%)</td>
<td>0</td>
</tr>
</tbody>
</table>

 carry the weak force and gluons (g) carry the strong force. The gluons do not have electric charge, but carry colour charge. Some important properties of the gauge bosons are given in Table 2.3.

In general, the number of independent mediators for a gauge theory is equal to the number of generators of the governing symmetry group. This terminology is somewhat technical, but the generators can be thought of as the dimensions needed to describe the gauge field [5]. For example, the electromagnetic force is governed by the $U(1)$ group and the symmetry of this group is described by a single real number (an angle of rotation). Consequently there is a single mediator (the photon) for the electromagnetic force. The electromagnetic and weak interactions are commonly unified together under the name of the ‘electroweak’ interaction.

Similarly, the gauge group corresponding to the strong force is called $SU(3)$ which has eight generators\(^1\) and hence there are eight allowed gluons, each carrying a different combination of colour charge.

Table 2.3: Some observed properties of the gauge bosons, from [1]. All bosons in the table have spin of 1. The Higgs boson has not been observed so is not included here, but it is described in the text.

<table>
<thead>
<tr>
<th>Force</th>
<th>Mediator</th>
<th>Mass (GeV/c²)</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electromagnetic</td>
<td>γ</td>
<td>&lt; $1 \times 10^{-27}$</td>
<td>0</td>
</tr>
<tr>
<td>Strong</td>
<td>g</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Weak</td>
<td>$W^\pm$</td>
<td>80.399 ± 0.023</td>
<td>±e</td>
</tr>
<tr>
<td></td>
<td>$Z^0$</td>
<td>91.1876 ± 0.0021</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^1\)The group $SU(n)$ corresponds to the group of $n \times n$ unitary matrices with determinant 1. The number of independent generators for $SU(n)$ is $n^2 - 1$. 

25
Photons and gluons are massless\(^1\), while the \(W\) and \(Z\) bosons have masses. In gauge field theories, it is most natural to have massless gauge bosons and theories were developed during the 1960s to explain how the \(W\) and \(Z\) bosons could be massive [6–8].

The explanation relies on ‘spontaneous symmetry breaking’ of the electroweak interaction. This symmetry breaking means it is possible for a theory to satisfy gauge symmetries and yet allow for ground states which do not. This property is analogous to what is commonly seen in ferromagnets. In the ‘ground’ (i.e. low-temperature) state of a ferromagnet, the electrons spins are aligned in one particular direction. This alignment in a particular direction ‘breaks’ the rotational symmetry that would exist if the spins were randomly aligned.

The application of spontaneous symmetry breaking to the electroweak interaction requires the existence of a so-called ‘Higgs’ field and at least one associated gauge boson, the Higgs boson. The simplest theories including a Higgs field require a single, uncharged, spin-0 boson. The existence of the Higgs field is the leading explanation proposed for the masses of \(W\) and \(Z\) bosons, but it has not yet been observed. The observation of the Higgs boson is one of the most important physics goals of experiments at the Large Hadron Collider.

The following sections contain a more mathematical description of the physics described later in the thesis. In particular, the primary decay channel considered in later chapters is \(D^{*+} \rightarrow \pi^+ D^0\) with \(D^0 \rightarrow K^- \mu^+ \nu_\mu\), a semileptonic decay which proceeds through the weak interaction.

### 2.2 Semileptonic Charm Decays

One of the strongest indicators of a weak process is the presence of a single charged lepton in the final state. The semileptonic decay studied in this thesis is illustrated in Fig. 2.1. Here the heavy \(c\) quark within a \(D^0\) meson decays to the lighter \(s\) quark and a lepton pair. A similar process occurs in nuclear \(\beta\)-decay where a \(d\) quark within a proton decays to a \(u\) quark and an \(e^-\bar{\nu}_e\) pair. For an extensive review of leptonic and semileptonic decays of heavy hadrons see [9].

An important feature of semileptonic decays is their simplicity. The weak interaction which underlies these decays is described well by theory and understood well. The matrix element governing the pseudo-scalar to pseudo-scalar meson semileptonic decay \(D^0 \rightarrow K^- \mu^+ \nu_\mu\) can be separated into a weak current (which describes the lepton pair) and a strong current (which describes the evolution from initial- to final-state hadrons). This simplifies the mathematics greatly.

For decays similar to Fig. 2.1 where the virtual \(W\) decays instead to hadrons, this

---

\(^1\)For gluons this is a theoretical value. Gluons have not been observed to have mass, but small non-zero masses have not been ruled out.
2.2. SEMILEPTONIC CHARM DECAYS

Figure 2.1: The Feynman diagram for the $W$-mediated semileptonic decay of $D^0 \rightarrow K^- \ell^+ \nu$.

separation is not possible. The interactions of the final-state hadrons make such a factorisation impossible. For this reason the type of semileptonic decay illustrated in Fig. 2.1 is a good channel through which to study the strong interaction.

Through semileptonic decays, the coupling between $W$ bosons and quarks can be studied. The coupling-strength of the $W$ boson to leptons ($W \rightarrow \ell \nu$) is observed to have a single value, $g$. This is analogous to the coupling-strength of photons to leptons in the electromagnetic interaction, which is proportional to the fine structure constant, $\alpha$. The coupling of the $W$ boson to quarks is more complicated and is governed by a $3 \times 3$ unitary matrix known as the Cabbibo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix is named after the three physicists who first proposed it [10, 11].

The CKM matrix takes the form

\[
V_{\text{CKM}} = \begin{pmatrix}
V_{ud} & V_{us} & V_{ub} \\
V_{cd} & V_{cs} & V_{cb} \\
V_{td} & V_{ts} & V_{tb}
\end{pmatrix}
\] (2.2.1)

where each $V_{ij}$ is complex. Here $u, d, s, c, t$ and $b$ represent up, down, strange, charm, top and bottom quarks respectively. When a quark, $q$, of charge $-1/3$ decays to one ($q'$) of charge $+2/3$ by emitting a virtual $W^-$, the coupling at the vertex is proportional to $V_{qq'}$. For example, the amplitude for the process $b \rightarrow c\ell^-\bar{\nu}$ is proportional to $V_{cb}$. Complex conjugates are taken for $W^+$ emission: the amplitude for $c \rightarrow s\ell^+\nu$ is proportional to $V_{cs}^*$. The CKM matrix thus provides the connection between mass ($q_i$) and weak ($q'_i$) eigenstates:

\[
\begin{pmatrix}
d' \\
s' \\
b'
\end{pmatrix} = V_{\text{CKM}} \begin{pmatrix}
d \\
s \\
b
\end{pmatrix}
\] (2.2.2)

Generally, an $n \times n$ unitary matrix has $n^2$ independent parameters. Six of these relate to the phases and phase differences between the quark fields. Overall phases of the quark fields are unobservable, so five parameters can be removed by absorbing the phases into the quark fields. This leaves the CKM matrix with four free parameters. These are most
commonly taken to be three mixing angles and a complex phase.

The parametrisation used to describe the CKM matrix is an arbitrary choice, but the standard parametrisation is:

\[
V_{\text{CKM}} = \begin{pmatrix}
c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\
-s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\
s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13}
\end{pmatrix}.
\] (2.2.3)

Under this parametrisation, \( s_{ij} = \sin \theta_{ij} \) and \( c_{ij} = \cos \theta_{ij} \), where \( \theta_{ij} \) are the mixing angles and \( \delta \) is the complex phase described above. For historical reasons, the angle \( \theta_{12} \) is also known as the Cabibbo angle, \( \theta_C \).

There is approximately a factor of 100 between the magnitude of the largest and smallest CKM elements. The largest elements lie on the diagonal, meaning that the most probable transitions are those that stay in the same generation: \( u \to d \), \( c \to s \) and \( t \to b \). The current best evaluation of the CKM element magnitudes from [1] is:

\[
|V_{ij}| = \begin{pmatrix}
0.97425 \pm 0.00022 & 0.2252 \pm 0.0009 & (3.89 \pm 0.44) \times 10^{-3} \\
0.230 \pm 0.011 & 1.023 \pm 0.036 & (40.6 \pm 1.3) \times 10^{-3} \\
(8.4 \pm 0.6) \times 10^{-3} & (38.7 \pm 2.1) \times 10^{-3} & 0.88 \pm 0.07
\end{pmatrix}.
\] (2.2.4)

An second parametrisation which more clearly illustrates this difference in element size is the Wolfenstein parametrisation:

\[
V_{\text{CKM}} = \begin{pmatrix}
1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\
-\lambda & 1 - \lambda^2/2 & A\lambda^2 \\
A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1
\end{pmatrix} + O(\lambda^4).
\] (2.2.5)

The four new parameters are related to the standard ones:

\[
\begin{align*}
\lambda &= s_{12} \\
A\lambda^2 &= s_{23} \\
A\lambda^3(\rho - i\eta) &= s_{13}e^{-i\delta}.
\end{align*}
\] (2.2.6)

It is through the complex phase \( \delta \) of the CKM matrix that CP violation appears in the Standard Model. CP is the product of two quantum mechanical operators. Charge conjugation \( (C) \) interchanges particles with their anti-particles and parity \( (P) \) changes the ‘handedness’ of a co-ordinate system, i.e. \( P \) transforms \( (t, x) \to (t, -x) \).

The complex phase \( \delta \) means that neither the mass nor weak eigenstates are eigenstates of CP. Since hadrons and the associated antihadrons are not eigenstates of this operation, the amount of CP violation (CPV) present in a particular decay can be studied by comparing the differences between the two particle types, or the way they interact with each
2.2. SEMILEPTONIC CHARM DECAYS

CPV was first observed in the decay $K^0_L \rightarrow \pi^+ \pi^-$, by Cronin and Fitch [12]. Since then CPV has been observed in other modes, notably in the decays and mixing of $B$ mesons. It is notable that CPV has never been observed in strong interactions – this is known as the ‘Strong CP Problem’ [13]. There are several ways through which CPV can arise naturally in Quantum Chromodynamics (QCD) and the lack of CPV in this interaction is not yet fully understood.

The unitarity of the CKM matrix forces several sum rules of the elements:

$$
\sum_{i=u,c,t} V_{ij}V_{ik}^* = \delta_{jk}, \quad \text{and}
$$

$$
\sum_{i=u,c,t} V_{ji}V_{ki}^* = \delta_{jk}.
$$

(2.2.7)

(2.2.8)

Where $\delta_{jk}$ represents the Kronecker delta. From these relations, the six where $i \neq k$ define triangles in the complex plane. An important part of the LHCb research programme is to make measurements of various components of these triangles to check the consistency of the SM and to search for New Physics.

The combined results of these measurements are collected by the CKM Fitter group and HFAG\textsuperscript{1}. For example, many experiments have contributed to measurements of the CKM elements described by:

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

(2.2.9)

This triangle is often referred to as the Unitarity Triangle. The CKM Fitter summary of measurements [14] from 2010 is shown in Fig. 2.2. Here, the three angles are given by:

$$
\alpha \equiv \arg \left( -\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right),
$$

$$
\beta \equiv \arg \left( -\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right),
$$

$$
\gamma \equiv \arg \left( -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right),
$$

(2.2.10)

(2.2.11)

(2.2.12)

and the triangle is drawn in the complex $\rho - \eta$ plane. The different measurements of each quantity in the triangle are shown as coloured bands.

The Unitarity Triangle is probably the most well-known of the CKM matrix triangles, and many $b$-decay channels studied at LHCb enable measurements of its elements. However the primary decay studied in this thesis, shown in Fig. 2.1, features a $V_{cs}$ ver-

\textsuperscript{1}Updated results and plots available at: http://ckmfitter.in2p3.fr and http://www.slac.stanford.edu/xorg/hfag/
CHAPTER 2. THEORETICAL INTRODUCTION

Figure 2.2: A graphical representation of the Unitarity Triangle in the complex plane. The symbols $\Delta m_s$ and $\Delta m_d$ correspond to mass differences from $B_{s,d}^0$ mixing, while $\varepsilon_K$ is a parameter governing kaon mixing [14].

Some of the simplest semileptonic modes to study are those with a pseudoscalar meson in the initial and final states. Pseudoscalar mesons are those with a total spin of zero and odd parity (denoted $J^P = 0^-$), and the decay $D^0 \rightarrow K^- \mu^+ \nu_\mu$ is a semileptonic decay of this form.

In the remainder of this section, some of the mathematics necessary to describe this type of decay will be discussed. First, it is important to define the symbols that will be used. The heavy (decaying) quark and the lighter (non-decaying) anti-quark are denoted $Q$ (which decays to $q'$) and $\bar{q}$ respectively. The masses of the initial- and final-state pseudoscalar mesons $P$ and $P'$ are labelled $m_P$ and $m_{P'}$, and they have four-momenta $p$ and $p'$. There are only two independent four-vectors in this decay, which we can take to be $q = p - p'$ and $p + p'$.

At energies much less than the $W$ mass, the amplitude for a semileptonic decay
2.2. SEMILEPTONIC CHARM DECAYS

\[ M_Q \rightarrow M'_q \ell \nu \] for a meson \( M \) into another meson \( M' \) is given by [15, 16]:

\[ \mathcal{M}(M_Q \rightarrow M'_q \ell \nu) = -i \frac{G_F}{\sqrt{2}} V_{q'Q} L^\mu H_\mu. \] (2.2.14)

Here \( G_F \) is the Fermi coupling constant, \( V_{q'Q} \) is an element from the CKM matrix and terms of order \( m_Q m_\ell / m_W^2 \) have been neglected. The amplitude is split into a leptonic current (\( L^\mu \)) and a hadronic current (\( H_\mu \)) given by:

\[ L^\mu = \bar{u}_\ell \gamma^\mu (1 - \gamma_5) v_\nu \] (2.2.15)

and

\[ H_\mu = \langle M(p) | J^\mu_{\text{had}} | M'(p') \rangle. \] (2.2.16)

The leptonic current is written in terms of the Dirac spinors \( u \) and \( v \) for the leptons.

The Standard Model weak interaction proceeds by a so-called “\( V - A \)” charged weak current operator. This means the hadronic current can be written: \( J^\mu = V^\mu - A^\mu \), where the vector coupling \( V^\mu \) carries linear momentum, while the axial vector (or pseudovector) coupling \( A^\mu \) carries both linear momentum and angular momentum. When pseudoscalar mesons are in both the initial and final states, the axial term disappears. Assuming that \( M \) and \( M' \) are both pseudoscalars \( P \) and \( P' \), the hadronic current can then be decomposed in terms of so-called “form-factors” [17] as

\[ \langle P'(p')|V_\mu|P(p)\rangle = F_1(q^2) \left( (p + p')^\mu - \frac{m^2_{P} - m^2_{P'}}{q^2} q^\mu \right) + F_0(q^2) \frac{m^2_{P} - m^2_{P'}}{q^2} q^\mu. \] (2.2.17)

Here, \( F_0(q^2) \) and \( F_1(q^2) \) are the longitudinal and transverse\(^1\) form-factors. The use of form-factors is quite common in effective field theories to condense some more complicated theoretical terms into a single function which is accessible experimentally. Here, they are defined such that \( F_0(0) = F_1(0) \) to remove singular behaviour at \( q^2 = 0 \).

A second common way to write (2.2.17) is:

\[ \langle P'(p')|V_\mu|P(p)\rangle = f_+(q^2)(p + p')^\mu + f_-(q^2)(p - p')^\mu, \] (2.2.18)

where \( f_+(q^2) = F_1(q^2) \) and

\[ F_0(q^2) = f_+(q^2) + \frac{q^2}{m^2_{P} - m^2_{P'}} f_-(q^2). \] (2.2.19)

\(^1\)These two terms are used because the form-factors can be associated with the transfer of a particle with quantum numbers \( J^P = 0^+ \) and \( J^P = 1^- \) respectively.
CHAPTER 2. THEORETICAL INTRODUCTION

This can be simplified even further in the limit of \( m_\ell \rightarrow 0 \). All terms in (2.2.17) featuring \( q^\mu \) are multiplied by a factor of \( m_\ell \), so can be neglected in this limit [18]. Thus, we are left with a simplified expression for the matrix element:

\[
\langle P'(p')|V_\mu|P(p)\rangle = f_+(q^2)(p+p')_\mu,
\] (2.2.20)

which is valid to a good approximation for \( \ell = e \) or \( \mu \). It should be noted that since \( f_+(q^2) \) is semi-empirical, it takes different values for each type of decay \( P_{Q\bar{Q}} \rightarrow P'_{Q'\bar{Q}}\ell \nu \), so more generally the form-factor might be written \( f_+[D, K](q^2) \) for the particular decay considered in this thesis.

Thus, using (2.2.14), the differential decay width (up to order \( m_\ell^2 \)) for the decay \( D^0 \rightarrow K^-\ell\nu \) is given by:

\[
d\Gamma = \frac{G_F^2|V_{cs}|^2}{24\pi^3}[f_+(q^2)]^2p_3^3K,
\] (2.2.21)

where \( p_3^3 \) is the magnitude of the kaon’s 3-momentum in the rest frame of the \( D^0 \).

Several models exist for the form of \( f_+(q^2) \). The most common is the simple single-pole form:

\[
f_+(q^2) = \frac{f_+(0)}{1 - q^2/m_{\text{pole}}^2}.
\] (2.2.22)

In the modified pole model [19], it takes the form:

\[
f_+(q^2) = \frac{f_+(0)}{(1 - q^2/m_{\text{pole}}^2)(1 - \alpha_p q^2/m_{\text{pole}}^2)}
\] (2.2.23)

where \( \alpha_p \) is a parameter to be determined by experiment. In the case of \( \alpha_p = 0 \), the two expressions are equivalent.

The pole mass \( m_{\text{pole}} \) is the mass of the lowest-lying \( Q\bar{Q} \) meson satisfying several conditions on its quantum numbers [9]. In the decay \( D^0 \rightarrow K^-\ell\nu \), the quark transition is \( c \rightarrow s \), and the pole mass is given by \( m(D_s^0) = 2.11 \text{ GeV} \).

Similarly, the Isgur-Scora-Grinstein-Wise (ISGW) model [20] gives

\[
f_+(q^2) = \frac{f_+(0)(1 + \alpha_I q_{\text{max}}^2)}{[1 - \alpha_I(q^2 - q_{\text{max}}^2)]^2}.
\] (2.2.24)

In this model, \( q_{\text{max}}^2 \) is the kinematic limit of \( q^2 \) (in the decay \( D^0 \rightarrow K^-\ell\nu \), this is the case where the \( K^- \) is created at rest in the \( D^0 \) rest-frame), and \( \alpha_I \) is a model parameter.

Finally, a fourth common model for the form-factor is the exponential form:

\[
f_+(q^2) = f_+(0)e^{\alpha q^2},
\] (2.2.25)

where again, \( \alpha \) is a parameter to be determined by experiment.

By integrating (2.2.21) over the kinematically allowed region of \( q^2 \) and assuming the
simple pole model with a pole at the $D_s^+$, we have:

$$\Gamma(D^0 \rightarrow K^- \ell \nu) = |V_{cs}|^2 |f_+(0)|^2 (1.54 \times 10^{10}) s^{-1}. \quad (2.2.26)$$

Thus, if one assumes a value for $|V_{cs}|$, the measured decay rate for $D^0 \rightarrow K^- \ell \nu$ can be used to extract $|f_+(0)|$.

One of the most interesting parameters to compare between experiment and theoretical predictions is $m_{pole}$. So when analysing data of this decay mode, the value of $m_{pole}$ can be allowed to vary and determined along with the form factor.

### 2.3 Form Factor Measurements

A variety of techniques have been used to determine the form factors for $D^0 \rightarrow K^- \mu^+ \nu_\mu$ (or for the topologically similar modes $D^0 \rightarrow \pi^- \mu^+ \nu$ and $D^0 \rightarrow K^- / \pi^- e^+ \nu$). A list of some of the more recent studies includes those by the CLEO [21], BaBar [22], FOCUS [23], Tagged Photon Spectrometer [24], E687 [25] and Belle [26] collaborations.

It is most common to measure the shape of the form factors as a function of $q^2$ instead of determining an overall normalisation. The definition used in this analysis of $q^2$ is the invariant mass squared of the lepton pair, $q^2 = (p_D - p_K)^2$, where $p_D$ and $p_K$ are the four-momenta of the $D^0$ and kaon respectively.

Working in the rest-frame of the $D^0$, $q^2$ is explicitly given by the equation

$$q^2 = (p_\mu + p_\nu)^2 = (p_D - p_K)^2 = m_D^2 + m_K^2 - 2m_D E_K, \quad (2.3.1)$$

where $E_K$ is the energy of the kaon measured in the $D^0$ rest frame.

Typically the kinematically allowed region of $q^2$ is divided into a certain number of bins. After compensating for effects of detector resolution and acceptance on $q^2$, an analytic integration of (2.2.21) can be performed over the momentum range in each bin. The expected number of events from the different model predictions can then be tested against the data.

There are several backgrounds (many with relatively high cross-sections) to consider when measuring the mode $D^0 \rightarrow K^- \mu^+ \nu_\mu$. In the cases where a muon and kaon are correctly identified, the most important contributions potentially come from three modes with a $\mu - K$ pair and additional particles in the decay. In these, the some of the ‘missing’ energy ascribed to the the neutrino is instead carried by the missed third particle, and the genuine pair can be combined with a mass below the $D^0$ mass. Firstly, $D^0 \rightarrow K^{*-} \mu^+ \nu$ (BR = 2%) and $D^+ \rightarrow K^{*0} \mu^+ \nu$ (BR = 6%), where the $K^*$ decays to a $K \pi$ pair. These two modes have the same final state as $D^0 \rightarrow K^- \mu^+ \nu_\mu$, but with the addition of an extra pion.
Similarly, the third important background with a muon in the final state is the decay $D^+_s \to \phi \mu^+ \nu$ (BR $\simeq 2\%$) where the $\phi$ decays to a $K^+ K^-$ pair. This has a high branching fraction and the same final state, but with an additional kaon.

The second class of important backgrounds are those where the kaon or muon is mis-identified. For example, were a pion to be mis-identified as a muon, the decay $D^0 \to K^\pi$ would be a potential background. However, the closeness of the muon and pion masses means this decay could be removed by excluding events where the $\mu$-$K$ mass is very close to the $D^0$ mass. The largest contributions to backgrounds from this second class are expected to be from three-body charm decays such as $D^0 \to K^\pi^+ \pi^0$ (BR $= 13\%$) or $D^0 \to \pi^- \mu^+ \nu$ (BR $= 0.3\%$).

A third component is the combinatorial background in a busy experimental environment – by this we mean candidates which do not come from any signal but are instead due to “random” groups of tracks which mimic the true signal. Typically analyses use wrong-sign candidates to understand this type of background. One can search for the physically impossible decay $D^0 \to K^+ \mu^+ \nu$ for example, and since one would expect two positive random tracks to have the same probability of forming a $D$ vertex as two “correctly-charged” ones, these events can provide a method of eliminating this background.

Some difficulties can arise in precisely calculating the $q^2$ value for each candidate, since it is necessary to add the momenta of the muon and the “missing” neutrino in the event. Several different methods have been used to solve this problem. The FOCUS Collaboration used a method called clone closure, while the Tagged Photon Spectrometer
Collaboration attempted to calculate the missing momentum algebraically, using the $D^*$ mass as a constraint.

For a three-body decay including one neutrino, enough can be known from the measured momenta to solve for the missing momentum algebraically, to give a quadratic equation with a pair of allowed solutions (see for example [27]). The flight direction of the $D^0$ can be reconstructed from the positions of the primary and $\mu - K$ vertices. This vector can be used to decompose the $\mu - K$ pair’s momentum ($p_{\mu K}$) into components parallel ($p_{\mu K}^\parallel$) and perpendicular ($p_{\mu K}^\perp$) to the unit vector along the $D^0$ flight direction $\hat{F}$.

\[ p_{\mu K}^\parallel = (\hat{F} \cdot p_{\mu K})\hat{F} \quad (2.3.2) \]
\[ p_{\mu K}^\perp = p_{\mu K} - p_{\mu K}^\parallel. \quad (2.3.3) \]

From conservation of momentum:

\[ p_{\mu K}^\perp = -p_{\nu}^\perp \quad (2.3.4) \]
\[ p_D = p_{\mu K}^\parallel + p_{\nu}^\parallel \quad (2.3.5) \]
\[ E_D = E_{\mu K} + E_{\nu} \quad (2.3.6) \]

Here, each $E$ represents the energy of the of the subscripted particle. By squaring (2.3.6),

\[ E_D^2 = m_D^2 + p_D^2 = (E_{\mu K} + E_{\nu})^2, \quad (2.3.7) \]

we arrive at a quadratic equation in $|p_D^\parallel|$:

\[ 0 = a|p_{\nu}^\parallel|^2 + b|p_{\nu}^\parallel| + c \quad (2.3.8) \]

where

\[ a = 4(p_{\mu K}^\perp + m_{\mu K}^2) \quad (2.3.9) \]
\[ b = 4|p_{\mu K}^\parallel|(2p_{\mu K}^\perp - (m_D^2 - m_{\mu K}^2)^2 \quad (2.3.10) \]
\[ c = 4p_{\mu K}^\perp(p_{\mu K}^\parallel + m_D^2 - (m_D^2 - m_{\mu K}^2)^2 \quad (2.3.11) \]

The quadratic equation (2.3.8) can be solved to give two solutions, which except for the case of non-physical solutions are generally difficult to separate.

Historically at LHCb, in addition to the algebraic solution method discussed above, two of the most common methods in semileptonic analyses are the $k$-factor and corrected mass methods. The first uses simulated events to provide a correction factor to the value of the mother (i.e. $D^0$) momentum based on the observed momenta of its two visible daughter particles. The second uses a simple formula to correct the mother particle’s
reconstructed mass ($M$):

$$M_{\text{corr}} = \sqrt{M^2 + |p_{\text{miss.}}|^2 + |p_{\text{miss.}}|},$$  \hspace{2cm} (2.3.12)$$

where $p_{\text{miss.}}$ is the missing momentum transverse to the direction of flight defined by a line between the primary and decay vertices.

To estimate the $D^0$ momentum, the E653 Collaboration [28] constructed a momentum-estimator:

$$|p_{\text{est.}}| = \frac{M_{D^0} E_{\text{vis.}}}{\sqrt{p_{\text{vis.}}^2 + m_{\text{vis.}}^2}},$$  \hspace{2cm} (2.3.13)$$

where $m_{\text{vis.}}$, $E_{\text{vis.}}$ and $p_{\text{vis.}}$ are the invariant mass, energy and transverse momentum (to the $D^0$ flight direction) of the visible decay products.

Each of these methods has positive and negative characteristics, which will be discussed in later chapters.

In addition to measuring the charm form factors, many experiments have also determined some properties of the overall $D^0 \to K^- \mu^+ \nu_\mu$ rate. For example, the number of $D^0 \to K^- \pi^+$ events is expected to be well measured at LHCb and this decay comprises a crucial part of the charm cross-section measurements. It is very common in the literature to measure the ratio $\Gamma(D^0 \to K^- \mu^+ \nu_\mu) / \Gamma(D^0 \to K^- \pi^+)$ since this number is more well-predicted by theory than many other quantities. If the branching ratio for the two-body decay is well-known, it enables a measurement of the branching ratio for $D^0 \to K^- \mu^+ \nu_\mu$.

### 2.4 $D^0 - \overline{D^0}$ Mixing in Semileptonic Decays

In the charm sector many channels allow $CP$-violation parameters to be measured. In this section, the mathematics of measuring $D^0 - \overline{D^0}$ mixing in decays to semileptonic final states will be described. More complete reviews of $CP$ violation in meson decays are available elsewhere (for example [1]). The search for charm-meson mixing has been an active area of research in charm physics since the first evidence for the process was presented in 2007 [29].

When produced, neutral $D$ mesons are in flavour eigenstates, either $D^0$ or $\overline{D^0}$.$^1$ These are however not eigenstates of the weak-interaction Hamiltonian, and one can write

$$|D_{1,2}\rangle = p|D^0\rangle \pm q|\overline{D^0}\rangle,$$  \hspace{2cm} (2.4.1)$$

where the mass eigenstates have masses and widths $m_1$, $\Gamma_1$ and $m_2$, $\Gamma_2$. The complex numbers $p$ and $q$ satisfy $|p|^2 + |q|^2 = 1$. The time evolution of the two mass eigenstates is

---

$^1$In this section we adopt the phase convention that $CP|D^0\rangle = +|\overline{D^0}\rangle$. 
2.4. \(D^0 - \bar{D}^0\) MIXING IN SEMILEPTONIC DECAYS

governed by the Schrödinger equation and

\[ |D_i(t)\rangle = e^{-im_i t - \frac{i}{2} \Gamma_i t} |D_i(t = 0)\rangle, \]

which has eigenvalues

\[ \omega_i \equiv m_i - \frac{i}{2} \Gamma_i. \]

In terms of the flavour eigenstates the Schrödinger equation for this system gives

\[ i \frac{\partial}{\partial t} \begin{pmatrix} D^0(t) \\ \bar{D}^0(t) \end{pmatrix} = \begin{pmatrix} M - \frac{i}{2} \Gamma & i \\ \frac{i}{2} \Gamma & \bar{M} \end{pmatrix} \begin{pmatrix} D^0(t) \\ \bar{D}^0(t) \end{pmatrix}, \]

where \(M\) and \(\Gamma\) are Hermitian matrices, representing a 2 x 2 effective Hamiltonian. The off-diagonal elements of the matrices \(M\) and \(\Gamma\) describe the mixing. If the symmetry \(CPT\) is conserved, this forces \(M_{11} = M_{22}\) and \(\Gamma_{11} = \Gamma_{22}\).

This equation can be solved to give:

\[ \left( \frac{q}{p} \right)^2 = \frac{M_{12}^* - \frac{i}{2} \Gamma_{12}^*}{M_{12} - \frac{i}{2} \Gamma_{12}}, \]

The real and imaginary parts of the eigenvalues represent the masses and widths of the two eigenstates. The eigenvalues are given by

\[ \omega_{1,2} = (M - \frac{i}{2} \Gamma) \pm \frac{q}{p} (M_{12} - \frac{i}{2} \Gamma_{12}) \]

which can be compared to (2.4.3) above. Here \(M = (M_1 + M_2)/2\) and \(\Gamma = (\Gamma_1 + \Gamma_2)/2\).

Two dimensionless parameters are normally used to characterise charm mixing: \(x \equiv \Delta m/\Gamma\) and \(y \equiv \Delta \Gamma/2\Gamma\), where \(\Delta m = |m_1 - m_2|\) and \(\Delta \Gamma = |\Gamma_1 - \Gamma_2|\). For mixing to occur, it is necessary that at least one of \(x\) or \(y\) be non-zero.

In many examples of charm mixing, doubly Cabibbo-suppressed (DCS) decays to the same final state provide interference terms. For a semileptonic decay of the form \(D^0 \rightarrow K^- \ell^+ \nu\), no DCS decays exist and hence the charge of the kaon uniquely tags the state of the decaying \(D^0\) or \(\bar{D}^0\).

In this case, assuming small mixing rates, the decay time distribution is [30, 31]:

\[ R_{\text{mix}}(t) \approx R_{\text{unmix}}(t) \frac{x^2 + y^2}{4} (\Gamma t)^2, \]

where \(t\) is the proper time of the decay and \(R_{\text{unmix}}(t) \propto e^{-\Gamma t}\). Often the characteristic \(D^0\) lifetime \(\tau_{D^0}\) is used in place of \(1/\Gamma\) in the above expression.

\[1\text{Technically, the two matrices describe the off-shell (or dispersive) and on-shell (absorptive) parts of the mixing.}\]
The time-integrated mixing rate\(^1\) is then given by:

\[
r_{\text{mix}} = \int_0^\infty R_{\text{mix}}(t) \, dt = \frac{x^2 + y^2}{2}.
\]

(2.4.8)

This quantity is the one most commonly measured by experiments searching for mixing through this channel. The world average for this number is \(r_{\text{mix}} = (1.30 \pm 2.69) \times 10^{-4}\) [1]. Some measurements of this quantity are given in Table 2.4.

<table>
<thead>
<tr>
<th>Year</th>
<th>Expt</th>
<th>Final state(s)</th>
<th>(r_{\text{mix}} \times 10^{-3})</th>
<th>90% C.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>Belle</td>
<td>(K^{(\ast)}e^-\nu_e)</td>
<td>0.13 ± 0.22 ± 0.20</td>
<td>&lt; 0.61 \times 10^{-3}</td>
</tr>
<tr>
<td>2007</td>
<td>BABAR</td>
<td>(K^{(\ast)}e^-\nu_e)</td>
<td>0.04(^{+0.70}_{-0.60})</td>
<td>(−1.3, 1.2) \times 10^{-3}</td>
</tr>
<tr>
<td>2005</td>
<td>Belle</td>
<td>(K^{(\ast)}e^-\nu_e)</td>
<td>0.02 ± 0.47 ± 0.14</td>
<td>&lt; 1.0 \times 10^{-3}</td>
</tr>
<tr>
<td>2005</td>
<td>CLEO</td>
<td>(K^{(\ast)}e^-\nu_e)</td>
<td>1.6 ± 2.9 ± 2.9</td>
<td>&lt; 7.8 \times 10^{-3}</td>
</tr>
<tr>
<td>2004</td>
<td>BABAR</td>
<td>(K^{(\ast)}e^-\nu_e)</td>
<td>2.3 ± 1.2 ± 0.4</td>
<td>&lt; 4.2 \times 10^{-3}</td>
</tr>
<tr>
<td>2002</td>
<td>FOCUS</td>
<td>(K^+\mu^-\nu_\mu)</td>
<td>−0.76(^{+0.99}_{-0.93})</td>
<td>&lt; 1.01 \times 10^{-3}</td>
</tr>
<tr>
<td>1996</td>
<td>E791</td>
<td>(K^+\ell^-\nu)</td>
<td>1.1(^{+3.0}_{-2.7})</td>
<td>&lt; 5.0 \times 10^{-3}</td>
</tr>
</tbody>
</table>

The PDG also gives world averages for the mass difference and width separately. Presently the values of these have been measured to be:

\[
\Delta m = x\Gamma = (2.35^{+0.59}_{-0.63}) \times 10^{-2} \text{ ps}^{-1}
\]

(2.4.9)

and

\[
\Delta\Gamma/\Gamma = 2y = 1.66 \pm 0.32 \times 10^{-2}.
\]

(2.4.10)

These numbers can be compared to mixing in the neutral \(B\) meson sector. The values for \(B^0_d\) mesons [32] are

\[
\Delta m_d = 0.508 \pm 0.004 \text{ ps}^{-1} \quad \text{and} \quad |\Delta\Gamma_d/\Gamma_d| = 0.011 \pm 0.037,
\]

(2.4.11)

while for \(B^0_s\) mesons, the numbers are different by an order of magnitude:

\[
\Delta m_s = 17.78 \pm 0.12 \text{ ps}^{-1} \quad \text{and} \quad \Delta\Gamma_s/\Gamma_s = 0.072^{+0.049}_{-0.051}.
\]

(2.4.12)

---

\(^1\)Some constants (in particular, factors of \(\Gamma\)) were neglected in the above discussion to simplify the explanation but are re-included in this expression.
3.1 The Large Hadron Collider

3.1.1 Introduction

The Large Hadron Collider (LHC) [33–35] is currently the world’s highest-energy particle accelerator. It was constructed approximately 100 m below ground at the European Centre for Nuclear Research (CERN) in the tunnel formerly used by the Large Electron-Positron collider (LEP) [36]. Figure 3.1 shows a schematic drawing of the position of the LHC (including the SPS and the four largest LHC experiments) beneath the Franco-Swiss border.

![Figure 3.1: An underground schematic of the SPS, the LHC and the four largest LHC experiments, from [37].](image)

Two of the most important parameters of any particle accelerator are energy and luminosity. The higher the energy of a particle collision, the more is available for the
CHAPTER 3. THE LHCB EXPERIMENT

production of new particles. It is usually characterised by the centre-of-mass energy, $\sqrt{s}$.

For a colliding beam machine such as the LHC, the particles make multiple passes through a ring of electromagnetic fields which provide acceleration in the forward direction and causes the particle trajectories to curve. The magnitude and frequency of the fields is increased and carefully controlled to increase the beam energy while still keeping it in a stable orbit.

Luminosity is a measure of particle flux and is proportional to the rate of interactions. It is important to distinguish the instantaneous luminosity ($\mathcal{L}$) from the integrated luminosity ($L$). The integrated luminosity is the integral over time of $\mathcal{L}$ and is the ratio of yield ($N$) to cross-section ($\sigma$) for any process:

$$N = L\sigma$$  \hspace{1cm} (3.1.1)

The usual units for instantaneous luminosity are cm$^{-2}$s$^{-1}$, while integrated luminosity is usually quoted using the traditional unit of the barn$^1$, for example in units such as the inverse picobarn (pb$^{-1}$).

Like all modern particle colliders, the LHC accelerates bunched beams. For two bunches containing $n_1$ and $n_2$ particles colliding head-on, spread uniformly over a cross-sectional area $A$, the instantaneous luminosity is given by

$$\mathcal{L} = f \frac{n_1 n_2}{A}$$  \hspace{1cm} (3.1.2)

where $f$ is the collision frequency. For multiple bunches, the luminosity increases proportionally to the number of colliding bunches in each beam. In reality, particles are not spread uniformly in space, but typically have some sort of Gaussian distribution.

Equation 3.1.2 is commonly rewritten in terms of two beam parameters: the amplitude function ($\beta$) and the transverse emittance ($\varepsilon$). Both are functions of the position along the path length in the beam direction.

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi \sqrt{\varepsilon} \beta^*}$$  \hspace{1cm} (3.1.3)

Here, $\beta^*$ represents the amplitude function evaluated at the IP. The amplitude function describes the beam size and focussing$^2$. Generally, low values of the amplitude function indicate a beam which is tightly focussed while higher values indicate the particles are less focussed and travelling more parallel. For high luminosity, a very tightly focussed beam with a correspondingly low $\beta^*$ value is desirable.

The transverse emittance is a measure of the transverse energy of accelerated particles.

$^1$1 barn $\equiv 10^{-28}$m$^{-2}$

$^2$The transverse motion of a particle traversing the accelerator is given by $x(s) \propto \sqrt{\beta(s)} \cos(\psi(s) + \delta)$, where $\delta$ is an integration constant and $\psi(s)$ is the phase.
3.1. THE LARGE HADRON COLLIDER

We work in the co-ordinate system where \( z \) is the beam direction and the \( xy \) plane is the transverse plane. If we assume the beams have Gaussian profiles of width \( \sigma_i \) in the \( i \) direction, the transverse emittance can then be defined component-wise:

\[
\varepsilon_x \equiv \pi \frac{\sigma_x^2}{\beta_x}, \quad \varepsilon_y \equiv \pi \frac{\sigma_y^2}{\beta_y}.
\]  

(3.1.4)

There are thus several variables which can be changed to achieve the high luminosities required by the LHC. The beam currents can be increased by increasing either \( f \) or the number of protons per bunch. When operating at peak design luminosity, the LHC beams will have approximately \( 10^{11} \) protons per bunch and almost 3000 bunches per beam. It is also possible to decrease \( \beta^* \) with the use of focusing magnets close to the IP or to decrease \( \varepsilon \) through damping. The LHC design for LHCb’s IP (named IP8) is to have tuneable \( \beta^* \) in the range 1–50 m.

3.1.2 LHC Design

The LHC is a 27 km circumference synchrotron designed to collide protons with \( \sqrt{s} = 14 \) TeV at a frequency of 40 MHz. The LHC is designed to achieve luminosities of order \( 10^{34} \text{ cm}^{-2}\text{s}^{-1} \). It is also able to collide heavy ions (lead nuclei), but since LHCb is not designed to acquire data from this second type of collision, in this thesis we focus only on proton-proton collisions.

A schematic diagram of the CERN accelerator complex (of which the LHC is just one part) is pictured in Figure 3.2. A complex chain of acceleration and storage is necessary to achieve the design energy of the LHC, but many components in this chain existed before construction of the LHC began. In the first step, protons are obtained from hydrogen gas and are accelerated to an energy of 50 MeV in a linear accelerator (LINAC2). Next, the energy is raised to 1.4 GeV as the protons pass through the Proton Synchrotron Booster and then to 25 GeV by the Proton Synchrotron (PS).

The final stage before injection into the LHC is the Super Proton Synchrotron (SPS), which accelerates the protons to 450 GeV. The protons are injected as two counter-rotating beams into the LHC by the SPS. The LHC is designed to then accelerate the protons to 7 TeV per beam. For the year 2010, the protons were only accelerated to a maximum of 3.5 TeV per beam.

The beams are accelerated and contained by 1232 dipole and 392 quadrupole electromagnets, each of which is super-conducting to provide sufficiently high fields. In addition, around 8000 smaller magnets are used for fine-tuning of the beams.

The LHC is divided into octants and beams are made to collide inside experiments at four IPs around the LHC ring. This is pictured in Figure 3.3. ALICE [39] is located at IP2, ATLAS [40] at IP1, CMS [41] at IP5 and LHCb [42] at IP8. A dump point for both
beams is located in octant 6.

Some important LHC parameters (reproduced from [33]) are given in Table 3.1.

### 3.2 The LHCb Experiment

Proposals to create a heavy flavour physics experiment at the LHC were first made in the early 1990s. In the mid-90s the LHC experiments committee (LHCC) requested a joint proposal from the groups working in this field, and the LHCb Letter Of Intent was released in 1995 [44].

This was followed in 1998 by the first technical proposal [45] and five years later by a re-optimised version [46]. The most important changes in the revised version were reductions in the overall amount of material traversed by particles and improved trigger efficiency.

The LHCb detector (shown in Figure 3.4) is a forward-spectrometer similar in design to many fixed-target detectors. This means it is designed to cover a relatively small solid angle around the beam axis, and is thus a very different design to the three other LHC detectors which are designed for almost complete angular coverage. The detector is pri-
3.2. THE LHCB EXPERIMENT

Figure 3.3: A schematic diagram of the LHC, from [43]. The beams collide inside the experiments located in octants 1, 2, 5 and 8.

...arily designed to enable study of CP violation and rare decays in the \( b \)- and \( c \)-meson sectors.

Although CP violation was first observed in the decays of kaons [47], the \( b \)-meson sector has comparatively large CP violation effects and with a large number of possible decay modes is experimentally accessible. Two large experiments – \( \bar{B}A\bar{B}R \) (at PEP-II) and BELLE (at KEK-B) – known as \( B \) factories were designed specifically to probe this sector. At present the results from the \( B \) factories and other more general-purpose detectors such as CDF and DØ have been consistent with the CKM mechanism within the Standard Model. One part of the LHCb experimental programme will be to test and complement the results of these experiments.

With very high statistics, LHCb will also be able to measure rare decays of \( b \) mesons. For example, one of the most accessible channels is the flavour changing neutral current decay \( B_s^0 \rightarrow \mu^+ \mu^- \). The CDF experiment has put a limit on the branching ratio of...
Table 3.1: LHC beam parameters relevant for the peak luminosity.

<table>
<thead>
<tr>
<th></th>
<th>At Injection</th>
<th>Collision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beam Data:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proton energy (GeV)</td>
<td>450</td>
<td>7000</td>
</tr>
<tr>
<td>Relativistic gamma</td>
<td>479.6</td>
<td>7461</td>
</tr>
<tr>
<td>Number of particles per bunch</td>
<td>$1.15 \times 10^{11}$</td>
<td>$1.15 \times 10^{11}$</td>
</tr>
<tr>
<td>Number of bunches</td>
<td>2808</td>
<td>2808</td>
</tr>
<tr>
<td>Longitudinal emittance ($4\sigma$) (eVs)</td>
<td>1.0</td>
<td>2.5</td>
</tr>
<tr>
<td>Transverse normalised emittance ($\mu$m rad)</td>
<td>3.5</td>
<td>3.75</td>
</tr>
<tr>
<td>Circulating beam current (A)</td>
<td>0.584</td>
<td>0.584</td>
</tr>
<tr>
<td>Stored energy per beam (MJ)</td>
<td>23.3</td>
<td>362</td>
</tr>
<tr>
<td><strong>Peak Luminosity Related Data:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS bunch length (cm)</td>
<td>11.24</td>
<td>7.55</td>
</tr>
<tr>
<td>RMS beam size at IP1 and IP5 ($\mu$m) ($\beta^* = 0.55$)</td>
<td>375.2</td>
<td>16.7</td>
</tr>
<tr>
<td>RMS beam size at IP2 and IP8 ($\mu$m) ($\beta^* \sim 10$)</td>
<td>279.6</td>
<td>70.9</td>
</tr>
<tr>
<td>Peak luminosity in IP1 and IP5 (cm$^{-2}$s$^{-1}$)</td>
<td>—</td>
<td>$1.00 \times 10^{34}$</td>
</tr>
<tr>
<td>Luminosity in IP8 (cm$^{-2}$s$^{-1}$)</td>
<td>—</td>
<td>$\sim 2.00 \times 10^{32}$</td>
</tr>
</tbody>
</table>

$B(B_s^0 \rightarrow \mu^+\mu^-) < 4.7 \times 10^{-8}$ at a 90% confidence level (CL) [48]. This is a very clean channel (with only two muons in the final state) which is expected to be measured well by LHCb. With approximately 0.3 fb$^{-1}$ (achievable within one year of running) it is expected that LHCb will have sufficient statistics to improve on the Tevatron measurements.

LHCb has the ability to measure many other rare decays and CP violation parameters. Full details of the experimental reach of LHCb can be found elsewhere [46].

It is helpful to define the system of co-ordinates used in LHCb. It is a right-handed Cartesian system with a $z$-axis in the direction of beam 1 (towards IP1) and an $x$-axis pointing away from the centre of the LHC ring. The LHC is built at a slight vertical angle (3.6 mrad from horizontal), so the $y-$ axis does not point directly upwards. The LHCb magnet is predominantly vertical, so ‘bending’ occurs in the $xz$ plane.

The detector geometry exploits the fact that inside LHCb, both $b$ and $\bar{b}$ hadrons are produced in the same forward or backward cone (i.e. close to the beam-pipe). Figure 3.5 shows the polar angles of $b$ and $\bar{b}$ quarks inside LHCb as simulated by PYTHIA. The angular acceptance of the LHCb detector is $\pm 250$ mrad in the $yz$ (non-bending) plane and $\pm 300$ mrad in the $xz$ (bending) plane, with a minimum acceptance of approximately 15 mrad. This enables tracks with pseudo-rapidities ($\eta$) in the range $1.6 < \eta < 4.9$ to be measured. LHCb is able to detect approximately 34% of the $B$ mesons while covering around 2% of the total solid angle.

The LHCb detector is designed to operate at a maximum luminosity of approximately $2 \times 10^{32}$ cm$^{-2}$s$^{-1}$, about fifty times lower than the maximum design luminosity of the LHC. To achieve this, the beam is intentionally defocussed close to the interaction point.
3.2. THE LHCb EXPERIMENT

Figure 3.4: Schematic diagram of the LHCb detector, from [42]. Labels are described in-text.

Figure 3.5: Angular distribution of $b$- and $\bar{b}$-quarks inside LHCb, from [45].
CHAPTER 3. THE LHCB EXPERIMENT

... 

The detector was designed with specific physics goals in mind to exploit such high production of $b\bar{b}$ pairs. In particular, LHCB has:

- an efficient trigger which is sensitive to the many possible final states of interest;
- precise vertexing and mass resolution (essential to making precise proper-time measurements and reducing backgrounds) and;
- good particle identification (PID) to distinguish final state particle types.

The major sub-detectors of LHCB are labelled in Figure 3.4. Closest to the interaction point (at left-of-picture) is the Vertex Locator (VELO). Outside there are the tracking system (TT, T1, T2, T3), two ring image Cherenkov detectors (RICH1, RICH2), electromagnetic and hadronic calorimeters (ECAL, HCAL) and a muon detection system (M1–M5). A full description of the LHCB detector is available elsewhere [42, 46]. In the following sections we give an overview of each major detector component.

3.2.1 The Vertex Locator

The innermost part of the LHCB detector is the VELO.¹ Later parts of this thesis describe work on the VELO, so slightly more space will be devoted to this subdetector. A more detailed description can be found in the VELO Technical Design Report [50].

The VELO is a silicon microstrip detector which can provide precise measurements of tracks near to the interaction point. It can be used to measure decay vertex locations, and especially can be used to locate the secondary vertices of $b$ (and $c$) decays. The VELO is comprised of two halves, each containing 21 sensor modules constructed from 300 $\mu$m thick n$^+$-on-n silicon, chosen for its radiation hardness. On average a particle traversing the VELO passes through 16% of a radiation length [51].

A novel feature of the VELO is that the detector halves are retractable and can be moved up to 30 mm away from the beam during periods of unstable or “dangerous” beam conditions which could potentially damage the detector – for example during beam injection. During stable running, the innermost active strips of the detector can be moved within 8 mm of the beam. This proximity to the interaction point allows the measurement of tracks with angles as low as 15 mrad from the beam.

Each module contains two half-disc silicon sensors back-to-back which measure $R$ (the distance from the beam in a perpendicular direction) and $\phi$ (the angle around the

¹Some parts of this section have been previously published in [49].
3.2. THE LHCb EXPERIMENT

beam axis) co-ordinates. Each detector half contains 21 of these modules. Four additional stations of $R$-type sensors behind the interaction point – the pileup counters – are used as part of the LHCb trigger to reject events with multiple interactions.

Each VELO half sits inside a thin-walled aluminium box (an RF-box) which allow the modules to sit in a secondary vacuum, separate to the LHC vacuum. The outsides of the RF boxes have a corrugated shape which, in addition to a 1.5 cm offset along the beam axis allows the two halves to slightly overlap and mesh together. A VELO half and its RF box are shown in Figure 3.6. The final overlapping position of the two RF boxes is shown in Figure 3.7. When closed, the foil sits approximately 5 mm from the beams.

![Figure 3.6: A VELO half shown with its RF box. The two are pictured slightly apart to demonstrate how they fit together, from [42].](image)

The ability of the two halves to overlap is an important design feature. Tracks which pass through this overlap region provide important information on the relative alignment between the halves.

The VELO was designed to enable high-precision reconstruction of vertices with a high signal to noise ratio ($S/N > 14$). The vertex measurement uncertainty depends on the number of tracks, but for an average event the VELO can achieve a resolution of about $42 \mu m$ in the $z$ direction (parallel to the beam axis) and about $10 \mu m$ in the transverse ($xy$) plane. The VELO is able to detect particles with pseudo-rapidities in the range $1.6 < \eta < 4.9$ for tracks originating near the interaction region ($|z| < 10.6$ cm) illustrated in Figure 3.8.

The VELO contains 42 modules (21 in each detector half) and the layout of these is pictured in Figure 3.8. Each module (one is pictured in Figure 3.9) contains two silicon strip detectors back-to-back, measuring $R$ and $\Phi$ co-ordinates. The secondary vacuum
inside the VELO is designed to minimise material inside the VELO and to protect the LHC vacuum. The LHC operates in a vacuum pressure of the order $\sim 10^{-8}$ mbar, while the secondary vacuum is kept below $10^{-4}$ mbar.

During operation, each VELO module consumes approximately 20 W of power, so the VELO has a cooling system to transfer heat to the outside of the system. Each module base (pictured in Figure 3.9) is made of low-mass carbon fibre on a carbon fibre base. This base (known as a ‘paddle’) is the mechanical and thermal link between the sensor and the VELO exterior. The paddles connect to a CO$_2$ cooling system through an aluminium coupler and a soft-metal indium joint. During operation, the VELO is kept within a temperature range of $-25$ to $+10^\circ$C, ideally close to $-5^\circ$C.

Silicon strip sensors have several useful properties. The sensors used in the VELO are 300$\mu$m thick, silicon n-strip on n-bulk (n-on-n) oxygenated sensors. A study detailing the choice of this sensor type and irradiation tests has been published in [52]. They are relatively thin, and do not significantly change particle trajectories, nor absorb large amounts of particle energy. The strips have pitches varying between 40 and 103 $\mu$m. These strip sizes provide excellent spatial resolution, with single-hit resolutions measured as low as 4 $\mu$m.

Each silicon sensor has 2048 strips, which are read out through 16 so-called ‘Beetle’ chips [53]. These chips were specifically designed to meet the LHCb performance requirements, and have been tested to be sufficiently radiation-hard to survive in the LHCb operating environment. The Beetle chips output to four ports (known as ‘links’), and each port contains 32 channels. The chips have multiple adjustable settings to allow calibration.
3.2. THE LHCB EXPERIMENT

Figure 3.8: Schematic cross-section of (top) the VELO in the $xz$ plane at $y = 0$ and (bottom) two sensors in the closed and open positions, from [42].

Figure 3.9: A VELO module. Each VELO half contains 21 modules.
and in particular, the ability to change gain settings is important for the work presented in this thesis.

![Figure 3.10: The raw output from a Beetle for some test pulse data. This shows the average of several events where positive and negative test pulse signals are sent into two channels. Note the four header bits at $x < 0$ preceding the channel output.](image)

The chips accept input data at a rate of 40 MHz. This is then fed into an analogue pipeline with a 4 $\mu$s latency until the Level-0 (L0) trigger decision is made. Level-0 is the hardware component of the LHCb trigger. Figure 3.10 shows a typical output from one link for a single event. The Beetle chips have a circuit able to inject a specified charge into the readout (a test pulse), and this plot shows such an output. Non-digitized data in this ‘raw’ format is called Non-Zero-Suppressed (NZS), and the gain studies described later in this report are based on data of this type.

Once an event passes the Level-0 trigger, the NZS analogue signal is set to TELL1 boards [54] which digitise the signal. These boards are used by almost all sub-detectors within LHCb to provide signal readouts. They have flexible hardware which can be modified to suit each sub-detector. In the case of the VELO, the TELL1 boards perform some pre-processing (such as removing header cross-talk, performing pedestal subtraction and common-mode correction) before outputting digitized signals to the higher-level trigger.

The units for hits in the VELO are Analogue-to-Digital Converter (ADC) counts. One ADC count is equivalent to a charge of 440 electrons ( [55]).

At the beginning of each link output there are four so-called ‘header bits’ which indicate the beginning of an event output. These can be seen below $x = 0$ in Figure 3.10.

---

[50] The LHCb trigger is discussed in greater detail in Section 3.2.5.
3.2. THE LHCB EXPERIMENT

Header cross-talk occurs when there is some leakage of charge between the header bits (only the last two have a significant effect) and the neighbouring channel. The TELL1 board has an algorithm built-in to correct the signal based on the values of the final two headers (H2 and H3).

A channel’s pedestal is the ‘zero’ (i.e. no signal) output value. In an ideal scenario, this would be equivalent to no output charge, but noise in each channel means this is non-zero (and different) for each channel. To account for this, pedestal subtraction is an algorithm applied to each channel in the TELL1 board to calibrate the pedestals.

Common-mode noise is a type of low-frequency noise which is shared by multiple signal channels. The TELL1 boards also correct for this with an algorithm called Linear Common Mode Suppression. For each event a correction is calculated which is then applied to a full analogue link.

Since LHCb began collecting data in 2009, the VELO has been able to measure performance through several metrics. Four important measures are the spatial alignment, vertex resolution, impact parameter (IP) resolution and single-hit resolution.

The VELO was used to reconstruct the first tracks from beam-absorber collisions at the LHC in August 2008 and a first alignment was obtained [56]. The VELO alignment procedure [57] determines the sensor, module and VELO half alignment. Sensor and module alignment are performed by a fit to the residual distribution and by using the Millepede algorithm [58] respectively. A combination of overlap track hits and the position of primary vertices reconstructed by each half alone are used to align the halves.

The motion of the two halves between fills and the hit precision required mean that alignment is a critical task. The module and sensor alignments are known to better than 5 µm. Figure 3.11 shows the sensor vertical (y) and horizontal (x) alignments, which have an RMS of 4.4 µm. With more study, the VELO group expects to reduce this number to less than 2 µm.

![Overview of misalignments](image)

Figure 3.11: Summary of the horizontal and vertical misalignments for all VELO sensors.
CHAPTER 3. THE LHCB EXPERIMENT

The half alignment is also typically known to 5 µm precision, varying on a run-to-run basis due to the VELO halves’ retraction and reinsertion for each fill.

The precision with which a vertex is known depends strongly on the number of tracks \((N)\) contributing to the vertex. To evaluate the vertex resolution, the tracks in a particular event are randomly split into two groups and the primary vertex is reconstructed for each. If the two vertices have the same \(N\), the vertex positions are compared and the width of the resulting residual histogram yields the resolutions.

The vertex resolutions in the \(x\) and \(y\) directions are shown in Figure 3.12 as a function of \(N\). For vertices with more than 20 associated tracks, the resolutions display a \(1/\sqrt{N}\) dependence, illustrated by the solid curves.

![Figure 3.12: Vertex resolutions in \(x\) (red) and \(y\) (blue). Above \(N = 20\) a curve proportional to \(1/\sqrt{N}\) is fitted to each.](image)

At the design luminosity and energy, the average number of tracks per primary vertex is expected to be close to 25. For vertices with 25 tracks, we measure resolutions of 15.7 µm, 15.4 µm and 90.4 µm in \(x\), \(y\) and \(z\) respectively.

The impact parameter (IP) of a track is the distance of closest approach between a track and the primary vertex. Identifying displaced vertices through track IP is an important requirement of LHCb.

As a result of multiple scattering, the width of the IP distribution depends on track transverse momentum \((p_T)\). For all tracks in an event, in each bin of \(1/p_T\), the IP distribution is fitted with a Gaussian function and the resolution is given by the width of the fitted Gaussian. As an illustration, the \(x\) component of IP resolution is shown in Figure 3.13 as a function of \(1/p_T\).

For the highest-momentum tracks, we measure IP resolutions below 25 µm. We see some difference between simulation and data, and work is on-going to understand this.

Tracks which pass through several strips in a sensor enable the calculation of weighted cluster centres, and single hit resolutions as low as 4 µm, much better than the binary resolution. This is illustrated in Figure 3.14. The resolution depends linearly on strip
3.2. THE LHCB EXPERIMENT

![Graph](image)

**Figure 3.13**: The $x$ component of IP resolution for 2010 data (black) and simulation (red) with linear best-fit curves overlaid.

pitch and also on the track angle.

![Graph](image)

**Figure 3.14**: Single hit resolution for the binary estimate (red) and for low-angle (blue) and high-angle (black) tracks from collisions.

### 3.2.2 The Tracking System

The VELO is often considered part of the tracking system, but since it is particularly important for later parts of this thesis it has been considered separately above. The LHCb tracking system uses two main technologies: silicon trackers and drift-time detectors in four tracking stations.

Silicon trackers are used in the Tracker Turicensis (TT), which is the tracking station closest to the interaction point, located before the magnet. The remaining three stations
(labelled T1–T3 in Figure 3.4) utilize silicon detectors closest to the beam in the Inner Trackers (IT) and drift-time detectors in Outer Trackers (OT) further from the beam.

The TT consists of two pairs of single-sided silicon detector layers. In each pair, the layers have a $5^\circ$ relative rotation between them. This enables transverse momentum components to be measured more easily. This structure is given the name $x-u-v-x$, as the two outer ($x$) layers have strips aligned vertically (and hence measure $x$ co-ordinates). The names $u$ and $v$ are given to the two rotated alignments. The layouts of both the TT and IT are illustrated in Figure 3.15, where the $x-u-v-x$ structure is specifically illustrated for the TT.

![Diagram of TT and IT layouts](image)

Figure 3.15: Illustration of the TT (left) and IT (right) layout with indications of the relative size (the two scales are not equivalent), from [59]. The $x-u-v-x$ structure is shown for the TT, while only an $x$ layer is shown for the IT.

The active area of the TT is approximately $8.3 \text{ m}^2$. The sensors of the TT are $500 \mu\text{m}$ thick, $9.64 \text{ cm}$ wide and $9.44 \text{ cm}$ long. They contain 512 signal strips which have a pitch of $183 \mu\text{m}$.

Layers of single-sided silicon detectors are also used in the IT where they are arranged in an approximate “+” shape around the beam, also using the $x-u-v-x$ layout. The sensors comprising the IT are $7.6 \text{ cm}$ wide and $11 \text{ cm}$ long, containing 384 readout strips with strip pitch of $198 \mu\text{m}$. This enables a spatial resolution of approximately $50 \mu\text{m}$ for the IT system to be achieved.

The IT does not cover the full acceptance (unlike the TT), but instead covers the high-multiplicity region closest to the beam pipe. The remainder of the acceptance region is covered by the OT.

Surrounding the IT parts of each tracking station is an Outer Tracker. Lower particle fluxes further from the beam mean that a different technology can be used in this region. Similarly to TT and IT detectors, each OT station is made of two pairs of relatively rotated modules in the $x-u-v-x$ layout, but in OT modules drift-tubes are used instead of
silicon strips. The OT is designed to provide precise momentum resolution over a large acceptance area (the total active area of one station is $5971 \times 4850$ mm$^2$).

The OT layers are made of drift-tubes with inner diameters of 4.9 mm. The tubes are filled with a mixture of Argon (70%) and CO$_2$ (30%) gases, which yield a fast drift time (of around 50 ns) and a spatial resolution of approximately 200 $\mu$m. This technology has a coarser resolution than the silicon strips used in the TT and IT, but is relatively inexpensive, which makes it suitable to cover a large acceptance area.

### 3.2.3 The Magnet

The LHCb magnet is a conventional (i.e. non-superconducting) dipole magnet, with an integrated magnetic field of approximately 4 Tm. The field has been thoroughly mapped, and is known within a precision of $4 \times 10^{-4}$. The magnet has the capability of switching polarity. This will be important to study systematic asymmetries of the detector, which could affect CP-violation measurements.

The magnetic field is illustrated in Figure 3.16. The magnet is comprised of two coils, each with 225 turns. They provide a total excitation current ($NI$) of 1.3 MA per coil.

![Illustration of the y component of the LHCb magnetic field and its position relative to the silicon trackers, from [46].](image)

Figure 3.16: Illustration of the $y$ component of the LHCb magnetic field and its position relative to the silicon trackers, from [46].
3.2.4 Particle Identification

Particle Identification (PID) at LHCb is performed by RICH1 and RICH2 (the two Ring Imaging Cherenkov detectors), by the muon system and by the electromagnetic and hadronic calorimeters.

RICH detectors record the Cherenkov light emitted as particles pass through the detector. The LHCb RICH detectors are designed to perform PID over a large range of momenta, and especially to separate pions and kaons. RICH1 sits between the VELO and the magnet and the much larger RICH2 sits after the magnet and silicon trackers.

Cherenkov radiation is the type of electromagnetic radiation emitted when a charged particle passes through a medium faster than the speed of light in that medium. The effect is an optical analogy of the sonic boom and the emitted light is in the form of a cone. The cone half-angle $\theta$ depends only on the particle velocity, $v = \beta c$, and the refractive index of the material, $n$:

$$\cos \theta = \frac{1}{n \beta}$$  \hspace{1cm} (3.2.1)

A combination of spherical and flat mirrors are used to reflect light from Cherenkov radiation cones to arrays of photo detectors. The ring diameter determines the Cherenkov angle ($\theta$), and hence the value of $\beta$ for each charged track. The layout of RICH1 is shown in Figure 3.17.

![Figure 3.17: The RICH1 detector seen in the vertical plane, from [46].](image-url)
3.2. THE LHCB EXPERIMENT

RICH1 is designed to identify particles with momenta between 1 GeV/c and 60 GeV/c. Two different radiators are used to achieve this. The first, aerogel, has a refractive index of \( n \approx 1.03 \) and can separate pions and kaons in the range of 1–15 GeV/c. Higher-momentum particles are identified using a C\(_4\)F\(_{10}\) radiator volume. This gas has a refractive index of \( n \approx 1.0014 \). The Cherenkov angles for all tracks in simulated \( B^0 \rightarrow \pi^+\pi^- \) events are pictured in Figure 3.18.

![Figure 3.18: Cherenkov polar angle \( \theta \) versus particle momentum in the RICH system for all tracks in simulated \( B^0 \rightarrow \pi^+\pi^- \) events, from [60].](image)

RICH2 is designed similarly, but performs PID in the momentum range 50–150 GeV/c. The radiator used in RICH2 is CF\(_4\) gas, which has a refractive index of \( n \approx 1.005 \). While RICH1 covers the full LHCb angular acceptance, the angular acceptance of RICH2 is slightly smaller and covers the range from approximately \( \pm 15 \) mrad to \( \pm 120 \) mrad.

The expected number of photons from a typical charged track (\( \beta = 1 \)) is approximately 16 for the C\(_4\)F\(_{10}\) radiator, 5 for Aerogel and 14 for CF\(_4\) gas. To measure the Cherenkov rings, the RICH system requires photodetectors which are sensitive to single photons and cover an area of approximately 3.3 m\(^2\). It was also a design requirement to have a high active-to-total area ratio. The detectors chosen for this task are Hybrid Photon Detectors (HPDs). They are able to detect wavelengths in the range 200–600 nm and are shielded so they can operate in magnetic fields as strong as 50 mT.

An illustration of an HPD is shown in Figure 3.19. Photoelectrons from a cathode are accelerated by a voltage in the range 10 to 20 kV onto a silicon detector. The pixel size at the HPD entrance window is \( 2.5 \times 2.5 \) mm\(^2\) and allows LHCb to achieve a Cherenkov angle resolution of 1.6 mrad.

After the two RICH detectors, the next sub-detectors providing particle ID are the calorimeters: the electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL).
CHAPTER 3. THE LHCB EXPERIMENT

Figure 3.19: A schematic (left) and a photograph (right) of a pixel HPD, from [42].

The calorimeters are comprised of cells which increase in size as distance from the beam axis increases.

The calorimeters perform several functions in LHCb. These include PID and identifying neutral particles such as photons and neutral pions. They also play an important role in the LHCb Level 0 trigger and are able to measure the energy of electrons, photons and hadrons. Some important statistics for the calorimeters are given in Table 3.2.

Table 3.2: Some important requirements of the three calorimeter sub-detectors (after [61]).

<table>
<thead>
<tr>
<th></th>
<th>SPS/PS</th>
<th>ECAL</th>
<th>HCAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of channels</td>
<td>2 × 5952</td>
<td>5952</td>
<td>1468</td>
</tr>
<tr>
<td>Lateral dimensions in x, y</td>
<td>6.2 m × 7.6 m</td>
<td>6.3 m × 7.8 m</td>
<td>6.8 m × 8.4 m</td>
</tr>
<tr>
<td>Depth in z</td>
<td>108 mm</td>
<td>835 mm</td>
<td>1655 mm</td>
</tr>
<tr>
<td>Layers (scintillator + absorber)</td>
<td>2+1</td>
<td>67+66</td>
<td>3+3</td>
</tr>
</tbody>
</table>

Closest to the interaction point are the Scintillator Pad Detector/Pre-Shower system (SPD/PS). This detector is made of two layers of scintillator, separated by a 15 mm lead plate. It works in conjunction with the ECAL to reject events with high $p_T$ neutral pions (which provide a large background) and to determine whether electro-magnetic showers occur before the ECAL. The SPD/PS contributes $0.1\lambda_I$ (where $\lambda_I$ is the nuclear interaction length) to the calorimeter’s material budget.

The ECAL consists of alternating layers of 2 mm thick lead, 120 $\mu$m thick Tyvek® (a synthetic paper-like material) and 4 mm thick scintillator tiles. Readouts of the 5952 channels are made by transmitting the scintillator light to photo-multiplier tubes through wavelength-shifting fibres. A view of the ECAL installed in the LHCb pit and the three sizes of ECAL modules are shown in Figure 3.20. The total length of the ECAL is equivalent to $25X_0$ (where $X_0$ is the characteristic radiation length) which ensures containment of electro-magnetic showers. For the materials used in the ECAL, this is equivalent to 1.1 nuclear interaction lengths.

The HCAL is designed similarly, but the modules are composed of alternating layers of 4 mm scintillators and 12 mm of iron. The total thickness of the HCAL is equivalent to 5.6 nuclear interaction lengths, which brings the total thickness of the calorimeter system
3.2. THE LHCB EXPERIMENT

Figure 3.20: A view of the ECAL (left) installed in the LHCb pit (it is not completely closed) and the three types of ECAL modules (right), from [42].

to $6.8 \lambda_I$. The total number of channels output from the HCAL is 1468.

The part of the LHCb detector furthest from the IP is the muon detection system, comprised of five detector stations (M1–M5 in Figure 3.4). Thick (80 cm) iron absorbers sit between the detector stations M2–M5 to only allow through highly-penetrating muons. The minimum momentum required for a muon to traverse the entire muon detection system is approximately 6 GeV/$c$. A schematic diagram of the muon system, showing its position relative to the calorimeters is shown in Figure 3.21.

The stations are segmented into areas of different resolution depending on proximity to the beam (and hence particle flux). The muon detection system has a (comparatively) large active area of 435 m$^2$, with 1380 detection chambers, and utilizes two detection technologies: Gas Electron Multipliers (GEMs) are used in the inner-most region of M1 and Multiple Wire Proportional Chambers (MWPCs) in all other regions.

Each consists of series of chambers of gas held inside a potential difference of several kV. As a charge passes through each, the gas becomes ionized and the resulting charge avalanche can be read out as a signal.

Efficient detection of muons is important for many LHCb analyses including the benchmark decays $B^0_s \rightarrow \mu^+\mu^-$ and $B^0_s \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$. It is also critical for some lepton flavour violating $\tau$ modes such as $\tau^- \rightarrow \mu^-\mu^+\mu^-$. For decays such as these it is critical to understand the muon efficiency and the mis-identification rates, particularly between $\mu$ and $\pi$. For this it will be necessary to combine information from all sub-detectors in the LHCb particle ID system.

3.2.5 The Trigger

The trigger is a part of LHCB as critical to data collection as any of the physical sub-detectors. The designed beam crossing rate inside LHCb is 40 MHz, with about a quarter of this rate visible by LHCb (it is necessary for protons to collide and for at least one track to be in the LHCb acceptance). To reduce this rate and keep only events of interest, LHCb
uses a two-tiered trigger system. The first tier, the Level-0 trigger, functions at a hardware level and analyses events at the LHC clock frequency of 40 MHz. Events which pass the Level-0 trigger are passed to the High Level Trigger (HLT), which further filters the data.

The Level-0 trigger is illustrated in Figure 3.22. It is designed to reduce the data rate to approximately 1 MHz, at which rate the entire detector can be read out for analysis by the HLT. The Level-0 performs a fast reconstruction of clusters in the calorimeters and in each event searches for signals meeting one of several criteria. It searches for clusters with high transverse energy ($E_T$) in the calorimeters, or high $p_T$ tracks in the muon detector. The pile-up counter is also used to veto events with very high numbers of primary vertices.

The HLT is a series of C++ algorithms which run on a dedicated event filter farm. It is divided into two levels, HLT1 and HLT2. The first task of the HLT is to quickly filter the incoming 1 MHz stream by refining the Level-0 selection (this is known as Level-0 confirmation). The remainder of HLT1 consists of several alleys as is illustrated in Figure 3.23. Each alley corresponds to one of the trigger lines from Level-0, and each partially-reconstructed event is compared to a set of specific selections. HLT1 is designed
3.2. THE LHCB EXPERIMENT

Figure 3.22: An overview of the Level-0 trigger showing the number of channels involved. Information from the pile-up sensors, calorimeters and the muon system are analysed at the LHC clock rate, from [42].

to output at a rate of 30 kHz. Finally, HLT2 applies a set of exclusive and inclusive selections to the events with a final output rate of 2 kHz. This final data stream is written to disk.

Figure 3.23: A flow-diagram of the LHCb trigger, from [42].

Since the HLT is based on C++ code, it is highly configurable and can be adjusted to efficiently manage data rates.

Section 6.1 describes the HLT2 line developed for the mode $D^0 \rightarrow h^- \mu^+ \nu_\mu$ (for $h = \pi$ or $K$).
3.2.6 The LHCb Software Environment

Later sections of this thesis mention the LHCb software environment, so it will be useful to briefly describe this here. The framework underlying the majority of the LHCb software is called Gaudi [63]. Software is kept and developed in packages and projects which are managed with the Apache Subversion revision control system (SVN)\(^1\).

Each step of the LHCb software chain, from the trigger to offline analysis, has a separate project. Some of the key components are:

- Gauss – which manages the generation of MC events. The initial event generation and the simulation of interactions with the detector are handled separately. A variety of external libraries are used in the *Generator Phase*, for instance proton-proton collisions are simulated with PYTHIA [64] and the decays of \(B\) mesons with EvtGen [65]. The *Simulation Phase* is based on GEANT 4 [66].

- Moore – the trigger. The packages relating to all aspects of the trigger (including L0) are included here.

- Boole – which performs digitization. The simulated detector interactions from Gauss are converted into digitized signals to emulate the response to real particles.

- Brunel – which performs reconstruction, treating MC and data events identically. Brunel performs pattern recognition to reconstruct the tracks of charged particles and energy clusters deposited by charged and neutral particles. Additionally, some preliminary particle identification is performed.

- DaVinci – which is used for physics analysis. Analysis tasks in DaVinci accept fully reconstructed tracks and energy clusters as input and the software allows users to manipulate the data and save outputs for further offline analysis.

Projects also exist for the monitoring and commissioning of subdetectors. For example, the silicon tracker and VELO use a project called Vetra.

Each project is composed of several packages. These are smaller sets of related software which performs a similar set of tasks. The packages are contained within *package groups*. For example, the code which defines the modes used in MC generation are kept in Gen/DecFiles, where Gen (for *Generation*) is the package group, containing the DecFiles (for *Decay Files*) package. This package, along with all others relating to simulation and MC generation is associated with the Gauss project.

In the following chapters, developments to VELO software will be discussed. A dedicated project group for VELO software exists, called simply Velo, which is associated with the Vetra project.

\(^1\)http://subversion.apache.org/
4.1 Introduction

This chapter and the following one concern studies of VELO performance. The basis for the present chapter is the LHCb note LHCb-INT-2010-032, which was released internally to the LHCb collaboration. This chapter features several changes from the note (for instance updated plots), but the content is mostly unchanged.

Each VELO silicon sensor has 2048 strips, which are read out through 16 Beetle chips [53]. Each Beetle chip utilizes four analogue links and thus each link outputs 32 analogue channels of data.

The channel data are preceded by four bits of header information and the links are output when a particular event is accepted by the L0 trigger. The headers encode the pipeline number of the event\(^1\) and are digital bits which are subsequently sent over an analogue link.

In this chapter, by “gain” we mean the conversion factor that connects a certain charge input to a silicon strip and the signal output by the Beetle link. When header bits are output, since a known digital signal has been sent through an analogue link, they can be used to monitor and calibrate the gain of each link.

All signals from test pulses or energy deposition by particles are affected by channel-dependent capacitance. Each VELO strip has a different capacitance due to properties including its size and state of depletion. Each link thus has different capacitance properties depending on its strip capacitances and properties of the Beetle pipeline. Since headers are added to the link output in the Beetle, they are not affected by strip capacitance or by non-uniformities in the Beetle pipeline.

First the three calibration methods are discussed. The algorithm to evaluate the gain of each link through the Full Header Swing (FHS) is described in Section 4.2. Section 4.3 describes the test pulse (TP) method, where instead of a calibration based on

\(^1\)This is the position of the event in the internal Beetle pipeline where the data awaits the L0 decision.
header heights the gain settings are chosen to equalise the height of test pulses from each link. Finally, Section 4.4 describes a method based on data from collisions. This method compares the Most Probable Values (MPVs) of link-by-link Landau distributions.

The aim of these methods is to normalise the gain link-by-link, so they do not correct for differences between channels in the same link. Variations in gain across the VELO are affected by several hardware properties including the strip capacitance and differences in each channel’s pre-amplifier circuit.

Sections 4.5 and 4.6 describe the procedure of finding gain settings needed to normalise the gain throughout the VELO under the three methods. As the VELO ages, we expect the gain will need to be recalibrated occasionally and the monitoring software is described in Section 4.7. We recommend using the FHS method to calibrate the gain because it is independent of strip capacitance, pipeline non-uniformity and many of the effects of radiation damage. It may be necessary to re-evaluate this decision on a year-by-year basis.

### 4.2 Full Header Swing

A header is classified either as Header High (HH) or Header Low (HL) depending on whether it corresponds to a ‘0’ or ‘1’ bit. Once Beetle outputs are output through TELL1 boards, the signal units are Analogue-to-Digital Converter (ADC) counts. In the VELO, one ADC count is approximately equivalent to a charge of 440 electrons.

When output in raw, Non Zero Suppressed (NZS) data, the headers sit on top of a pedestal. The pedestals in the VELO system are set to be approximately 512 ADC counts and more details about VELO pedestals are available elsewhere [67]. Typical HH and HL values are approximately 560 and 450 ADC counts respectively. Figure 4.1 shows some typical header distributions for one link. Shown on the left is the header distributions over the full range of physical ADC values, and two narrow distributions can be seen. These two distributions are then magnified separately on the right of the figure.

The quantity chosen to measure the gain of each link is the Full Header Swing or FHS. This is defined for each link as the difference between the means of the HH and HL distributions and for the uncalibrated VELO has a value between 80 and 120 ADC counts. Figure 4.2 shows the FHS distributions for R and Phi VELO sensors for a poorly calibrated VELO before a gain calibration.

The FHS is a fairly robust quantity which is independent of many variables (such as pedestal values) which differ between the links. If \( \langle HH \rangle \) and \( \langle HL \rangle \) represent the means of the HH and HL distributions, then we can write

\[
FHS \equiv \langle HH \rangle - \langle HL \rangle .
\]
4.3 Test Pulse Normalisation

A second method for monitoring and calibrating the gain has been proposed which substitutes the header heights for test pulses (TP).

The Beetle chips have a circuit able to inject a specified charge (i.e. a test pulse) into the individual channels. Similarly to headers, the test pulses have a known input value...
and thus offer an alternative method to monitor the gain. After a gain calibration using the FHS method, the R and Phi test pulse distributions (shown in Figure 4.3) have means of 32.4 and 33.3 ADC counts respectively. The RMS values of the distributions are 1.4 and 1.7 ADC counts.

By following a similar method to the gain calibration by FHS values, it is possible to calibrate the VELO gain by the test pulse heights and the two methods are correlated since the outputs are affected by the same variations in strip capacitance. The FHS method was chosen to monitor and calibrate the gain for the years 2009 and 2010, but test pulses offer an alternative, though not completely independent method. For example, unlike header heights, the test pulse output from each strip is dependent on strip capacitance and
4.4 MPV Normalisation

Before LHCb had recorded many events from collisions, FHS and test pulse normalisation were the only two reasonable methods for gain normalisation. Since 2010, LHCb has recorded enough events that a third method can be used for monitoring and calibrating the VELO gain through link-by-link Landau distributions.

A normalisation based on the response to real tracks could be considered more ‘physical’ than the two methods described above which rely on electronic output from the VELO itself. This method measures the most probable value (MPV) of the Landau distributions for each link and normalises the links to the same MPV\(^1\). An attempt is made to account for track-angles affecting the Landau distribution. Tracks which intersect the silicon sensors at high angles pass through more silicon as they traverse the sensor and hence leave more energy on average than tracks travelling perpendicular to the sensors. To correct for this effect, the final ADC count from each cluster is multiplied by the cosine of the track’s angle $\theta$.

\(^1\)This is done by using the gain scan curves described in Section 4.5.

Figure 4.3: The test pulse distributions for the VELO calibrated with the FHS method represented as a 2D plot (left) and histograms (right) for R (top) and Phi (bottom) sensors. ‘Unweighted Summary’ means that the plot has not been re-weighted by the noise of each link.
A typical Landau distribution (for VELO sensor 48, link 41) is pictured in Figure 4.4. A fit (red curve) of the Landau distribution convoluted with a Gaussian is made to the data to provide the MPV and full width at half maximum (FWHM) estimates. In this chapter, for brevity we refer to the fit of a Landau distribution convoluted with a Gaussian simply as a Landau fit.

If the VELO is calibrated using the Landau distributions, we would expect the MPV and FWHM estimates to be the same between links. For the current method, we take all VeloClusters associated to tracks in collisions data. These are then associated to a particular link and the fitting algorithm can be applied to the resulting histograms.

The occupancy of each link depends very strongly on its position relative to the interaction point. In practice to collect at least $2 \times 10^4$ hits in the lowest occupancy links requires around $10^8$ tracks with VELO hits from the minimum bias stream. Minimum bias is the name given events which are selected with as little trigger bias as possible. Since minimal trigger decisions are made to filter the more “interesting” physics signal events, this stream (called MINBIAS) is essentially a sample of all events that occur within the LHCb detector. However, some bias is practically unavoidable. For instance, it is necessary for the LHCb trigger to be fired, and this does not occur for events when all products remain in the beam pipe and do not interact with the detector.

Several classes of track exist in LHCb [71] and an illustration of the different types is
shown in Figure 4.5. The main difference between the types relates to which sub-detectors the track traversed. The plots in this section are based on a sample of 97 million long, VELO and upstream tracks recorded in August 2010 from the minimum bias stream.

![Figure 4.5: The track definitions used in LHCb.](image)

The number of hits in each link is shown in Figure 4.6, where effects due to VELO geometry are visible. Sensor are numbered sequentially in a downstream direction – the most upstream $R$ sensor is number 0, while the furthest downstream is number 41. For $\Phi$ sensors, these are sensors 64 and 105.

Particles which leave the IP and traverse a sensor very close to the IP travel at high angles, so have much higher probability of not leaving sufficiently many hits for a track to be reconstructed. Thus, in Figure 4.6 the sensors with higher numbers typically have a greater number of hits associated with tracks, and sensors upstream of the IP have very few hits. There is an order-of-magnitude difference between the highest and lowest occupancy links.

To verify the precision of the fitted Landau curves, we can compare the MPV to the largest bin in the Landau distribution, which for sufficiently high statistics will be within one ADC count. The difference between the maximum bin and the fitted MPV is shown in Figure 4.7. The majority of links are within one ADC count and by comparing Figure 4.6 and Figure 4.7 evidently the distribution does not strongly depend on statistics.

These data were taken with a VELO calibrated using the FHS method – each link had a FHS value calibrated to 100 ADC counts (more details on how the calibration is performed are given in the following sections). The distributions of fitted Landau MPV and the full width at half maximum (FWHM) of the fit are shown in Figure 4.8 and Figure 4.9. There is some variation in MPVs, and in particular the $R$ sensors (numbers 0–41) have a lower mean MPV than the $\Phi$ sensors. This effect is largely due to higher strip capacitance in $R$ sensors.
Figure 4.6: Summary of the number of hits in each link for the minimum bias sample.

Figure 4.7: Difference between the Landau maximum bin and the fitted MPV.
4.4. MPV NORMALISATION

Figure 4.8: Distribution of fitted MPVs.

Figure 4.9: Distribution of FWHM of fit for each link.
The distribution of MPVs for $R$ and $\Phi$ sensors are summarised in Figure 4.10 and Figure 4.11. The data were taken in late 2010 when we can assume radiation damage was minimal, with a VELO calibrated under the FHS method. The means (RMS) of the two distributions are 37.1 (1.2) and 38.8 (1.2) and enable an estimate of the FHS precision – the RMS divided by the mean in each case is approximately 3%.

![Figure 4.10: Distribution of MPVs for $R$ sensors for a VELO calibrated under the FHS method.](image)

We expect an agreement between the FHS and MPV methods of approximately 3% to be sufficient for physics analysis and any potential $dE/dx$ measurements from the VELO.

### 4.5 Gain Scan and Gain Parameters

The algorithm discussed in Section 4.2 allows users to monitor the VELO gain, but GainMon also allows users to run over Gain Scan (GS) data to determine gain curves. These curves give the relationships between an input gain setting on the TELL1 card and the resulting FHS (or test pulse) value. The form of the gain curve for each link enables the optimum gain settings to be determined for the FHS and TP methods.

A Gain Scan is a PVSS recipe which takes NZS data (without beams), with incremented, decreasing gain settings. This recipe takes 2000 events at each of 23 steps in gain. The steps are applied to the DAC_DATA setting on the TELL1 ARX card [54, 72]. This setting controls the upper limit of the voltage range digitized by the ADC and hence the proportionality of an input voltage to ADC counts.
4.5. GAIN SCAN AND GAIN PARAMETERS

The gain is altered by a simple linear function. If we let \( s \) represent the step number, then the gain setting \( G(s) \) is given by:

\[
G(s) = 61376 + 245(11 - s),
\]

which increments the gain factor from 104% to 95% around the default DAC_DATA setting of 61376 (or EFC0 in hexadecimal – the first two bits are used for configuration). The GS data for step 11, corresponding to a \( G(s) \) of 100%, is equivalent to no change. The GS algorithm sets the gain setting of each link to this value, which causes a shift in the HH and HL distributions and hence also in the FHS distributions.

Software in the Velo/VetraScripts package can then be used to analyse the resulting ROOT file to find the necessary value of \( s \) to normalize each link’s FHS to a constant value. The chosen value is a somewhat arbitrary choice, but at present it is set to be 100 ADC counts. Setting his value too low means a loss in signal-to-noise, while too high reduces the dynamic range of the VELO. A typical sensor response to such a gain scan is shown in Figure 4.12. The decrease with step number is not linear, so we fit an exponential curve of the form:

\[
FHS = c_1 e^{c_2 x} + c_3,
\]

where \( x \) represents the gain setting. An example fit for link 63 is superimposed on the
data.

![Figure 4.12: FHS response to a Gain Scan for one sensor (64 links). To demonstrate the exponential fit, the best-fit curve for link 63 is superimposed on the bottom right plot.](image)

The gain settings are stored in XML files in the VELOCOND database and some related technical details are given in appendix A. The python script used to create the XML files also performs some re-ordering tasks. It uses the TELL1Map module to map sensor to TELL1 numbers from the VELO TELL1 database. It also maps link numbers associated with a sensor to the correct link within the TELL1 – the $i^{th}$ link associated to a sensor is mapped to a TELL1 link number via:

$$\text{link} = 16(3 - \lfloor i/16 \rfloor) + 8(1 - \lfloor (i \mod 16)/8 \rfloor) + i \mod 8.$$  

(4.5.3)

### 4.6 Gain Parameters From the MPV Method

It is relatively straight-forward to use the link-by-link MPV information to recalibrate the VELO gain, using the code described in the previous section.

We assume that the effect of the gain setting on FHS and MPV are the same (i.e. if a gain change increases the FHS by $N\%$, then the MPV will also increase by $N\%$). Under this assumption, the gain curves described in Sec. 4.5 can also be used to calibrate the links to a chosen Landau MPV.

1. 

http://hep.ph.liv.ac.uk/velodb/VELO/
If the current MPV value is $M$, then the gain setting needed to normalise the MPV to the value $M'$ is given by

$$x' = \frac{1}{c_2} \ln \left[ \frac{1}{c_1} \left( \frac{M'}{M} \cdot 100 - c_3 \right) \right].$$ (4.6.1)

The number 100 appears in the above equation because this was the FHS normalisation value used during data taking.

Once this setting is calculated for each link, the values can be uploaded as XML files following the procedure described in Sec. 4.5.

### 4.7 Performance

The software written to monitor the performance of gain is called `drawFHSPlots` and is available in the `Velo/VetraScripts` package. It can be run either through the ROOT interpreter, or in the VELO GUI. This code produces several plots which shifters and experts can examine to monitor the VELO gain.

Of the three methods discussed above, the FHS and TP methods are the easiest to monitor for shifters in the pit. Only a small sample of $10^4$ NZS events are required to produce the monitoring histograms, and this data can be taken during any period without LHC beams.

The gain of each link monitored through the link FHS depends strongly on the VELO timing. We measure the spread of link FHS values by the RMS of the R and Phi FHS histograms, shown in Figure 4.13. When the timing is precise, a gain calibration can achieve FHS values around the desired mean (100 ADC counts) with an RMS of around 0.4. Under this calibration, the corresponding spread of the MPV distributions is approximately 3%, which we expect is sufficiently precise for physics analysis and any potential $dE/dx$ developments.

One of the most important tasks of monitoring the VELO gain is to determine how the link FHS values vary over time. As the VELO ages, it is possible that the FHS distributions will change, and a recalibration may be required at certain times. However, changes caused by radiation damage will occur mainly in the sensors and not the Beetle chips, so will be less severe than if the VELO were normalised under the MPV method. It is not yet known how often a recalibration will be required, although it is expected to be of order 3–6 months.

The FHS method appears to provide stability for months at a time and has been used successfully to monitor and calibrate the gain. When the VELO has minimal radiation damage, the MPV method provides an alternative, but the VELO group plans to use the FHS method for the long-term because it is much less sensitive to the effects of radiation damage.
CHAPTER 4. VELO GAIN CALIBRATION AND MONITORING

Figure 4.13: FHS distributions after a gain calibration represented as a 2D plot (left) and histograms (right) for R (top) and Phi (bottom) sensors. An uncalibrated distribution (red) is overlaid for comparison.

Additionally, the speed of performing FHS and TP analyses means these methods are useful for monitoring in the pit. Compared to the relatively small number of NZS events required under the FHS or TP methods, the MPV method requires at least 100 million minimum bias tracks to accumulate sufficient data. This task is performed by a dedicated TupleTool and two analysis scripts in Velo/VetraScripts.

4.8 Summary

Three methods for monitoring the VELO gain were discussed. Two of these (FHS and TP methods) allow for a calibration to be performed by running a gain scan in the absence of beams. The third method of monitoring the MPV for link-by-link Landau distributions uses the gain scan curves and collision data to normalise the MPVs of each link to the same value. To ensure each link in the VELO has a sufficient number of hits, approximately $10^8$ tracks are required to perform this task.

The FHS-based gain calibration and parameter uploading have been tested and have so far provided stability over several months and the recommendation is to use this method for the long-term calibration and monitoring of the VELO gain.
CHAPTER
FIVE

PROSPECTS FOR PARTICLE IDENTIFICATION WITH
THE VELO

5.1 Introduction

The three other large LHC experiments (ALICE, ATLAS and CMS) have studied particle identification (PID) through energy loss (see for example [73]) and have presented results publicly. This type of PID is common in gas or in silicon (for a summary of the latter see [74]). In this chapter we investigate the possibility of using this method in LHCb.

Measurements of a particle’s momentum and the energy lost per unit of material traversed \(\left(\frac{dE}{dx}\right)\) have been used in many experiments to identify particle species in the low momentum region below 2 GeV/c. Here, \(x\) is usually measured in units of density per unit length (for example g/cm\(^2\)). Primarily the technique can be used to discriminate between charged hadron species and to a lesser extent leptons. There is also the possibility of detecting long-lived massive charged particles, which are predicted by some New Physics models.

If \(\frac{dE}{dx}\) identification could be achieved it would complement the existing PID provided by the RICH detectors (see [42] for details). For example, RICH1 provides PID down to approximately 1 GeV, and this is where contributions from \(\frac{dE}{dx}\) would provide the best discrimination.

The LHCb VELO (as described in [50]) is composed of two halves of double-sided, semi-circular silicon strip detectors. Each half consists of 21 stations of back-to-back \(R\)- and \(\Phi\)-measuring sensors. On average, long tracks leave clusters in 12 of these sensors, each of which provides a measurement of energy loss. The distribution of energy deposited by charged particles per unit length in thin silicon typically follows the so-called Landau distribution.

In Section 5.2 we describe the Landau distribution and eight estimates used to characterise it. In the case of low statistics, it is preferable to use an arithmetical quantity to
approximate the most probable value (MPV) of the energy deposition distribution rather than fitting a curve. Section 5.3 describes the track selection for this analysis.

In Section 5.4 the analysis of a sample of minimum bias Monte Carlo (MC) is described. We find the distributions using MC truth and reconstructed momentum values. A brief comparison to collision data is made in Section 5.5 and finally two other approaches tried during this study are described in Section 5.6.

5.2 Landau Distributions

The silicon in each VELO sensor is 300 $\mu$m thick. When charged particles pass through a thin layer of silicon, the distribution of measurements of $dE/dx$ over small ranges of $x$ typically follows the Landau distribution. Detailed summaries of this process are available elsewhere [75].

The Landau distribution $f(x)$ has the following form:

$$f(x) = \frac{1}{\pi} \int_0^\infty \exp(-t \log t - xt) \sin(\pi t) dt,$$

where $t$ is simply an integration variable. This distribution is known for some unusual properties: a long high-end tail and an undefined mean. It is commonly characterized by its width (the full-width at half-maximum or FWHM) and most probable value (MPV).

In the VELO, instead of measuring the deposited energy directly, we measure Analog-to-Digital Converter (ADC) counts. One ADC count corresponds to a charge of approximately 380 electrons collected in the silicon.

In practice, the distributions of clusters seen in the VELO are closer to a Landau distribution smeared with a Gaussian. In the case of high statistics (such as the distribution of all clusters in a sensor) the MPV and FWHM can be found by fitting such a function over a chosen range. A typical distribution was shown in Figure 4.4 and is reproduced here in Figure 5.1.

The process of fitting a full Landau function can give excellent results for distributions with high-statistics in cases where the amount of CPU time consumed is not particularly important. When PID is performed on a track-by-track level, statistics are low (approximately ten clusters per track), and with many hundreds of tracks per event the potential CPU consumption would be prohibitively high. For these reasons using a full Landau fit is inappropriate for this application.

To study $dE/dx$ in the VELO, we have used eight arithmetic quantities to characterise each Landau distribution. All approximate the MPV without requiring a full Landau fit. The eight quantities are the standard arithmetic mean and median, two truncated means and four generalised means.

The median is simply the middle value after the hits have been sorted by size. If there
5.2. LANDAU DISTRIBUTIONS

Figure 5.1: A reproduction of Figure 4.4: a typical Landau distribution for VELO sensor 38, link 41 for clusters on tracks. The characteristic high-end tail is clearly visible.

is an even number of hits, the median is taken as the average between the two central values.

Truncated means lessen the effect of high outliers but at the cost of losing statistics. To calculate a truncated mean, the highest $N\%$ of values (rounded to the nearest integer number) are discarded then a standard arithmetic mean is taken of the remaining values. For the two truncated means used in this study, the top 20% and 40% of hits are discarded.

A generalised mean of grade $k$ of a set of positive real numbers $x_1, \ldots, x_n$ is given by:

$$M_k(x_1, \ldots, x_n) = \left(\frac{1}{n} \cdot \sum_{i=1}^{n} x_i^k \right)^{1/k}.$$  \hfill (5.2.2)

The four generalised means studied here have grades $k \in \{-0.5, -2, -4, -6\}$. Note that the case of $k = -1$ corresponds to the harmonic mean. Generalised means can (especially for large, negative values of $k$) reduce the effect of high outliers, although it is at the expense of introducing some dependence on the arbitrary parameter $k$.

The properties of the Landau distribution (in particular the undefined mean) indicate that the mean and median of samples from a Landau distribution will provide poor estimates for the MPV. The most common estimate in past experiments has been some variety of truncated mean.
5.3 Track Selection

Based on the LHCb tracking definitions [71], we study two classes of track: long and upstream. These two meet the criteria of leaving hits in the VELO while also travelling through the magnet and thus allowing a momentum measurement. For clarity, the track definitions shown in Figure 4.5 are reproduced in Figure 5.2.

![Track Definitions](image)

Figure 5.2: The track definitions used in LHCb.

Due to the geometry, the two track types have a different mean number of associated clusters, with upstream tracks having slightly fewer. For tracks with momentum below 5 GeV/c the average number of clusters is 12.3 for long tracks and 11.0 for upstream tracks. The distributions are shown in Figure 5.3.

For this study we are interested in all tracks, so we make very few physics cuts. The tracks we examine have the following requirements:

- track is long or upstream
- momentum (p) is below 5 GeV/c
- at least 10 associated VELO clusters
- the PID DLL(e − π) is below zero.

By DLL(e − π) we mean the difference between the combined delta log likelihood values for the electron and pion hypotheses. This requirement is made to remove some of the electrons in a fairly crude way, and results in a sample with the most ‘obvious’ (or high DLL value) electrons removed. It will be shown in the following chapters the electron distributions from MC have a different shape to the other particle species. The comparatively large numbers of electrons suggests that removing some through this method is useful.
5.4 Monte Carlo Results

First, we examine the eight MPV estimates with the true MC momentum and particle ID from a sample of long, upstream and downstream tracks from approximately one million 3.5 TeV minimum bias MC events. The number of each species found in this sample which pass the selection requirements are given in Table 5.1. We study five species of particle: proton (red), electron (black), pion (grey), kaon (blue) and muon (green). The distributions for the eight estimates (using both long and upstream tracks) are shown in Figure 5.4 and Figure 5.5.

As expected, the mean and median distributions look much poorer (i.e. wider) than many of the others. For demonstration purposes, we examine more closely the distribution of 40% truncation. The distributions for each particle type are separated in Figure 5.6. The shapes of the proton and kaon distributions are especially distinctive below 2 GeV/c.

The large number of points in these scatter plots can make them difficult to analyse and the information is often better presented in a profile plot, such as Figure 5.7. A profile plot enables a two-dimensional histogram or scatter plot (such as Figure 5.6) to be presented in a form resembling a one-dimensional histogram. The profile of a scatter plot in the $xy$ plane can be obtained by plotting the $y$ values against the $x$ values for each point in the scatter plot.
CHAPTER 5. PROSPECTS FOR PARTICLE IDENTIFICATION WITH THE VELO

Figure 5.4: MC distributions for the four generalised means as a function of the true momentum for the five species of particle: proton (red), electron (black), pion (grey), kaon (blue) and muon (green).

Figure 5.5: MC distributions for mean, median and the two truncated means as a function of the true momentum for the five species of particle (colours as above).
Figure 5.6: The 40% truncation distribution for each particle species separately. The particles (top-to-bottom); left column: pion, kaon, proton; right column: electron, muon.

plane is found by calculating the mean and RMS of the $y$ values in each $x$ bin. The profile plot then displays the associated mean $y$ value in each $x$ bin, with the RMS associated with each point presented as error bars.

This plot shows how clearly the protons (and to a lesser extend kaons) are separated from the other species in the low momentum range.

Even with the simple PID cut to remove electrons, those which remain have a much wider distribution than the other species. The numbers of each species found in this sample which pass the selection requirements are given in Table 5.1.

<table>
<thead>
<tr>
<th>Species</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron</td>
<td>43530</td>
<td>0.3</td>
</tr>
<tr>
<td>Pion</td>
<td>2927710</td>
<td>18.3</td>
</tr>
<tr>
<td>Proton</td>
<td>155359</td>
<td>1.0</td>
</tr>
<tr>
<td>Kaon</td>
<td>288744</td>
<td>1.8</td>
</tr>
<tr>
<td>Muon</td>
<td>13920</td>
<td>0.1</td>
</tr>
</tbody>
</table>

The plots provide a proof-of-concept that several of the trialled MPV estimates could enable the VELO to provide some PID information to LHCb. It is entirely dependent on the precision of momentum measurements for this particular class of track – i.e. very low momentum long or upstream tracks which have a large (> 10) number of associated VELO clusters.

The geometry of LHCb causes a minimum momentum cutoff for long tracks. Particles with very low momentum are swept out of the acceptance by the magnet. For tracks
with momentum below approximately 1.3 GeV/c, upstream tracks are all that remain. However, the momentum reconstruction of such tracks is difficult.

We can plot the reconstructed momentum for these two classes of tracks against the true MC momentum and this is shown as a profile plot in Figure 5.8. The data points are plotted against a reference line of slope 1 – i.e. if momentum reconstruction was perfect, each point would fall on this line.

Upstream tracks are those which satisfy solely the VELO and Tracker Turicensis (TT) acceptance (and hence not the T stations). They thus do not leave hits in any tracking station after the magnet. To precisely measure momentum, LHCb must determine the track’s
curvature through the magnet (i.e. after the VELO and TT). For this reason, though upstream tracks can be found in the useful momentum range, the momentum measurements are not as reliable. Long tracks show excellent agreement between true and reconstructed momentum but there are few below 1.2 GeV/c.

In changing from MC true momentum to MC reconstructed momentum, the particle species become less distinct in each MPV estimate. Figures 5.9 and 5.10 are the scatter and profile plots for the 40% truncated mean using long tracks alone. The same plots for long and upstream tracks combined are shown in Figure 5.11 and Figure 5.12.

The effect of the difference between true and reconstructed momentum can be seen by comparing Figure 5.7 and Figure 5.12. The separation power between species is much less with reconstructed momentum. With the above very simple selection requirements,
which include only one PID requirement, it is difficult to achieve both high purity and efficiency for selecting any particular species.

We select two bins of momentum: the first 0.7–0.9 GeV/c, and the second 0.9–1.1 GeV/c which are both in the region where \( \frac{dE}{dx} \) is expected to give the greatest separation. We simply count the number of MC particles of each species type: electron, muon, pion, kaon, proton and “other”. By other we mean all tracks which have a MC ID different to the other five, which are almost entirely ghost tracks.

Ghost tracks (for more details see [76]) are collections of pseudo-random hits in LHCb which are reconstructed as a track, although they do not originate from a single genuine particle. They can arise through several processes: from random electrical noise in the LHCb detectors, from the mis-reconstruction of genuine physics tracks or through a combination of these.

This class of track is detrimental to physics analysis, so several techniques are used by
the LHCb collaboration to identify potential ghost tracks (necessary when one does not have access to the true MC particle ID). These include associating VELO tracks with hits in the other silicon trackers and using a likelihood method to assign a ghost probability to each track [77].

At the lowest momenta protons have the highest MPV values, so to test the viability of $dE/dx$ particle ID we can attempt to isolate the protons with high MPVs. The distributions of the MPV estimate in each momentum bin are shown in Figure 5.13 and Figure 5.14. At high values of the truncated mean, only electrons and protons (and ghost tracks) remain.

![Figure 5.13: Distribution of the 40% truncated mean estimate in the momentum bin 0.7–0.9 GeV/c. The particle species are proton (red), electron (black), pion (grey), kaon (blue) and muon (green).](image)

![Figure 5.14: Distribution of the 40% truncated mean estimate in the momentum bin 0.9–1.1 GeV/c. The particle species are proton (red), electron (black), pion (grey), kaon (blue) and muon (green).](image)

We impose five different minimum values for the 40% truncated mean (including 0, which corresponds to no cut). The results are given in Table 5.2 and Table 5.3 and are
similar for most of the other MPV estimates.

Table 5.2: Numbers of each species in the momentum bin 0.7–0.9 GeV/c, with a minimum 40% truncated mean value.

<table>
<thead>
<tr>
<th></th>
<th>40% Trunc. Mean Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Electron</td>
<td>8743</td>
</tr>
<tr>
<td>Muon</td>
<td>158</td>
</tr>
<tr>
<td>Pion</td>
<td>44429</td>
</tr>
<tr>
<td>Kaon</td>
<td>1737</td>
</tr>
<tr>
<td>Proton</td>
<td>722</td>
</tr>
<tr>
<td>Other</td>
<td>19060</td>
</tr>
<tr>
<td>Total</td>
<td>74849</td>
</tr>
</tbody>
</table>

Table 5.3: Numbers of each species in the momentum bin 0.9–1.1 GeV/c, with a minimum 40% truncated mean value.

<table>
<thead>
<tr>
<th></th>
<th>40% Trunc. Mean Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Electron</td>
<td>6676</td>
</tr>
<tr>
<td>Muon</td>
<td>294</td>
</tr>
<tr>
<td>Pion</td>
<td>87534</td>
</tr>
<tr>
<td>Kaon</td>
<td>4227</td>
</tr>
<tr>
<td>Proton</td>
<td>1830</td>
</tr>
<tr>
<td>Other</td>
<td>20745</td>
</tr>
<tr>
<td>Total</td>
<td>121306</td>
</tr>
</tbody>
</table>

These numbers indicate the sort of performance one might expect from this method. For example, in the 0.7–0.9 GeV/c momentum bin of this MC sample, particles selected by a minimum truncated mean cut of 90 are 36% electrons and 52% protons. However, achieving this purity of protons has required rejecting 99.6% of all tracks.

Performance is somewhat better for some of the generalised means. For example, the results in the 0.7–0.9 GeV/c bin for the generalised mean with \( k = -6 \) are shown in Table 5.4. In this case, requiring the generalised mean value be above 90 leaves a sample with 13% electrons and 78% protons. This selection rejects 99.7% of all tracks.

It is possible to impose additional requirements on track quality by cutting on the track fit \( \chi^2 / \text{d.o.f.} \). However, since many of the tracks with high \( \chi^2 \) have only a few associated hits, the requirement of at least ten associated VELO clusters removes most of the poor-quality tracks. We find that even requiring \( \chi^2 / \text{d.o.f.} < 1.5 \) does not significantly improve the purity of protons in the above sample.

Clearly, in this form \( dE/dx \) PID is impractical to use, even to select protons. The effect of a large uncertainty in upstream track momentum measurements causes smearing
### 5.5 Data Results

Useful particle ID through $dE/dx$ is clearly problematic even in Monte Carlo and in data we see the same difficulties. We examine a sample of 29 million minimum bias (MB) events of 3.5 TeV collisions. For this type of analysis the particular stream of data used is mostly irrelevant. We select the same tracks described above and compare the MC and data distributions, which are shown in Figure 5.15 and Figure 5.16.

There are several things to note in these plots. In general the distributions are similar, with a couple of important differences. Even with a logarithmic $z$-axis, the area below 1 GeV/$c$ where the proton curve becomes prominent in MC is hardly visible in data.

Secondly, there is an overall normalisation difference in the average MPV value. This is due to a factor input to Brunel which models the proportionality between MC energy deposited in a VELO sensor and the resulting ADC count output. This difference has been corrected in more recent MC releases.

Evidently, without improvements to momentum measurements of very low momentum tracks (which is mostly limited by geometry) and the ability to reject electrons in this range, the VELO cannot provide much additional PID information.

---

#### Table 5.4: Numbers of each species in the momentum bin 0.7–0.9 GeV/$c$, with a minimum $k = -6$ generalised mean value.

<table>
<thead>
<tr>
<th></th>
<th>Gen. Mean $(k = -6)$ minimum</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>Electron</td>
<td>8743</td>
<td>542</td>
</tr>
<tr>
<td>Muon</td>
<td>158</td>
<td>0</td>
</tr>
<tr>
<td>Pion</td>
<td>44429</td>
<td>3</td>
</tr>
<tr>
<td>Kaon</td>
<td>1737</td>
<td>241</td>
</tr>
<tr>
<td>Proton</td>
<td>722</td>
<td>692</td>
</tr>
<tr>
<td>Other</td>
<td>19060</td>
<td>328</td>
</tr>
<tr>
<td>Total</td>
<td>74849</td>
<td>1706</td>
</tr>
</tbody>
</table>
Figure 5.15: Distribution of MC for 40% truncated mean with no requirements on ID to replicate data.

Figure 5.16: Distribution of collision data for 40% truncated mean.
5.6 Other Methods

This section describes a few other methods investigated during this study, in particular using the existing global PID to select tracks and performing the $dE/dx$ analysis on a sample of $\Lambda \rightarrow p\pi$ events.

The global PID cuts are optimised to work above 1 GeV/$c$, but it is still possible to make plots similar to Figure 5.12 but selecting the particle species with the global PID instead. We select tracks with the same requirements as before, but without any requirement on DLL$(e - \pi)$. The very simplistic requirements used are given in Table 5.5. In the table, DLL$(e - \pi)$ is notated by ‘PIDe’ for simplicity, and similarly for the other particle species.

Table 5.5: Identification of particles using global PID.

<table>
<thead>
<tr>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron</td>
</tr>
<tr>
<td>Pion</td>
</tr>
<tr>
<td>Proton</td>
</tr>
<tr>
<td>Kaon</td>
</tr>
<tr>
<td>Muon</td>
</tr>
</tbody>
</table>

These selection requirements are far from optimised, but are designed to test whether this method has potential. The results for long and upstream tracks are shown in Figure 5.17. The most useful comparison is with Figure 5.12 (on page 86). The electron curves are quite similar between the two, down to the cutoff at 1 GeV/$c$. In the global PID plot, all other species lie over each other and are not distinguishable.

The reasons for this are that the global PID is not optimised to work in this region, and in particular for very low momentum upstream tracks.

A second approach studied was to select a sample of pion and proton tracks by searching for the decay $\Lambda \rightarrow p\pi$. Depending on the tightness of selection cuts, this can yield a clean sample of pions and protons. However due to the decay kinematics, it is very rare to find protons from this decay with momentum below 2 GeV/$c$.

The very simple selection criteria were based on the StdLooseProtons and StdLoosePions lists. Two ‘daughter’ cuts were made on the track $\chi^2$/d.o.f. and impact parameter, and four cuts were made on the candidate $\Lambda$ (on the vertex $\chi^2$, DIRA, mass and impact parameter). Running these cuts on a sample of 15 million signal MC events yielded 42298 candidate decays, but no selected proton had a momentum below 2 GeV/$c$. This method appears to be ineffective to isolate a sample of protons.
5.7 Summary

At present, it appears that the VELO cannot offer a contribution to LHCb PID. Several candidate estimates for Landau MPV estimates work successfully with MC Truth momentum, but uncertainty of momentum measurements for the very lowest momentum upstream tracks drastically reduces the effectiveness. With reconstructed momentum, rejection of more than 99% of tracks in a momentum bin is necessary to select a sample of protons as pure as 78%.

At present, gain in the VELO is calibrated based on the heights of header bits in non-zero-suppressed data. The gain is the multiplication factor between the charge deposited in a silicon sensor and the resulting output ADC count. To perform $dE/dx$ with data, it is important that the gains are calibrated to give a uniform response throughout the VELO.

The primary cause of difficulty relates to the geometry of LHCb. This technique of $dE/dx$ PID would work best for particles with momentum below 2 GeV/$c$, but these very low momentum particles are swept out of the LHCb acceptance before interacting with most of the detector and are thus classified as upstream tracks.
TRIGGER AND STRIPPING FOR $D^0 \to K^- \mu^+ \nu_\mu$

The signature of the decay $D^0 \to K^- \mu^+ \nu_\mu$ is a moderate-momentum $K^-\mu$ pair originating from a secondary vertex, with a mass below the $D^0$ mass, $M(D^0)$. This chapter describes methods used to extract this signal from the LHCb data.

The Branching Ratio (BR) for the decay is just over 3%, while the integrated cross-section for $D^0$ production within the LHCb acceptance is approximately 1.5 mb [78]. Thus, at the LHCb nominal integrated yearly luminosity of 1 fb$^{-1}$, approximately $10^{11}$ signal events will be produced in each year of nominal running. For this reason, fairly low efficiencies of signal selection are acceptable provided the resulting sample of events is clean and relatively free from unwanted backgrounds.

Three stages are required to collect a clean sample of signal events: the trigger, stripping and finally offline selection. Offline selection is described in Chapter 7 and the first two steps are described in the present chapter.

LHCb runs an exclusive HLT2 trigger for this signal, which is only run on events passing particular combinations of the lower-level L0 and HLT1 triggers and selects events with an optimal ratio of $S/\sqrt{S+B}$ (where $S$ and $B$ represent the number of signal and background events respectively).

Stripping is the nickname used in LHCb for what is essentially an offline trigger which runs after HLT2 and full data reconstruction. This stage of processing is necessary due to the large amount of data collected by LHCb, since the processing power required for an individual to run on all HLT2 events would be too demanding [79]. Stripping sorts the data into streams of a more manageable size specific to each analysis.

Finally, the offline selection is the final stage which runs on the greatly reduced number of events in the stripped data and selects the events used in the final fitting procedures.
6.1 Trigger

As was described in Section 3.2.5, both hardware-based (Level-0 or L0) and software-based (High Level Trigger or HLT) levels are used in the LHCb trigger [80]. The trigger is a constantly evolving system which must be flexible enough to manage changes in running conditions. During early data-taking this was particularly important, and the increases in beam energy and luminosity as the LHC matured required frequent reassessment of the trigger.

A definite limit for the trigger is imposed by the rate at which LHCb can write events to tape, approximately 3 kHz. The function of the trigger is to reduce the nominal interaction rate from 40 MHz to this figure. Nominally, L0 and HLT1 are designed to reduce the rate by a factor of 40 to approximately 1 MHz and then by a factor of 30 to around 30 kHz. The final reduction to a few kHz is performed by HLT2.

Analysing a dataset where events have been accepted through a large number of different triggers presents complications when assessing systematic errors. For this reason, the exclusive HLT2 line only accepts events which pass a small set of LHCb lower-level triggers. With a final state \(K^-\mu^+\nu\) pair, the choice of which L0 lines to permit is evident, and only events passing L0Hadron or L0Muon triggers are accepted.

The trigger requirements on HLT1 require a Trigger on Signal (TOS). This means that one of the \(D\) daughters must trigger the line in question. The reverse of this situation is called Trigger Independent of Signal (TIS) where an event is selected by a trigger but it is not one of the \(D\) daughters which caused the trigger to fire.

The Monte Carlo (MC) samples used to study and develop the trigger were MC10, Sim-01-Strip12 samples of signal \(D^{*+} \rightarrow \pi^+ D^0(K^-\mu^+\nu)\) and minimum bias (MB) simulated data. Equal amounts of data with the LHCb magnet polarity “up” and “down” were used. In total there were 20.6 million MB events (MC identifier 30000000), and 2.0 million signal events (MC identifier 27173001).\(^1\) All events had the same LHCb reconstruction and trigger software: Gauss v39r0, Boole v21r9, Moore v10r2 and Brunel v37r8p5. The samples all had a particular trigger configuration (TCK) applied: TCK 0x002e002a. It is from this TCK that trigger efficiencies were calculated.

A slight complication is that due to the nature of MB events, some genuine signal was contained in the so-called “background” sample. Monte Carlo truth information was used to filter out such events. Any event referred to as signal in this section was from the signal sample, and also was required to be the genuine event under MC truth, while background events are those from the background sample which fail the MC truth requirement for signal.

\(^1\)The Identifier numbers are those defined in [81].
6.1. Loose Preselection for Trigger

To select a sample for testing purposes, some pre-selection was applied to the two MC samples. A loose set of cuts specifically designed to select prompt (i.e. originating from the primary vertex) $D^*$ tagged $D^0 \to K^- \mu^+ \nu_\mu$ decays was applied to the MB and signal samples. These two samples were used to determine which variables were most effective in separating background and signal. Once a set of effective variables was found, a second, even looser, pre-selection was applied only to the signal MC, to arrive at a set of cut values that minimally biased the signal $q^2$ distribution\(^1\) (the $q^2$ optimisation).

The first pre-selection had requirements on the pointing angle and the impact parameter of the slow pion and the $D^o$, all of which loosely constrained the decay to be prompt. The pointing angle is defined as the angle between the momentum of the partially reconstructed $D^0$ (since the neutrino is not reconstructed at this point, this is equivalent to the momentum of the reconstructed $K\mu$ pair) and the direction of flight from the best primary vertex to the decay vertex. The Impact Parameter (IP) is defined as the perpendicular distance between a particle’s direction and a particular vertex.

The cuts applied are given in Table 6.1 and Table 6.2. The cut quantities used in this pre-selection were similar to those finally chosen for the HLT2 line discussed in Section 6.1.2, although one notable difference is that the HLT2 line does not require prompt charm decays. Here, FD means ‘flight distance’, the distance of the vector connecting the best primary vertex to the $D^0$ decay vertex. We require the $z$ component of this to be greater than zero to ensure the $D^0$ is moving in a forward direction.

Many of the cuts in Table 6.1 and Table 6.2 use standard variables such as particle momentum ($p$) or transverse momentum ($p_T$). However, some cuts are on variables with names which are not so self-evident and use LHCb jargon. The IP $\chi^2$ for example is defined as a track’s impact parameter from a particular primary vertex measured in units of $\chi^2$, calculated from the uncertainties in track and primary vertex position. In the pre-selection, the IP $\chi^2$ of the muon and kaon were simply required to be positive.

The cuts also use particle ID (PID) information for the kaons. Full details of the LHCb PID system are available elsewhere (see for example [82]). Here we simply require the Difference Log Likelihood ($DLL$) between the particle being a kaon or a pion to be greater than five. For some context, Figure 6.1 gives an example of PID performance for $p-\pi$ and $K-\pi$ separation as a function of track momentum in 2010 data [83].

The input daughter particles are from standard LHCb lists: StdLooseMuons, StdLooseKaons and StdLoosePions. These are the tracks reconstructed in LHCb, which are given the corresponding mass hypothesis and pass a very loose set of PID requirements. Muons are also required to pass an additional boolean quantity known as ISMUON. This quantity is an additional muon PID variable calculated by the LHCb muon group, which

\(^1\)As described in section 2.2, $q^2 = (p_D - p_K)^2$. 

95
Table 6.1: List of loose preselections for \( D^0 \rightarrow K^- \mu^+ \nu_\mu \) applied to signal and minimum bias MC for optimising an HLT2 line.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Cut type</th>
<th>Cut Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Track fit</td>
<td>( \chi^2_{\text{tr.}/n.d.f.} )</td>
<td>( \mu^\pm )</td>
<td>&lt; 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K^\pm )</td>
<td>&lt; 6</td>
</tr>
<tr>
<td>IP or IP ( \chi^2 )</td>
<td>( \chi^2_{IP} )</td>
<td>( \mu^\pm )</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>from any PV</td>
<td></td>
<td>( K^\pm )</td>
<td>&gt; 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( D^0 )</td>
<td>&lt; 0.2 mm</td>
</tr>
<tr>
<td>Transverse momentum</td>
<td>( p_T )</td>
<td>( \mu^\pm )</td>
<td>&gt; 300 MeV/c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( K^\pm )</td>
<td>&gt; 300 MeV/c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( D^0 )</td>
<td>&gt; 800 MeV/c</td>
</tr>
<tr>
<td>Momentum</td>
<td>( p )</td>
<td>( K^\pm )</td>
<td>&gt; 2.0 GeV/c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( D^0 )</td>
<td>&gt; 12.0 GeV/c</td>
</tr>
<tr>
<td>PID DLL(( K - \pi ))</td>
<td>PIDK</td>
<td>( K^\pm )</td>
<td>&gt; 0</td>
</tr>
<tr>
<td>Mass window</td>
<td>( M(\mu K) )</td>
<td></td>
<td>(800, 2000) MeV/c^2</td>
</tr>
<tr>
<td>Distance of closest</td>
<td>DOCA</td>
<td>( K - \mu )</td>
<td>&lt; 0.15 mm</td>
</tr>
<tr>
<td>approach</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertex fit</td>
<td>( \chi^2_{\text{vtx}} )</td>
<td>( D^0 )</td>
<td>&lt; 10</td>
</tr>
<tr>
<td>Pointing angle</td>
<td>DIRA</td>
<td>( D^0 )</td>
<td>&gt; 0.9994</td>
</tr>
<tr>
<td>z co-ordinate of FD</td>
<td>( \delta_z )</td>
<td>( D^0 )</td>
<td>&gt; 0</td>
</tr>
</tbody>
</table>

Table 6.2: List of loose preselections for \( D^{*+} \rightarrow \pi^+ D^0 \) applied to signal and minimum bias MC for optimising an HLT2 line.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Cut type</th>
<th>Cut Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Track fit</td>
<td>( \chi^2_{\text{tr.}/n.d.f.} )</td>
<td>( \pi^+ )</td>
<td>&lt; 10</td>
</tr>
<tr>
<td>PID DLL(( \mu - \pi ))</td>
<td>PIDMu</td>
<td>( \pi^+ )</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Distance of closest</td>
<td>DOCA</td>
<td>( \pi - D^0 )</td>
<td>&lt; 0.4 mm</td>
</tr>
<tr>
<td>approach</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Momentum</td>
<td>( p )</td>
<td>( \pi^+ )</td>
<td>&gt; 1.0 GeV/c</td>
</tr>
<tr>
<td>Vertex fit</td>
<td>( \chi^2_{\text{vtx}} )</td>
<td>( D^{*+} )</td>
<td>&lt; 15</td>
</tr>
<tr>
<td>Minimum IP</td>
<td>( \chi^2_{IP} )</td>
<td>( \pi^+ )</td>
<td>&lt; 100</td>
</tr>
<tr>
<td>( M(\pi K - K \mu) )</td>
<td>( \delta_m )</td>
<td></td>
<td>&lt; 300 MeV/c^2</td>
</tr>
</tbody>
</table>

is set to true if the candidate has certain combinations of hits in the muon stations, as shown in Table 6.3.

It is worth noting that in this section the symbols \( D^0 \) and \( \mu K \) may have similar meanings, since the candidate \( D^0 \) particles considered are simply \( \mu K \) pairs which pass the above selection cuts. The measured mass \( M \) of the candidate \( D^0 \) will not have a value of the true \( D^0 \) mass (1865 MeV/c^2) however, but is instead equal to the invariant mass of the \( \mu K \) pair. Even for a true \( D^0 \rightarrow K^- \mu^+ \nu_\mu \) decay, the value of \( M \) will be lower than the \( D^0 \) mass due to the missing energy carried by the neutrino.
Figure 6.1: RICH mis-identification and efficiency as a function of track momentum (from [83]). Proton identification (red) and pion mis-identification are shown with a requirement of $D LL(p − π) > 5$ and are shown in (a). Similarly, in (b), $K$ and $π$ ID and mis-ID are shown for $D LL(K − π) > 5$.

Table 6.3: Requirements of hits in muon stations for the ISMUON flag (momentum measured in GeV/c).

<table>
<thead>
<tr>
<th>ISMUON Requirements</th>
<th>3 $&lt; p &lt; 6$</th>
<th>M2 + M3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6 $&lt; p &lt; 10$</td>
<td>M2 + M3 + (M4 or M5)</td>
</tr>
<tr>
<td></td>
<td>$p &gt; 10$</td>
<td>M2 + M3 + M4 + M5</td>
</tr>
</tbody>
</table>

The DOCA variable (distance of closest approach) is simply the smallest distance calculated between the two daughter tracks. Clearly, for tracks genuinely originating from the same vertex, this number should be small. There is also a requirement on the $δ_z$ variable.

The final variable is the cosine of the so-called pointing angle (DIRA). For decays where all daughters are correctly reconstructed, this quantity will be very close to 1. For the decay $D^0 \rightarrow K^- μ^+ ν_μ$, apart for the case where the neutrino carries little momentum, the DIRA will be somewhat lower. A cut at 0.9994 on this quantity allows a 35 mrad difference between the momentum- and flight-vectors. This value may appear severe, but typical charm decays in LHCb are highly boosted, so almost all signal decays meet this requirement. The DIRA distributions for signal and MC (normalised to have the same number of total candidates) are shown in Figure 6.2.

Although the cuts were designed to select $D^{*+} \rightarrow π^+ D^0 (K^- μ^+ ν_μ)$ events, the trigger was optimised using only one $D^0$ per event. With very loose requirements on the slow pion, typical $D^0$ decays were selected to make a $D^{*+}$ more than once in combinations with various slow pions. So, to optimise the trigger, an additional offline requirement was added selecting just one $D^0$ candidate in each event.

An additional requirement on pre-selected events was that the selected $D^0$ decay be a
correctly reconstructed genuine decay. This was performed using the LHCb Background Category tool \((IBackgroundCategory)\) [84]. This is a sophisticated tool which classifies each reconstructed candidate based on a comparison to the generator-level Monte Carlo. In the LHCb jargon, generated MC particles are referred to as \(MCParticles\), while reconstructed particles are referred to as \(Particles\).

At the time of writing this thesis, no publicly-available documentation exists for this tool, but LHCb collaborators can view online documentation through the DaVinci Doxygen\(^1\) portal\(^2\).

It is helpful to quote a passage from this online documentation to describe the functions of the IBackgroundCategory tool. Fourteen conditions are evaluated and each is given a letter from \(A\) to \(N\) (the following indented passage is taken verbatim from the online documentation):

- **A**: all final-state particles used to form the candidate are matched to decay products of the same true MC particle (not necessarily the signal)
- **B**: all final-state MC particles originating from the true MC particle defined in \(A\) are matched to particles used to form the candidate (except photons generated by PHOTOS\(^3\)); in case the decay descriptor of the true MC particle defined in \(A\) corresponds to an inclusive decay, “all final-state MC particles” only include the particles required in this (semi-)inclusive decay.
- **C**: all final-state particles used to form the candidate are correctly identified, i.e. have been assigned their correct (true) mass

---

\(^1\)http://doxygen.org
\(^2\)http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/
\(^3\)http://wasm.home.cern.ch/wasm/f77.html
6.1. TRIGGER

- D: the true MC particle defined in $A$ is a signal decay according to the decay descriptor, or the head of a decay chain which differs from the signal decay chain only by the presence or absence of intermediate resonances but has other otherwise the same head, same final state particles and same topology.

- E: the true MC particle defined in $A$ is a signal decay according to the decay descriptor, and all intermediate states of this decay are correctly reconstructed (as listed in the decay descriptor)

- F: the true MC particle defined in $A$ has a mass which does not exceed the mass of the head of the decay descriptor by more than 100 MeV/$c^2$ (tunable parameter)

- G: at least one final-state particle used to form the candidate is a ghost

- H: final-state particles used to form the candidate are matched to true particles from at least two different collisions (pileup)

- I: at least one final-state particles used to form the candidate is matched to a true decay product of a b-hadron (following mother-daughter relationships all the way through)

- J: at least one final-state particles used to form the candidate is matched to a true decay product of a c-hadron (following mother-daughter relationships all the way through)

- K: at least two final state daughters are matched to the same MCParticle.

- L: at least one final state daughter is matched to an MCParticle which is the MCMother of an MCParticle matched to another final state daughter.

- M: at least one final state daughter is associated to an MCParticle from the primary vertex

- N: every final state daughter is associated to an MCParticle from the same primary vertex

These conditions are used to classify candidates into different categories, labelled by a flag called BKGCAT. Using the shorthand of "pseudo-code", the “signal” (BKGCAT==0) flag for the IBackgroundCategory tool is satisfied if:

$$\text{BKGCAT} = 0 \Rightarrow \lnot G \land \lnot K \land \lnot L \land A \land B \land C \land D \land E.$$ (6.1.1)

In addition to the cuts in Table 6.1 and Table 6.2, candidate events were required to have a $D^0$ with a BKGCAT value of 0 for signal and non-zero for background. The number of true signal events selected was 147,867 (7.4%), and the number of background
events was 162,593 (0.8%). In each case a maximum of one candidate $D^0$ was selected per event, to avoid double counting.

The majority of signal events passing the lower-level triggers do so through the muon channels L0Muon and Hlt1TrackMuon. The efficiencies for these lines with the definitions used in the TCK 0x002e002a are shown in Tab. 6.4. For convenience we write the combination of the triggers “L0Muon or L0Hadron” simply as “L0”. For the HLT1 triggers a TOS is implied unless otherwise mentioned.

<table>
<thead>
<tr>
<th>Table 6.4: Efficiencies for lower-level triggers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>L0</td>
</tr>
<tr>
<td>L0 + Hlt1TrackMuon</td>
</tr>
<tr>
<td>L0 + Hlt1Track(Muon OR AllL0)</td>
</tr>
</tbody>
</table>

A different ‘very’ loose selection was applied to the signal MC to give a larger signal sample for the $q^2$ optimisation. To reduce any bias on the $q^2$ distribution, very few cuts were applied. The cuts are listed in Table 6.5. The number of signal events selected by these cuts was 284,999. Again, a maximum of one candidate was accepted per event. This number corresponds to a retention 1.9 times higher than the previous tighter selection. This signal sample was used to perform the $q^2$ optimisation for the HLT2 line.

<table>
<thead>
<tr>
<th>Table 6.5: List of very loose pre-selection for $D^{*-} \rightarrow \pi^+ D^0$ applied to signal MC for $q^2$ optimisation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Track fit</td>
</tr>
<tr>
<td>Transverse momentum</td>
</tr>
<tr>
<td>PID $DLL(K - \pi)$</td>
</tr>
<tr>
<td>PID $DLL(\mu - \pi)$</td>
</tr>
<tr>
<td>ISMUON</td>
</tr>
<tr>
<td>Distance of closest approach</td>
</tr>
<tr>
<td>Mass window</td>
</tr>
<tr>
<td>Momentum</td>
</tr>
<tr>
<td>Vertex fit</td>
</tr>
<tr>
<td>Vertex fit</td>
</tr>
<tr>
<td>$M(\pi K\mu - K\mu)$</td>
</tr>
<tr>
<td>$z$ co-ordinate of FD</td>
</tr>
</tbody>
</table>
6.1.2 Exclusive HLT2 Line

As described above, two steps were used to decide upon cut variables and values used in the HLT2 line (Hlt2CharmSemilepD02HMuNu). The effectiveness of many variables were evaluated using the CROP software [85], and the pre-selected sample of MB and signal MC events. The goal was to arrive at a small number of quantities which were able to separate signal and background while biasing $q^2$ distributions as little as possible.

The second step was to use the very loosely pre-selected signal MC sample to verify that the HLT2 line did not bias the $q^2$ distribution. The final choice of cut set was the result of several iterations of this procedure, while aiming to have an ‘acceptable’ rate on data.

The original plan for this trigger line was to trigger on prompt $D^0 \rightarrow K^-\mu^+\nu_\mu$ decays. However, after discussions within the Charm working group, this was extended to searches for both prompt and non-prompt $D^0 \rightarrow h^-\mu^+\nu_\mu$ (where $h = \pi$ or $K$). The HLT2 line thus has no cuts which restrict the $D^0$ to originating from the primary vertex. The $D^0 \rightarrow \pi^-\mu^+\nu_\mu$ line is identical to the $D^0 \rightarrow K^-\mu^+\nu_\mu$ line apart from the hadron mass hypothesis, and its addition allows scope for further measurements in this similar (but Cabbibo-suppressed) mode.

There are two main reasons for including non-prompt charm decays in this trigger. Firstly, since a large number of $b$ mesons are produced at LHCb, secondary $D$ decays such as $B^- \rightarrow D^0(K^-\mu^+\nu_\mu)h^-$ provide a cleaner sample of semileptonic $D^0$ decays.

Secondly, such modes would also provide a channel through which one could search for $CP$ violation. For example, the sign of the hadron in the previous decay uniquely tags the created $D$ as either $D^0$ or $\bar{D}^0$, and similarly the sign of the muon from the $D$ decay tags the $D$ when it decayed. A comparison of the number of created and decaying $D$ mesons of each type could enable a measurement of $CP$ violation in this system.

The optimisation was performed to a Figure of Merit (FoM) of $S/\sqrt{S+B}$ with a weighting of 100 for background events. The final choice of which variables to use in the optimisation procedure was made over several iterations with CROP. The aim was to arrive at a small set of variables which were able to achieve a relatively high FoM. The background weighting is included because we expect a large amount of signal, so it is reasonable to sacrifice some overall trigger rate to improve signal purity. The weighting means the optimisation is more favoured towards rejecting background than accepting signal. This is particularly important considering the bandwidth restrictions of the LHCb HLT2.

The seven key quantities used in the HLT2 line are: the $D^0$ flight distance, the $p_T$ sum of $K$ and $\mu$, the $p_T$ of the $K$ and $\mu$ separately, the $D^0$ momentum, the distance of closest approach (DOCA) between the $K$ and $\mu$ tracks and the corrected mass, $M_{corr}$ of the $D^0$. The distributions for signal (blue) and background (red) for the seven cut quantities
selected are shown in Figures 6.3–6.6. The distributions have been normalised to have the same total number of candidates.

![Figure 6.3: Signal (black) and background (red) distributions for HLT2 cut variables: $D^0$ flight distance in mm (left) and the logarithm of the sum of daughter $p_T$ (right).](image)

![Figure 6.4: Signal (black) and background (red) distributions for HLT2 cut variables: logarithm of kaon $p_T$ (left) and logarithm of muon $p_T$ (right).](image)

The corrected mass is a quantity designed to return a value close to the true mass for decays with missing energy, and was introduced in (2.3.12):

$$M_{\text{corr}} = \sqrt{M^2 + |p_{T\text{miss}}|^2 + |p_{T\text{miss}}|},$$  

(6.1.2)

where $M$ is the original mass reconstructed for the particle and $p_{T\text{miss}}$ is the missing
Some variables such as kaon and muon transverse momenta and mother momentum are common quantities to use in the LHCb trigger and proved powerful for this trigger also. Others, for example the sum of muon and kaon transverse momenta, are used less commonly in the trigger, but were retained due to good performance separating background and signal.
Several cut variables were used in the trigger although they were not optimised:

- the $D^0$ vertex $\chi^2$ – the cut was set to a maximum of 10;
- track $\chi^2$/d.o.f.. This is not perfectly simulated in the Monte Carlo, and the cut was set to the commonly used maximum of 3.0;
- the $\delta z$ distance between the reconstructed $D^0$ and primary vertices was required to be positive;
- the mass of the $\mu^-K^+$ pair was required to be less than 1900 MeV/$c^2$. The $D^0$ mass is 1.86 GeV/$c^2$, so this cut removes candidates which are obviously background.

The cut on $\delta z$ rejects the unusual class of events where the $D^0$ flight direction is reconstructed to be backwards, while the cut on the mass of the daughters rejects background events by cutting at a value above the $D^0$ mass. A specific trigger path is also required. The HLT2 line requires either L0Muon or L0Hadron to fire, and additionally TOS on Hlt1TrackMuon.

Additionally the possibility of making a global event cut (GEC) is allowed, to accept only events with fewer than 120 tracks. This is essentially a rate limiter and at present it is turned off by default. If the line begins to take too much HLT2 bandwidth, the GEC could be turned on (this could occur as the luminosity and number of primary vertices increases).

No requirements were made on the number of tracks during HLT2 optimisation.

The linear correlations between the variables are shown in Figure 6.7. Here, the correlation is the correlation factor used in ROOT [69], which evaluates the covariance between the two variables, scaled by the RMS of each distribution. The scaling means that quantities which are 100% correlated (anti-correlated) have a correlation value of 1.0 (-1.0). Completely un-correlated variables have correlation values of 0.

The strongest correlations are between the sum of daughter $p_T$ and the kaon $p_T$ (0.552) and between the $D^0$ momentum and the corrected mass (0.523). All other pairs of variables have correlations below 0.5.

The cuts used in the HLT2 line are shown in Table 6.6. Applied to the MC samples, these cuts accepted 5,203 of the 9,779 signal events with TOS on Hlt1TrackMuon (53.2%). For background events, the number was 21 out of 290 (7.2%). The percentages were slightly higher for the events which passed the combination of either TOS on Hlt1TrackMuon or Hlt1TrackAllL0: 55.0% for signal events and 8.9% for background events.

Allowing only a single HLT1 line significantly reduces difficulties in assessing systematic uncertainties and biases from the trigger. Since the combination of allowing either Hlt1TrackMuon or Hlt1TrackAllL0 lines does not significantly improve the ratio of signal to background events, the decision was made to only consider events TOS on the Hlt1TrackMuon line in this HLT2 line.
6.1. TRIGGER

Figure 6.7: Correlations between variables used in the HLT2 line.

Table 6.6: Cuts used in the $D^0 \rightarrow K^-\mu^+\nu_{\mu}$ Hlt2 line.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Cut type</th>
<th>Cut Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^0$ flight distance from ‘best’ PV</td>
<td>FD($D^0$)</td>
<td>&gt;</td>
<td>4.0 mm</td>
</tr>
<tr>
<td>Corr. mass</td>
<td>$M_{\text{corr}}(D^0)$</td>
<td>&gt;</td>
<td>1400 MeV/c$^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;</td>
<td>2700 MeV/c$^2$</td>
</tr>
<tr>
<td>Mass of $\mu - K$ pair</td>
<td>$M(\mu K)$</td>
<td>&lt;</td>
<td>1900 MeV/c$^2$</td>
</tr>
<tr>
<td>Sum of daughter $p_T$</td>
<td>Sum $p_T$</td>
<td>&gt;</td>
<td>2800 MeV/c</td>
</tr>
<tr>
<td>Muon transverse momentum</td>
<td>$p_T^\mu$</td>
<td>&gt;</td>
<td>800 MeV/c</td>
</tr>
<tr>
<td>Kaon transverse momentum</td>
<td>$p_T^K$</td>
<td>&gt;</td>
<td>600 MeV/c</td>
</tr>
<tr>
<td>Distance of closest approach ($K - \mu$)</td>
<td>DOCA</td>
<td>&lt;</td>
<td>0.07 mm</td>
</tr>
<tr>
<td>$D^0$ momentum</td>
<td>$p_D$</td>
<td>&gt;</td>
<td>20 GeV/c</td>
</tr>
<tr>
<td>Vertex fit</td>
<td>$\chi^2_{\text{vtx}}$</td>
<td>&lt;</td>
<td>10</td>
</tr>
<tr>
<td>$z$ co-ordinate of $D^0$ FD</td>
<td>$\delta_z$</td>
<td>&gt;</td>
<td>0</td>
</tr>
<tr>
<td>Track fit of daughters</td>
<td>$\chi^2_{tr}/n.d.f.$</td>
<td>&lt;</td>
<td>3</td>
</tr>
</tbody>
</table>

The bias on $q^2$ introduced by cutting on the muon $p_T$ is one of the most significant ($q^2$ biases will be discussed in more detail later in this section). To prevent additional biases being introduced by this HLT2 line, the $p_T$ cut on muon momentum is set to be the same as that used in Hlt1TrackMuon.

The line for $D^0 \rightarrow \pi^-\mu^+\nu_{\mu}$ is identical to the kaon line apart from the hadron mass hypothesis. Unlike $D^0 \rightarrow K^-\mu^+\nu_{\mu}$, this decay is Cabbibo suppressed, and the branching fraction is an order of magnitude lower (0.2% compared to 3.3%). No particle ID is
performed at trigger-level, so the line with the pion mass hypothesis triggers on a very similar set of events to the kaon line and bandwidth requirements are not significantly increased.

Also included are wrong-sign lines (i.e. searches for the decays $D^0 \rightarrow h^+ \mu^+ \nu_\mu$) for both decays. There is potential for confusion in using the term ‘wrong-sign’, which could also be used to describe a ‘mixed’ $D^0$ decay where the slow pion and muon have opposite signs: $D^{*+} \rightarrow \pi^+ D^0 \rightarrow \pi^+ D^0 (K^+ \mu^- \nu_\mu)$. In this thesis, the term ‘wrong-sign’ refers to the same-sign decay $D^0 \rightarrow h^+ \mu^+ \nu_\mu$ unless otherwise specified.

The wrong-sign lines are prescaled by a factor of 10 and can be used to estimate the combinatorial background. The prescale means only one event in every ten is retained. This is a simple way to reduce bandwidth for this class of events, which we assume contains only background events.

If we assume the LHCb acceptance and particle ID are identical for positively- and negatively-charged particles of the same species, to a good approximation the number of opposite-sign and same-sign random $\mu^- K^-$ pairs (i.e. not from the same genuine decay) passing the selection cuts should be the same. The approximation is more exact if, as expected, the numbers of events recorded under each magnet polarity are the same.

Since the running conditions and L0 trigger settings of LHCb change frequently, it is difficult to calculate the overall rate of this HLT2 line except by running it in real time. This HLT2 line was developed in early 2011 and at the time the LHCb Trigger and Stripping group provided a testing sample for HLT optimisation. The sample was recorded in the SDST format, which is a partially reconstructed format used in the HLT that is not normally available for offline analysis. This sample contained around 114,000 events which had passed a certain L0 trigger configuration commonly used in early 2011 (configuration 0x002a under the standard LHCb TCK notation).

When tested on this training sample, the HLT2 lines accepted ten events in both the kaon and pion right-sign channels. The numbers retained in the wrong-sign channels were three and one in the pion and kaon lines respectively. The SDST sample was constructed by the Trigger and Stripping group to replicate a L0 output rate of 10Hz. Thus the acceptance rate of the kaon line was approximately 100 Hz, for data at the instantaneous luminosity of the testing sample.

The HLT2 line contains a configurable option to prescale the line. During the period of this line’s development, the instantaneous luminosity in LHCb increased markedly while the maximum allowed output rate of HLT2 stayed constant. In June 2011, the prescale for the right-sign lines was set to retain one event in twenty and the wrong-sign lines were set to 1/100. Since the kaon and pion lines only differ by a single mass hypothesis, there is significant overlap between each pair of lines. The chosen prescales give an overall rate for the four lines of approximately 20 Hz.

The most precise measurements of the form factors for $D^0 \rightarrow K^- \mu^+ \nu_\mu$ are at electron-
6.1. TRIGGER

positron colliders, which have a very different background environment to the LHC, so the amount of statistics necessary is not directly comparable. Nevertheless, the number of events used in these analyses can be used as a benchmark for the statistics required at LHCb. For example the 2009 study by CLEO [21] used a sample of around 700,000 $D^0$ candidates from the full CLEO dataset of 818 pb$^{-1}$. Belle [26] and BaBar [22] used samples of 56,000 and 85,000 $D^0$ candidates respectively.

If we assume a conservative signal-to-noise of 4, a trigger operating at 10 Hz would accept 8 Hz of signal. At this rate, LHCb would record approximately 700,000 candidate decays (equivalent to the CLEO dataset) every 24 hours. For this reason, it was decided that an overall rate in the region of 10 – 20 Hz was acceptable for these channels.

In addition to the four triggers described above, a fifth, similar trigger runs at LHCb and is designed to be used for $D$ mixing searches in the channel $D^{*+} \to \pi^+ D^0 (K^- \mu^+ \nu_\mu)$. This is a rare process, so the strategy for this analysis is different than for a form factor measurement. In particular, high statistics are more crucial than the elimination of $q^2$-bias.

To meet these requirements the fifth trigger line (with the suffix ‘tight’) runs with identical cuts to those described above, except without a prescale and with a much harder cut on $D^0$ flight distance, of 40 mm. The trigger rate decreases almost exponentially as a function of flight distance. For example, increasing the flight distance cut from 4 mm to 15 mm approximately halves the trigger rate, and a cut at 70 mm accepts no events in the training sample.

There is a greater chance of backgrounds from the primary vertex polluting the signal for candidates with low flight distance. So, the very hard flight distance cut results in higher purity and a low rate without the need for a prescale, although it significantly biases $q^2$. A cut of 40 mm was chosen for the tight line because at this value, the rate is similar to the standard $D^0 \to K^- \mu^+ \nu_\mu$ line with a prescale of 0.05.

6.1.3 Bias on $q^2$

The trigger cut variables and values were chosen to bias the $q^2$ distribution as little as possible. In particular, the cut set was optimised to bias $q^2$ minimally after HLT1. From section 2.2, $q^2$ in the decay $D^0 \to K^- \mu^+ \nu_\mu$ is the invariant mass of the lepton pair, or equivalently:

$$q^2 = (p_D - p_K)^2. \tag{6.1.3}$$

Kinematic constraints mean the allowed values of $q^2$ for this decay have a range of less than 2 GeV$^2$. In the rest-frame of the $D^0$, low $q^2$ values occur when the leptons travel parallel to each other – in this case the daughter $s$ quark has a large recoil against the virtual $W$. The highest $q^2$ values occur when the $K$ is created almost at rest and the
leptons are produced almost back-to-back. The limits (assuming $c = 1$) are given by:

$$q^2_{\text{min}} = m_{\mu}^2 \approx 0.01 \text{ GeV}^2 \quad (6.1.4)$$

and

$$q^2_{\text{max}} = (m_D - m_K)^2 \approx 1.9 \text{ GeV}^2 \quad (6.1.5)$$

The configurations for $q^2_{\text{max}}$ and $q^2_{\text{min}}$ are illustrated in Figure 6.8, an illustration adapted from [9].

Figure 6.8: Adapted from [9], the kinematics for the semileptonic decay of a $c$ meson in its rest-frame: (a) before decay; (b) for high $q^2$ values ($q^2 \rightarrow q^2_{\text{max}}$) the $s$ quark is created almost at rest and the leptons are back-to-back; (c) for $q^2 \rightarrow q^2_{\text{min}}$, the lepton momenta are parallel and the $s$ quark receives a corresponding momentum ‘kick’.

The true $q^2$ distribution for generator-level Monte Carlo is shown in Figure 6.9. The sample contains approximately one million $D^*^+ \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_{\mu})$ events generated in Gauss v39r1 using PYTHIA 6 [64].

In this study, the word bias refers to preferentially rejecting or accepting events in certain regions of $q^2$. A selection which was equivalent to scaling each bin of $q^2$ by
6.1. TRIGGER

![Distribution Plot]

Figure 6.9: Generator-level true $q^2$ distribution for approximately one million $D^{*+} \rightarrow \pi^+D^0(K^-\mu^+\nu_\mu)$ events with daughters within the LHCb acceptance.

the same constant is considered un-biased. Although the HLT2 line is designed to bias $q^2$ as little as possible after the lower-level triggers, it is almost inevitable that the final distribution of events seen after HLT1 will be somewhat biased from the generator-level distribution.

The sample used to assess $q^2$ biases had 284,999 true $D^0$ events. Of these, 76,269 (26.8%) pass with TOS on either L0Muon or L0Hadron. Of the sample passing L0, 12,889 (16.9%) pass with TOS on Hlt1TrackMuon.

The probability for tracks from a decay with a certain $q^2$ to be correctly reconstructed varies in a complex way. For instance, at high $q^2$ values, the kaon is created almost at rest relative to the $D^0$ and for prompt charm events this could mean an increased possibility of a track which is close to the beam pipe and more difficult to reconstruct. Although the Monte Carlo sample is designed to replicate the acceptance region of LHCb, at generator-level this is performed with a simple cut on the maximum and minimum track angle and does not account for the exact shape of the detector (this is performed during the reconstruction stage). For this reason, biases are evident for events with a $q^2$ which favours daughters in certain regions of the detector (for instance close to the beam pipe).

Additionally, some momentum cuts on $D^0$ daughters are used in both stages of the lower-level triggers. Since $q^2$ is a quantity highly dependent on momentum, such cuts are potentially significant. The final choice of cuts in the HLT2 line was designed to minimise the $q^2$-bias after HLT1.

The effect of applying a single additional cut to certain base $q^2$ distributions (generator-level, reconstructed MC with loose cuts and after HLT1) was quantified with a simple $\chi^2$
value. To compare two distributions, histograms were made for each distribution in 100 bins of $q^2$ between 0 and 2 GeV. The two distributions were subtracted and the resulting histogram was rescaled by the original base distribution.

If one distribution $g(q^2)$ is equivalent to a base distribution $f(q^2)$ with events removed without $q^2$-bias, this is equivalent to

$$g(q^2) = \alpha f(q^2),$$

(6.1.6)

where $\alpha < 1$ is a positive constant. In this case, the above subtraction and rescaling is equivalent to

$$\frac{f(q^2) - g(q^2)}{f(q^2)} = 1 - \alpha,$$

(6.1.7)

assuming non-zero $f(q^2)$. Thus, the resulting distribution for cut without $q^2$-bias should be a straight line. For simplicity the resulting distribution will be referred to as the $q^2$-bias distribution. A horizontal line is fitted to each $q^2$-bias distribution over the range 0–2.0 GeV, and a $\chi^2$ value is calculated from the spread of data around this line.

There are several possibilities for how to define the uncertainty on each point in these distributions. Since the resulting $\chi^2$ values calculated from these errors are used only for comparisons among a set of similar distributions (and not for an overall calculation of statistical or systematic uncertainty) a fairly simple choice is made.

For comparisons where one distribution is a subset of another, we assume binomial errors. Letting $f_i$ and $g_i$ represent the number of entries in bin $i$ of the base and subset distributions, the proportion of events from the base distribution accepted (equivalent to the binomial probability) is $p_i = g_i / f_i$, and the value of bin $i$ in the $q^2$-bias distribution is $1 - p_i$.

The binomial uncertainty on $g_i$ is then given by:

$$\sigma(g_i) = \sqrt{f_ip_i(1-p_i)},$$

(6.1.8)

and we take the uncertainty on each point in the $q^2$-bias distribution to be:

$$\sigma(1-p_i) = \sigma(g_i)/f_i = \sqrt{\frac{p_i(1-p_i)}{f_i}}.$$  

(6.1.9)

Some comparisons are also made between two independent distributions, between generator-level MC and independently reconstructed and generated MC. In this case the generator-level MC (which has many more entries) is rescaled to have the same integral as the reconstructed distribution. If bin $i$ contains $N$ entries, a statistical error of $\sigma_i = \sqrt{N}$ is assigned to each bin. When two histograms are added or subtracted, the errors are added in quadrature. When two histograms are divided, the errors are assumed to be
uncorrelated using the standard ROOT sum of squares of weights method\(^1\).

For a best-fit line at \( y = \mu \), a \( \chi^2 \) value is calculated for the 100 bins in the \( q^2 \)-bias histogram:

\[
\chi^2 = \sum_{i=1}^{100} \left( \frac{x_i - \mu}{\sigma_i} \right)^2
\]

(6.1.10)

where \( x_i \) and \( \sigma_i \) are the value and error in bin \( i \). In the case where there are no entries in bin \( i \) (and \( \sigma_i \) is zero) the bin is ignored. Typically this happens for a no more than five bins at high \( q^2 \).

The \( q^2 \) and \( q^2 \)-bias distributions for generator-level compared to pre-selected signal MC are shown in Figure 6.10. The generator-level distribution has been rescaled to have the same total number of events as the signal MC sample. At \( q^2 > 1.5 \) GeV there is a clear difference between the two distributions, with a relative deficit of pre-selected events in this region. For other values of \( q^2 \) the comparison is much closer, with the exception of the lowest few bins.

![Figure 6.10: Generator-level (black) and loose pre-selected (green) MC \( q^2 \) distributions (left) and the corresponding \( q^2 \)-bias distribution (right). The generated distribution has been rescaled to have the same total number of events as the pre-selected distribution. The best-fit horizontal line is drawn in red. The \( \chi^2 \) value is 1090.](image)

The MC samples contain prompt \( D^{*+} \to \pi^+ D^0 (K^- \mu^+ \nu_\mu) \) decays, which are produced preferentially at high rapidity (i.e. almost parallel to the beam pipe). Thus, kaons from high \( q^2 \) decays, where they are produced almost at rest relative to the \( D^0 \), are also

produced preferentially close to the beam pipe. Tracks close to the beam pipe have a lower probability of being reconstructed, and we see a decrease between generated and reconstructed MC in the high $q^2$ region.

The biases evident in Figure 6.10 are difficult to compensate for with the trigger, since they relate to the probability of track reconstruction for certain $q^2$ values. However, provided the effect of reconstruction efficiency is known sufficiently well, the effects can be unfolded during offline analysis. The $\chi^2$ value over the full range of the $q^2$-bias plot in Figure 6.10 is 1090 (in this section, all $\chi^2$ values will be given to three significant digits unless otherwise stated).

Similarly, the biases added by the lower-level triggers can be determined by comparing the pre-selected distribution (which has no trigger applied), to the same distribution after L0 and HTL1. The triggers required in the HLT2 line are Hlt1TrackMuon, in addition to either L0Muon or L0Hadron. The $q^2$ distributions for signal MC after pre-selection, L0 and HLT1 are shown in Figure 6.11, and the $q^2$-bias distributions for the L0 and HLT1 distributions compared to pre-selected MC are shown in Figure 6.12.

![Figure 6.11: Pre-selected (green), passing L0 (red) and passing HLT1 (blue) $q^2$ distributions. To demonstrate the relative efficiencies, the distributions have not been re-scaled.](image)

The $\chi^2$ values of the L0 and HLT1 $q^2$-bias distributions compared to pre-selected MC (Figure 6.12) are 472 and 332. This indicates that although some (perhaps unavoidable) bias is introduced by the pre-selection, the additional biases introduced by L0 and HLT1 are smaller.

The size of the error bars in the bins of highest $q^2$ are a reminder that the statistics in these bins are very limited, so one should not draw conclusions from the last few bins of the $q^2$-bias distributions. Nevertheless, from Figure 6.12 it is apparent that the biases are most evident at high $q^2$, and are in the opposite direction to the bias introduced by the...
6.1. TRIGGER

Figure 6.12: Pre-selected–L0 (left) and pre-selected–HLT1 (right) $q^2$-bias distributions for Figure 6.11. The $\chi^2$ values for the distributions are 472 and 334 respectively.

pre-selection shown in Figure 6.10. Both L0 and HLT1 select slightly more events at high $q^2$ than if they were un-biased.

Similarly, the effect of each HLT2 cut on the HLT1 distribution can be quantified. The cut on muon transverse momentum is the same in HLT1 and HLT2, so the $q^2$-bias distribution is uniformly zero by design. The $q^2$ and $q^2$-bias distributions for the other optimised cut variables ($D^0$ flight distance, $D^0$ corrected mass, sum of daughter transverse momenta, kaon transverse momenta, daughter DOCA and $D^0$ momentum) compared to HLT1 are shown in Figures 6.13–6.18, for the cuts listed in Table 6.6.

Of the cuts listed in Table 6.6, the two with the greatest $q^2$ biases are the sum of daughter transverse momenta and $D^0$ corrected mass, which have $\chi^2$ values of 135 and 142 respectively. The distributions of these shown in Figures 6.3 and 6.6 show that these are two of the most effective at distinguishing signal and background. For this reason, the cuts were made as tight as possible while maintaining a $\chi^2$ value of order 1 per degree of freedom. To achieve a similar level of signal/background separation on other variables such as muon momentum or transverse momentum resulted in $\chi^2$ values that were typically greater than 2 per degree of freedom.

The final choice for the HLT2 line was to use $D^0$ flight distance to select for the displaced vertex. During the optimisation procedure, two other quantities were assessed: $D^0$ flight distance $\chi^2$ and the minimum IP $\chi^2$ of the daughters to the best primary vertex. After HLT1, the HLT2 flight distance cut retains 83.6% of signal events. A similar efficiency for flight distance $\chi^2$ is achieved with a cut at 90, and for minimum daughter IP $\chi^2$...
Figure 6.13: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on $D^0$ flight distance (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 103.

Figure 6.14: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on $D^0$ corrected mass (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 142.
6.1. TRIGGER

Figure 6.15: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on the sum of daughter transverse momenta (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 135.

Figure 6.16: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on kaon transverse momentum (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 95.
CHAPTER 6. TRIGGER AND STRIPPING FOR $D^0 \rightarrow K^-\mu^+\nu_\mu$

Figure 6.17: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on daughter DOCA (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 107.

Figure 6.18: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on $D^0$ momentum (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 89.
6.1. TRIGGER

an equivalent cut is at 10.5.

The $q^2$ and bias distributions for these two cuts are shown in Figure 6.19 and 6.20. The $\chi^2$ values of the distributions are 219 and 393 respectively, which compare poorly to 103 for the equivalent cut on flight distance. Both display positive biases at high $q^2$. For this reason, a simple flight distance cut was chosen for the HLT2 line.

![Figure 6.19: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on flight distance $\chi^2$ of 90 (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 219.](image)

The combined HLT2 line outlined in Table 6.6 can be compared to the output HLT1 and the generator-level MC. The $q^2$ and bias distributions are shown in Figure 6.21, with the generator-level distribution rescaled to have the same number of events. The distributions can be compared to the equivalent ones for the pre-selected distribution without any additional cuts in Figure 6.10.

The HLT2 cuts accept 6229 of the 12,889 events that pass with TOS on Hlt1TrackMuon (48.3%). Thus from the original sample of 284,999 very loosely pre-selected true signal MC events, the percentage passing the HLT2 line is 2.2%. Comparatively, the statistics in Figure 6.21 are very much lower than in Figure 6.10, but many of the properties of the distributions are similar. In particular, the $q^2$-bias seems to be small from approximately 0.05 GeV$^2$ (around the third histogram bin) to 1.5 GeV$^2$. In this figure, if a bin in the HLT2 distribution has no entries, the uncertainty in the bias plot is set to 100%.
Figure 6.20: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on minimum daughter $\chi^2$ of 10.5 (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 393.

Figure 6.21: Generator-level (black) and HLT2 selected (red) MC $q^2$ distributions (left) and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 135.
6.2 Stripping

As mentioned above, stripping is a further stage of data skimming which takes place after the high-level triggers and full data reconstruction. The stripping line for $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ firstly reconstructs right-sign and wrong-sign $D^0 \rightarrow K^\pm \mu^+ \nu_\mu$ decays using slightly tighter cuts than exist in the HLT2 trigger (including some on particle ID), then searches for a slow pion to reconstruct a $D^{*+} \rightarrow \pi^+ D^0$ decay.

Stripping lines are kept in the package Phys/StrippingSelections, and the line for $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ has the name StrippingDstarD02KMuNu. The line retains a relatively high number of HLT2 events and for this reason is written to the so-called microDST – a file format similar to LHCb’s standard DST, but with less information stored per event (for more details see [87]).

The microDST is designed for high-statistics channels, and the charm working group is a major user of this data format. Many charm channels are stored to the microDST, including $D^0 \rightarrow \mu^+ \mu^-$ and $D^0 \rightarrow h^+ h^-$. The format is the same as the DST, but with certain information removed to reduce the size. In particular, only information directly relevant to each analysis is included in the microDST. For example for an analysis selecting $B$ mesons, only the selected candidates and decay products are retained, along with other important information including the related vertices.

Although some information is lost in comparison to the full DST, the microDST has the benefit of allowing an order of magnitude more statistics to be written to disk after stripping. One drawback of the microDST is that it is not possible to make new vertices with other tracks in the original event, since this information is not retained. For example, it is not possible to combine a reconstructed $D^{*+} \rightarrow \pi^+ D^0$ with another hadron ($h$) in the event to make a $B \rightarrow h D^{*+}$ vertex. This is generally possible in the full DST.

This information is not lost permanently, since each triggered event is written to tape. To perform the sort of recombination described above, it is necessary to restrip the data with the additional requirements included. At the time of writing, new releases of the stripping software at LHCb are made approximately every six months.

The cuts used in the stripping line are listed in Table 6.7. With the exception of particle ID, the quantities used in selecting $D^0$ candidates are the same as for the trigger line. As for the trigger line, the daughter particles are taken from the StdLoose lists.

The selection requirements for $D^*$ candidates are listed in Table 6.8. The pion is often referred to as ‘slow’ and given the symbol $\pi_s$ because the typical momentum is relatively low.

Distributions of the three most important variables for $D^*$ selection ($\delta_m$, DOCA and vertex $\chi^2$) are shown in Figures 6.22 and 6.23. The strongest variable for distinguishing signal and background is the mass difference between the $D^*$ and $D^0$ candidates, written $\delta_m$. For genuine signal events, this peaks very sharply at approximately 150 MeV/$c^2$. 

119
Table 6.7: List of cuts used for $D^0$ selection in the $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_{\mu})$ stripping line.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Cut type</th>
<th>Cut Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^0$ flight distance</td>
<td>$FD(D^0)$</td>
<td>$&gt;$</td>
<td>4.0 mm</td>
</tr>
<tr>
<td>from ‘best’ PV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corr. mass</td>
<td>$M_{corr}(D^0)$</td>
<td>$&gt;$</td>
<td>1400 MeV/c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$&lt;$</td>
<td>2700 MeV/c</td>
</tr>
<tr>
<td>Mass of $\mu - K$ pair</td>
<td>$M(\mu K)$</td>
<td>$&lt;$</td>
<td>1900 MeV/c</td>
</tr>
<tr>
<td>Sum of daughter $p_T$</td>
<td>$Sum p_T$</td>
<td>$&gt;$</td>
<td>2800 MeV/c</td>
</tr>
<tr>
<td>Kaon transverse momentum</td>
<td>$p_T^K$</td>
<td>$&gt;$</td>
<td>800 MeV/c</td>
</tr>
<tr>
<td>Distance of closest approach</td>
<td>$DOCA$</td>
<td>$&lt;$</td>
<td>0.07 mm</td>
</tr>
<tr>
<td>$(K^- - \mu)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D^0$ momentum</td>
<td>$p_D$</td>
<td>$&gt;$</td>
<td>20 GeV/c</td>
</tr>
<tr>
<td>Vertex fit</td>
<td>$\chi^2_{vtx}$</td>
<td>$&lt;$</td>
<td>9</td>
</tr>
<tr>
<td>$z$ co-ordinate of $D^0$ FD</td>
<td>$\delta_z$</td>
<td>$&gt;$</td>
<td>0</td>
</tr>
<tr>
<td>Track fit of daughters</td>
<td>$\chi^2_{tr. / n.d.f.}$</td>
<td>$&lt;$</td>
<td>3</td>
</tr>
<tr>
<td>Kaon PID $DLL(K - \pi)$</td>
<td>$K_{PIDK}$</td>
<td>$&gt;$</td>
<td>5</td>
</tr>
</tbody>
</table>

Since stripping takes place after reconstruction has been completed, a significant difference between the stripping and trigger lines is the presence PID information. PID information is usually given as a difference log likelihood (DLL) between two species, and by default the comparison species is a pion. The DLL for muon identification has the symbol $PID_{\mu\mu}$, which represents $DLL(\mu - \pi)$.

The PID response is not yet particularly well-modelled in the LHCb Monte Carlo, and the PID requirements in the stripping line are set relatively loose to allow for this. In particular, no PID variable was used during the optimisation stage in developing the trigger and stripping lines. The PID response depends strongly on particle momentum, so cutting unnecessarily hard on such quantities could potentially affect the final $q^2$ distributions in unpredictable ways.
The ability of LHCb to separate kaons and pions depends on the precision with which Cherenkov rings in the RICH can be measured. For very high-momentum kaons and pions (over 50 GeV/c), the Cherenkov angles are very similar between the two and the species become almost indistinguishable to the PID system. For this reason, a strong cut on kaon PID is implicitly a cut on kaon momentum.

Despite the PID not being well-modelled in MC, the PID values can still be used to quantify and estimate the response of data to certain PID cuts. Similarly to the $q^2$-bias plots shown previously, the effect of applying a PID $>5$ cut to HLT1 output signal MC is shown in Figure 6.24. The $\chi^2$ value of the $q^2$-bias distribution is 115, which is lower than
the $\chi^2$ effect of the $D^0$ flight distance cut. Thus, at least in Monte Carlo, the bias caused by a kaon PID cut is not significantly more than that caused by the cut on flight distance.

![Figure 6.24: HLT1 MC $q^2$ distribution with (red) and without (black) a cut on kaon PID (left), and the corresponding $q^2$-bias distribution (right). The $\chi^2$ value is 115.](image)

The stripping line also has a loose requirement on the slow pion being not ‘muon-like’, requiring $\text{PID}_{\text{mu}} < 1$. A particle which is very likely to be a muon has a $\text{PID}_{\text{mu}}$ value significantly above 1. The stripping line also reproduces the trigger requirement that the muon passes the isMuon selection. A requirement is also made on the PID of the muon for it to be ‘unlike’ a pion, by requiring $\text{PID}_{\text{mu}} > -1$.

Muon identification is based on a complex set of algorithms (see for example [88]), but the most critical is how far a candidate particle penetrates the iron absorbers in the muon system. At high momentum, muons have a high probability of traversing the entire muon detector. Figure 6.25 (from [88]) shows the efficiency and mis-identification rates for two common muon identification algorithms as a function of momentum: isMuon and isMuonLoose. In the mis-identification plot of Figure 6.25, a distinction is made between hadrons which do (circles) and do not (triangles) decay in flight. Clearly, in all cases the best muon identification occurs at high momentum.

As with the trigger, each stripping line must meet criteria for retention and processing time. Since stripping is performed offline (i.e. not in real-time), these requirements are slightly more flexible than for the trigger. All data passing HLT2 are stored, so if necessary through error, oversight or the addition of new modes of interest, old data can be restrippled. Clearly this is different for the trigger, and any event not passing an HLT2
6.2. STRIPPING

Figure 6.25: The efficiency (left) and mis-ID rates (right) for the muon PID algorithms isMuon and isMuonLoose, from [88]. In the mis-ID plots, the contribution from hadrons which do/do not decay in flight are shown separately.

The stripping line guidelines established by the stripping convenors allow up to 1 ms of CPU time per event, and a total retention of 0.05% of HLT2-passed events for the full DST. In the microDST, the guideline is less strict, and lines are required not to have retention significantly above 0.5%. In tests for Stripping 15 (July 2011) all stripping lines were tested on a sample of 48,215 ‘good’ HLT2-passed events. Here the word ‘good’ is shorthand for events which successfully pass the reconstruction stage and are thus eligible for the stripping stage.

The two stripping lines considered here are given the labels RS (right-sign) and WS (wrong-sign) for the decay modes $D^{*+} \rightarrow \pi^+ D^0(K^- \mu^+ \nu_\mu)$ and $D^{*+} \rightarrow \pi^+ D^0(K^+ \mu^+ \nu_\mu)$ respectively. The retentions on the testing sample were 243 (0.50%) and 53 (0.11%) events in the RS and WS modes respectively. It is important to note that the $D^0 \rightarrow h^- \mu^+ \nu_\mu$ trigger was not running when the testing sample was recorded – so the true rate of the stripping line on data with this trigger functioning is slightly higher. The average CPU time per event on the two lines was 0.592 ms and 0.292 ms respectively.

Thus, although the total retention is close to the recommended limit, the lines are acceptable for writing to the microDST. The retention is comparable to other charm modes written to the microDST. For example, the lines for $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D^+ \rightarrow K^- K^+ \pi^+$ retained 0.69% and 1.65% of events respectively.

The limit on rate is due the number of events LHCb can afford to write to storage. Should the retention rate of this line become too high, similarly to the trigger, the stripping lines have configurable prescales which can be set to reduce the rate to any number.

\footnote{Technically, LHCb also keeps a small stream of no-bias data, which is a random sampling of events without any trigger requirements.}
CHAPTER 6. TRIGGER AND STRIPPING FOR $D^0 \rightarrow K^- \mu^+ \nu_\mu$

desired.

The stripping cuts were applied to MC10 signal and background MC samples. The signal MC contained approximately four million simulated signal events, with almost equal proportions of data with each magnet polarity (usually labelled Magnet Up or Magnet Down). The number of candidate $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ decays found in the Magnet Up and Magnet Down samples was 20,841 and 20,899 respectively while for the equivalent wrong-sign decays the numbers were 504 and 534 respectively.

These numbers do not however correspond to the total efficiency, because of events with multiple candidates. There are several strategies commonly employed to manage multiple candidates (see for example [89]) and this will be discussed in more detail in the following chapter. For analyses with tight selection, multiple-candidate events usually come from one of three classes: genuine multiple, clones and overlapping candidates.

Clones occur when a single genuine track is found twice (or more) during the reconstruction phase. LHCb employs ‘clone killing’ software to remove as many of this class as possible, but it is very difficult to remove them entirely. For clones, the majority of hits left by one genuine track are shared by several reconstructed tracks. This means the reconstructed tracks have very similar properties so either could be combined with other tracks to construct the same mother particle.

Overlapping candidates occur when tracks are shared between them, but it is not due to clones. For example, a $K^+ / \pi^-$ pair could potentially also be reconstructed as a $K^- / \pi^+$ pair and provide multiple candidates when reconstructing $D^0 \rightarrow K^- \pi^+$ and its charge conjugate decay. A more complex example of this category is when two genuine tracks share hits for some part of their trajectory – for example, it is possible for two tracks to share the same muon track segment.

As the name suggests, genuine multiples are events where the signal decay occurs twice. In signal MC, genuine multiple events are rare but non-negligible. Only one generated signal event is present in each signal MC event, but other signal decays occur in the underlying event at rates proportional to their branching fractions.

It is also possible for multiple candidates to occur even if all daughters are reconstructed correctly, if the selection cuts are too loose. This is especially problematic when associating a slow pion to a $D^0$ semileptonic vertex. Dozens of pions typically originate at the primary vertex during an LHC collision, so if a $D^0$ is correctly reconstructed there may be many candidate slow pions which can form a $D^{*+}$ vertex. Typically this problem is solved by requiring the ‘delta mass’ $(M(\pi D^0) - M(D^0))$ to have a value close to the PDG value of 145 MeV/$c^2$. The value of $M(D^0)$ in a semileptonic decay is not well defined because energy carried by the neutrino is missed, so the delta mass cut is required to be necessarily looser. This will be discussed more in the following chapter.

The presence of genuine multiple candidates in the signal MC can be quantified through several methods. Charm quarks are almost entirely produced in pairs, so if an
event has a genuine $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$, there is a certain probability that the other charm quark produced a similar decay. This effect can be assessed by searching for MC true $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ decays in the very loosely pre-selected MC which are in an event with another true, distinct signal decay. Ten such candidates (0.004\%) were found.

The parameters used by Pythia and EvtGen (through the Gauss interface) can also provide an order-of-magnitude estimate for the number of signal events to expect. Several parameters are assigned values for each set of MC created. The branching fraction $\mathcal{B}(D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu))$ used during the generation phase was 2.26\%. Similarly, the total cross-section $\sigma_{\text{tot.}}$ was assumed to be 91 mb, and the charm cross-section $\sigma(pp \rightarrow c\pi X)$ was assumed to be 5.4 mb, or 5.9\% of the total cross-section.

If we use the transition probability of $D^{*+}$ production from $c$ decays from the PDG [90] of $f(c \rightarrow D^{*+}) \approx 0.224 \pm 0.028$, this yields a value of

$$\frac{\sigma_{\text{sig.}}}{\sigma_{\text{tot.}}} \approx 2 \times \frac{0.0226 \cdot 0.224 \cdot 5.4 \text{mb}}{91 \text{mb}} \approx 0.06\% \quad (6.2.1)$$

where $\sigma_{\text{sig.}}$ is shorthand for $\sigma(pp \rightarrow c\pi \rightarrow D^{*+} X \rightarrow D^0 (K^- \mu^+ \nu_\mu) \pi^+ X)$ and the factor 2 accounts for charge conjugate decays. The transition probability used was measured around the $\Upsilon(4S)$ energy, but is sufficient here for an order-of-magnitude estimate. It is worth noting that since this calculation only accounts for prompt charm decays, it is probably a slight underestimate. Using this estimate, the expected number of signal events in a sample of minimum bias MC is approximately the same was what is observed.

Assuming one genuine $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ decay is present in an event, the probability of the other charm particle hadronising in the same fashion is $f(c \rightarrow D^{*+}) \cdot \mathcal{B}(D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)) \approx 0.05\%$. This corresponds to approximately one event per three million MB events having a pair of charm quarks which hadronise to the signal. With a further reduction due to selection inefficiencies, we treat the presence of two genuine signal decays in an event as a negligible effect. We assume the probability of two $c\pi$ pairs being generated in an event is negligibly small.

Of the 41,740 candidates selected from the signal MC sample, 39,255 (94.0\%) contained true $D^0 \rightarrow K^-\mu^+\nu_\mu$ decays. Of these, 31,790 (80.1\%) were the products of correctly reconstructed $D^{*+} \rightarrow \pi^+ D^0$ decays. Clearly, the stripping line selects $D^0$ mesons with high purity, but the association of an incorrect slow pion causes a reduction in purity when the $D^{*+}$ is reconstructed. This can be seen most clearly in Figures 6.26 and 6.27.

Figure 6.26 shows the number of candidates per event for events in the signal MC sample which pass the stripping cuts. For all events, 13.8\% have more than one candidate per event. When we consider only candidates where a MC true $D^0$ is associated with a false $D^{*+}$ (shown in Figure 6.27), 70.0\% come from events with more than one $D^{*+}$ candidate. Note that the sample in the two plots are not identical since the first includes one entry per event, while Figure 6.27 includes one entry per false candidate.
For events with exactly one $D^{*+}$ candidate, 87.7% were true $D^{*+}$ resonances. This suggests a simple way to improve purity, by only accepting events with a single $D^{*+}$ candidate. However, there are difficulties with this. For example, it is difficult to prove this cut would affect the overall fraction of signal and background in the same way and to understand how it would affect the final $q^2$ distribution.

In particular, the number of candidates per event will depend on the number of primary vertices. Each set of Monte Carlo data is generated at a set value of $\mu$, the number of visible proton-proton interactions per bunch crossing. For instance, the signal MC used in this study was generated with 2.5 interactions per bunch crossing on average.\footnote{This number ($\nu$) is slightly different to $\mu$ because it includes ‘invisible’ interactions which are not seen by the detector.} However, especially with early data, this number has varied substantially. Once the LHC and LHCb reach their design instantaneous luminosities, this number will be more stable. But for early data, it remains a difficult property to simulate accurately.

Typically, once cuts are as tight as possible to reduce the effect of multiple candidates, techniques such as random killing are used, and the effect is assessed as a systematic uncertainty.

The stripping cuts were also tested on samples of 6.05 million generic $b\bar{b}$ and 3.64 million generic $c\bar{c}$ MC10 events. Without any requirements on triggers, the number of candidates found were 2,066 and 512 respectively. Of these candidates, 1,054 (51.0\%) and 348 (68.0\%) were of MC true $D^0$ mesons, while 577 (27.9\%) and 208 (40.6\%) were true $D^{*+}$ decays.

### 6.2.1 Raw Distributions

The raw output of the stripping line can be observed through two key distributions: $\delta_m$ and $M(K\mu)$. The output of right-sign and wrong-sign distributions for signal Monte Carlo can be seen, and these can be compared to those of background events. During the
stripping process, no correction is made for the missing energy in the neutrino, so there are significant differences between the stripping output and the true MC distributions. In particular, after stripping, the $\delta_m$ distribution is simply the $K\mu\pi - K\mu$ invariant mass difference, which is similar to but not identical to the $D^0\pi - D^0$ mass difference.

The $\delta_m$ distribution for MC true signal events is shown in Figure 6.28. The stripping cuts selected 1,038 wrong-sign candidates, and the $\delta_m$ distributions for all right-sign and wrong-sign candidates are shown in Figure 6.29 and Figure 6.30. Here, the error bars are purely statistical.

The stripping cut requires $\delta_m < 200$ MeV/$c^2$, and the minimum possible $\delta_m$ value is at the pion mass (140 MeV/$c^2$). The number of true signal candidates rejected by the cut on $\delta_m$ is small (approximately 3% in signal MC).

Figure 6.28: $\delta_m$ distribution for true $D^{**}$ decays in signal MC.

Figure 6.29: $\delta_m$ distribution for all right-sign candidates in signal MC.

Figure 6.30: $\delta_m$ distribution for all wrong-sign candidates in signal MC.

The wrong-sign $\delta_m$ distribution is relatively flat across the allowed range. The same
plots can be made for all candidates selected from the $b\bar{b}$ and $c\bar{c}$ MC samples. The plots for all non-signal\(^1\) candidates are shown in Figure 6.31 (right-sign) and Figure 6.32 (wrong-sign). The equivalent plots for right-sign and wrong-sign distributions from $c\bar{c}$ MC are shown in Figure 6.33 and Figure 6.34. The effectiveness of the stripping cuts is illustrated by the very low statistics in these plots.

\[^1\text{Determined by requiring D0 BKGCA\(\neq 0.}\]

Similarly, the raw distributions of $M(K\mu)$ in signal MC are shown in Figures 6.35 (for true $D^{*+}$ candidates), 6.36 (all right-sign candidates) and 6.37 (all wrong-sign candidates). Plots for right-sign and wrong-sign $b\bar{b}$ MC are shown in Figures 6.38 and 6.39, while those for $c\bar{c}$ MC are shown in Figures 6.40 and 6.41.

The minimum and maximum physical values for $M(K\mu)$ occur in the extreme cases when the kaon and muon are created at rest relative to each other ($598\text{ MeV}/c^2$), or when the neutrino is created with low energy in the $D^0$ rest frame and the $K\mu$ pair has an invariant mass equal to the $D^0$ mass ($M(D^0) = 1865\text{ MeV}/c^2$). The importance of the $\delta_m$ distribution for separating signal and background can be seen by comparing the $M(K\pi)$ distributions of background and signal. All background distributions with sufficient statistics exhibit peaks in the allowed range, but only the right-sign signal MC exhibits a characteristic peak near 150 MeV/$c^2$ in the $\delta_m$ distribution.
The peaks in $M(K\pi)$ are caused by several factors, but an important one is genuine decays with a missing particle, which was discussed in Section 2.3. The three most important decays in this category are $D^0 \to K^{*-}(K^-\pi^0)\mu^+\nu$ (BR $\sim 1\%$), $D^+ \to K^{*0}(K^-\pi^+}\mu^+\nu$ (BR $\sim 4\%$) and $D_s^+ \to \phi(K^+K^-)\mu^+$ (BR $\sim 1\%$). Each of these has an oppositely-charged $K\mu$ pair in the final state, with an invariant mass near the $D^0$ mass. These decays can also cause mass peaks in the wrong-sign stripping line, for instance if the $K^-$ is missed in the $D_s^+ \to \phi\mu^+$ decay.

Since the reconstruction simply searches for a $K\mu$ pair sharing a vertex and not the absence of other tracks from the same vertex, many of these decays are accepted by the trigger and stripping lines. These effects will be discussed in more detail in the next chapter.

Finally, the MC distributions can be compared to what is seen in data. For offline analysis, we take only data accepted by the dedicated HLT2 line, but the output of the stripping
line for data from all triggers is also useful to ensure the trigger functions correctly. So it is useful to compare the right- and wrong-sign stripping output for all triggers and for $D^0$ mesons which are TOS for the dedicated HLT2 line. The data plotted here are from a sample of 47.8 million magnet down charm microDST events from July 2011. Right-sign events selected by the $D^0 \rightarrow K^- \mu^+ \nu$ stripping line account for approximately 5% of this microDST (2.49 million events). During this period of data taking, a prescale of 0.05 was applied to the right-sign trigger line.

The data plots are shown in Figures 6.42–6.49. Figures 6.42 and 6.43 show the right- and wrong-sign raw output of the stripping line for $\delta_m$. The equivalent plots for $M(K\mu)$ are shown in Figures 6.44 and 6.45. Although these plots include events from any trigger, the $D^{*+}$ peak is still prominent in the right-sign $\delta_m$ distribution.

One notable feature of the $M(K\mu)$ right-sign distribution is the peak just below the $D^0$ mass. This is due to two-body hadronic decays with a mis-identified daughter. The decay with the highest branching fraction is $D^0 \rightarrow K^- \pi^+$, at 4%. When the pion in a $K\pi$ pair is mis-identified as a muon, the small difference in the mass hypothesis ($M(\mu) = 105$ MeV/$c^2$ and $M(\pi^+) = 140$ MeV/$c^2$) causes the invariant mass of the pair to be slightly lower than the true value.
6.2. STRIPPING

Entries 2492538
Mean 166
RMS 17.8

Figure 6.42: $\delta_m$ distribution for all right-sign candidates in the sample of charm microDST data events.

Entries 377635
Mean 174
RMS 16.3

Figure 6.43: $\delta_m$ distribution for all wrong-sign candidates in the sample of charm microDST data events.

Entries 2492542
Mean 1.36e+03
RMS 267

Figure 6.44: $M(K\mu)$ distribution for all right-sign candidates in a sample of charm microDST data events.

Entries 377638
Mean 1.29e+03
RMS 300

Figure 6.45: $M(K\mu)$ distribution for all wrong-sign candidates in a sample of charm microDST data events.

A noticeable peak does not occur until around 1.83 GeV/$c^2$, so a simple way to remove this effect is simply to cut at this value in the $M(K\mu)$ distribution. This rejects almost all mis-identified $D^0 \rightarrow K^-\pi^+$ events, while rejecting 0.05% of true signal events. Alternatively it is possible to reconstruct the $K\mu$ pair under the $K\pi$ hypothesis and reject all events around a tight $D^0$ mass window. This will be discussed in more detail in the next chapter.

Requiring TOS on the HLT2 line reduces statistics considerably, but provides a cleaner signal that is less biased in $q^2$. The right- and wrong-sign $\delta_m$ distributions for TOS data are shown in Figures 6.46 and 6.47. The $M(K\mu)$ plots are shown in Figures 6.48 and 6.49.

By requiring TOS on the dedicated trigger, the triggers specifically selecting $hh$ pairs are no longer considered, and there is a relative reduction the height of the $hh$ peak compared to the bulk of the distribution.

It is useful to examine which triggers fired on events selected by the stripping line. In this sample of right-sign stripped data, 159,678 (6.4%) of the stripped candidate decays were from an event where the dedicated HLT2 line fired. Of these, 154,866 (96.9%) were TOS and 378 (0.2%) were TIS. The remaining events were classified neither TOS nor
CHAPTER 6. TRIGGER AND STRIPPING FOR $D^0 \rightarrow K^- \mu^+ \nu_\mu$

Table 6.9 summarises some of the most frequent triggers. The ‘tight’ $D^0 \rightarrow K^- \mu^+ \nu_\mu$ HLT2 line fires more than three times as often as the prescaled standard line. The fraction of stripped right-sign events passing at least one of the listed triggers is 74.6%. The remaining events are passed in small quantities by many of LHCb’s other triggers (such as the three-hadron charm hadronic triggers).

The triggers from the family Hlt2Topo are so-called topological triggers. These are generic triggers optimised for $B$ decays into final states with muons and charged hadrons [91]. The most common line in the stripping is called Mu3BodyBBDT. The suffix BBDT indicates that a so-called Bagged Bonsai Decision Tree [92] was used in the trigger. Taken together, 43.5% of events are passed by at least one of these triggers.

Candidates passed by the di-hadron charm lines have narrow mass windows of $\pm 50$ MeV/$c^2$ (or $\pm 150$ MeV/$c^2$ for the WideMass versions) around the $D^0$ mass. This means that stripped candidates passed by these triggers lie at the high end of the $M(K\mu)$ distri-
Table 6.9: A list of the most frequently-firing triggers in stripped, right-sign $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ data. There is some overlap between these percentages since more than one trigger can fire for each event.

<table>
<thead>
<tr>
<th>Trigger family</th>
<th>Name</th>
<th>% of stripped events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hlt2CharmHad</td>
<td>D02KK</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td>D02KPi</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>D02KKWideMass</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>D02KPiWideMass</td>
<td>1.6</td>
</tr>
<tr>
<td>Hlt2CharmSemilep</td>
<td>KMuNu</td>
<td>6.4</td>
</tr>
<tr>
<td></td>
<td>PiMuNu</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>KMuNuTight</td>
<td>20.3</td>
</tr>
<tr>
<td>Hlt2Topo</td>
<td>Mu2BodyBBDT</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td>Mu3BodyBBDT</td>
<td>28.2</td>
</tr>
<tr>
<td></td>
<td>Mu4BodyBBDT</td>
<td>22.5</td>
</tr>
</tbody>
</table>

The reduction in the peak near the $D^0$ mass between Figure 6.44 and Figure 6.48 is in part due to rejecting events passing the di-hadron charm triggers.

A third notable feature of Table 6.9 is the relation between the KMuNu and PiMuNu semileptonic charm triggers. Although the only difference between the two is a single mass hypothesis, the prescales work independently, and few events are passed by both triggers. The fraction of events passed by at least one of these triggers is 11.4%.
7.1 Missing momentum

As mentioned in Chapter 2, two methods for managing the missing momentum carried by the neutrino in the decay $D^0 \rightarrow K^- \mu^+ \nu_\mu$ were examined: solving the kinematics to return two allowed solutions for the neutrino momentum (Neutrino Reconstruction) and using MC to correct the $D^0$ momentum based on the reconstructed $K\mu$ mass ($k$-factor).

The $k$-factor has been used by several semileptonic $B$ analyses in LHCb (see for example [27]). The $k$-factor is a correction factor based on a quadratic fit to the ratio of reconstructed $K\mu$ momentum to the true $D^0$ momentum, and is a quadratically increasing function of the reconstructed $K\mu$ mass.

For convenience, we define $p_{\text{rec}}, p_{\text{true}}$ and $p_{\text{corr}}$ as the reconstructed $K\mu$ momentum, the true $D^0$ momentum and the $k$-factor corrected, reconstructed $K\mu$ momentum respectively. An example scatter plot of $p_{\text{rec}}/p_{\text{true}}$ is shown in Figure 7.1, based on a sample of around 32,000 true $D^{*+} \rightarrow \pi^+ D^0 (K^- \mu^+ \nu_\mu)$ decays from the signal MC sample.

For decays where the neutrino carries a small amount of momentum, the reconstructed $M(K\mu)$ approaches the true $D^0$ mass. Figure 7.2 is a profile plot of Figure 7.1, with a quadratic best-fit overlaid. The error bars on the points represent the RMS spread of data in each $k$-factor slice, divided by the square-root of the number of entries in that slice (i.e. $\sigma_i/\sqrt{N_i}$, where $\sigma_i$ and $N_i$ are the RMS of slice $i$ which has $N_i$ entries).

Several points appear outside the bulk of the distribution, with some returning a ratio $p_{\text{rec}}/p_{\text{true}}$ above one. This can occur in cases where a genuine signal event is poorly reconstructed, i.e. the daughters are correctly identified, but the momentum measurements have high uncertainty. In these cases it is possible for the reconstructed $K\mu$ momentum to be higher than the true momentum of the mother $D^0$.

The best-fit curve is forced to pass through 1 at the $D^0$ mass ($M_{D^0} = 1865$ MeV/$c^2$).
The $\chi^2$/NDF of this fit 43.59/30, with a fit function:

$$k(M(K\mu)) = 1 + a(M(K\mu) - M_{D^0}) + b(M(K\mu) - M_{D^0})^2,$$  \hspace{1cm} (7.1.1)

with $a = (6.018 \pm 0.049) \times 10^{-6}$ and $b = (4.059 \pm 0.092) \times 10^{-7}$. Note that at $M(K\mu) = M_{D^0}$, the above expression equals 1 by definition. The fit was only performed in the higher-statistics region of masses above 1.1 GeV/$c^2$.

For a given reconstructed $K\mu$ mass, the $k$-factor can be used as a correction factor to estimate the true $D^0$ momentum. The uncertainty in this correction factor depends on the $K\mu$ mass, and clearly the spread of values is much greater at low mass. Figure 7.3 shows the effect of this correction on reconstructed MC. Here, the relative difference between the uncorrected reconstructed $K\mu$ momentum and true $D^0$ momentum is shown by plotting $(p_{\text{rec.}} - p_{\text{true}})/p_{\text{true}}$. The equivalent plot for the corrected momentum is shown on the right-hand side.

An accurate correction would result in a sharp distribution around zero, perhaps smeared by a Gaussian to account for detector resolution effects. The corrected distribution is clearly not well represented by a simple Gaussian, and the distribution has a mean or $(-1.75 \pm 0.09) \times 10^{-2}$ and an RMS of $(1.639 \pm 0.007) \times 10^{-1}$.

The second method for accounting for the missing momentum is neutrino reconstruction. Recall from (2.3.8) that by decomposing the particle momenta into components parallel and perpendicular to the $D^0$ flight direction, the kinematics of the decay...
7.1. MISSING MOMENTUM

Figure 7.2: A profile plot of Figure 7.1, with a quadratic best-fit curve forced to pass through 1 at the $D^0$ mass.

Figure 7.3: The effect of a $k$-factor momentum correction. The difference between the uncorrected and true $D^0$ momentum distributions (left) and between the $k$-factor corrected and true $D^0$ momentum distributions (right).
$D^0 \rightarrow K^- \mu^+ \nu_\mu$ can be solved to give two quadratic solutions for the neutrino momentum:

$$|p_\nu^\parallel| = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad (7.1.2)$$

where

$$a = 4(p_{\mu K}^\perp + m_{\mu K}^2) \quad (7.1.3)$$
$$b = 4|p_{\mu K}^\parallel|(2p_{\mu K}^\perp - (m_D^2 - m_{\mu K}^2)^2) \quad (7.1.4)$$
$$c = 4p_{\mu K}^\perp(2p_{\mu K}^\parallel + m_{\mu K}^2) - (m_D^2 - m_{\mu K}^2)^2. \quad (7.1.5)$$

A choice remains between the two solutions. In some situations one or both solutions can be rejected immediately. For instance, it is possible for $b^2$ to be smaller than $4ac$, which is non-physical and leads to complex solutions to (7.1.2). Typically $b^2$ and $4ac$ are numerically close, so the effects of finite detector resolution can cause complex solutions to occur for genuine signal events. For signal MC passing the stripping cuts, approximately 44% of true $D^0$ decays have a complex solution to (7.1.2).

Similarly, for background, complex solutions can arise. For this class, we do not necessarily expect physical solutions, since the underlying physics is not a genuine $D^0$ decay (although it may closely resemble one). For background $b\bar{b}$ MC, approximately 55% of the few events remaining after stripping have complex solutions.

Additionally, any solution with an unphysically high neutrino momentum can be immediately rejected.

There are several possibilities to choose between the remaining solutions. In the LHCb analysis of $B_s \rightarrow D_s^+ \mu \nu$ [27], no method was found which was more successful than a random choice, so the smaller solution was always chosen. When chosen correctly, the smaller solution leads to lower associated uncertainty, since less momentum was taken by the neutrino and more was measurable in the visible charged tracks. For signal events with real solutions, this method selects the correct neutrino momentum around 50% of the time.

In the case of $D^{*+}$ tagged $D^0$ decays, the mass difference ($\delta_m$) between the reconstructed $D^{*+}$ and $D^0$ can be employed as an additional constraint on the decay kinematics. When all decay products are reconstructed (for example when the $D^0$ decays to two charged hadrons), the distribution of $\delta_m$ has a characteristic shape. The distribution exhibits a narrow peak around 145 MeV/$c^2$, as shown in Figure 7.4, taken from [93].

The PDG value for the $D^*$ mass is $M(D^{*\pm}) = 2010.25 \pm 0.14$ MeV/$c^2$. For real solutions (after rejecting the small number of events for which the MC true solution is complex), choosing the neutrino momentum solution which yields a mass closest to this value selects the correct value 83.4% of the time, which is clearly preferable to random choice.
To verify the neutrino reconstruction method, the same algorithm was run over MC true values and reconstructed values for the $D^0$ flight vector and muon/kaon momenta. For reconstructed candidates, the $D^0$ decay vertex is clearly defined as a displaced $K\mu$ vertex. But to determine the flight direction, it is also necessary to define the $D^0$ origin vertex.

For this calculation, we take the ‘best’ primary vertex (PV) as the creation vertex. As in Chapter 6 the ‘best’ PV is defined as the PV with which the reconstructed $D^0$ has the lowest IP $\chi^2$. Clearly this assumption is most appropriate for $D^0$ decays originating from the PV in events with a single PV. In the case of events with multiple primary vertices or secondary $D^0$ decays, this assumption is correct less frequently.

In LHCb, approximately 9% of hadronic charm decays seen are from secondary charm decays [94]. These decays occur when a charmed meson is not produced at the primary vertex, but from the decay of another particle, typically a $b$ meson. The most common variable used to separate prompt and secondary charm is the mother IP $\chi^2$ [95].

To evaluate the effectiveness of using the best primary vertex in the flight vector calculation, we can compare the reconstructed flight direction to the MC true flight direction as a function of the number of primary vertices. MC10 Monte Carlo is generated with 2.5 proton-proton interactions on average per beam crossing (this number is given the symbol $\nu$). For data, LHCb is only able to measure the average number of visible interactions per beam crossing ($\mu$), which varies by fill. For data taken in July 2011, $\mu$ was typically close to 1.4.

Figure 7.5 shows how the values of $\nu$ and $\mu$ affect the number of primary vertices found for triggered and stripped data and for stripped signal MC. In data, 80% of candidates come from events with either one or two primary vertices. In signal MC, this number is 63%.

The angle between the MC true and reconstructed $D^0$ flight vector (taking the best PV as the origin) is shown in Figure 7.6 for candidates from events with one or four primary
Figure 7.5: The number of primary vertices in signal MC (left) and stripped data (right). The MC events are generated with $\nu = 2.5$, while the data typically have $\mu = 1.4$.

Figure 7.6: Angle between reconstructed $D^0$ flight vector and the true flight vector (in radians) for candidates in events with one primary vertex (left) and with four primary vertices (right).
vertices. Although there are significantly fewer candidates from events with four primary vertices, there is no significant difference between the shapes of the two distributions. In both case, the distribution peak is close to 1.5 mrad. For this reason, we assume the ‘best’ PV position is a suitable choice as the origin of the $D^0$, independent of the number of primary vertices in the event.

Figure 7.7 shows the distribution of the true solutions for the two choices of reconstructed solution. After reconstructing the $D^0$ momentum, the distribution for correct choices is very close to a delta function at the correct momentum value, while the distribution for incorrect choices is spread widely around zero.

From (7.1.2), the difference between the incorrect and true solution for the neutrino momentum is always $\pm \sqrt{b^2 - 4ac/a}$, which is momentum-dependent. The momentum dependence can be seen in Figure 7.7 (left), where a tail is evident for incorrect solutions above the true $D^0$ momentum value.

![Figure 7.7: The effect of the neutrino reconstruction method for MC true quadratic solutions. The difference between the MC true $D^0$ momentum and the incorrectly chosen (left) and correctly chosen (right) MC true quadratic solutions are shown, rescaled by the true momentum.](image)

The equivalent plot to Figure 7.7 but for (real) reconstructed solutions for incorrect and correct choice of MC true solution is shown in Figure 7.8. Figure 7.9 shows the chosen and rejected real, reconstructed solutions. The key comparison is between the $D^0$ momentum distributions under the two methods ($k$-factor and neutrino reconstruction). This comparison (with the x-axes rescaled to be the same) is shown in Figure 7.10.

Neither the $k$-factor distribution nor the reconstruction method distribution are well-modelled by Gaussian distributions.
Figure 7.8: The effect of the neutrino reconstruction method for reconstructed quadratic solutions. The difference between the MC true $D^0$ momentum and the incorrectly chosen (left) and correctly chosen (right) reconstructed quadratic solutions (from the MC truth correct choice) are shown, rescaled by the true momentum.

Figure 7.9: The effect of the neutrino reconstruction method for rejected (left) and chosen (right) reconstructed quadratic solutions. The difference between the MC true $D^0$ momentum and each solution is shown, rescaled by the true momentum.
Figure 7.10: Comparison of the $D^0$ distributions between the k-factor and neutrino reconstruction methods. The k-factor method (left) necessarily results in larger statistics compared to the neutrino reconstruction method (right). The reconstruction method distribution is clearly more symmetrical around zero.

From Figure 7.10, it is clear that although neutrino reconstruction results in lower statistics due to rejected complex solutions, the method enables $D^0$ momenta to be determined with higher precision. There is contamination from the approximately 17% of events for which the incorrect root is chosen. From Figure 7.8, we can expect that the $D^0$ momentum value from these incorrect events will have a non-Gaussian distribution around the true $D^0$ momentum.

7.2 Reconstructed Monte Carlo

To estimate the relative levels of signal and background (S/B) in data, it is necessary to determine the distributions of signal and background Monte Carlo under the neutrino reconstruction method. There are various of classes of background to consider: generic, combinatoric, and peaking backgrounds.

LHCb simulates a large number of ‘generic’ MC events. These are samples of simulated $c\bar{c}$, $b\bar{b}$ and minimum bias events generated in similar proportions to what one would expect in data. These MC samples can be used to estimate the continuum contribution beneath a particular signal. Backgrounds in this class are referred to as generic backgrounds. Backgrounds due to incorrect particle identification are included in this category.

The combinatoric backgrounds are candidate events which simulate the signal, but in reality are random combinations of tracks from several decays or primary vertices. It is
reasonable to assume that the probability of two random tracks passing the selection is independent of the track charge, and the primary purpose of the wrong-sign trigger and stripping lines is to assess this background.

The third class of background to consider is the ‘peaking’ category. As mentioned in Section 2.3 there are several decays which have a relatively high branching ratio and a very similar final state to $D^{*+} \rightarrow \pi^+ D^0(K^+ \mu^+ \nu_\mu)$. To estimate the effect of these backgrounds, several samples (each of two million events) were generated for the most important of these:

- $D^{*+} \rightarrow \pi^+_s D^0 \rightarrow \pi^+_s K^{*-} \mu^+ \nu_\mu$, where $K^{*-} \rightarrow K^0 \pi^-$ or $K^- \pi^0$,
- $D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$, where $K^{*0} \rightarrow K^- \pi^+$ is forced, and
- $D_{s}^+ \rightarrow \phi \mu^+ \nu_\mu$, where $\phi \rightarrow K^+ K^-$ is forced.

Each of these decays can potentially mimic the signal, since each final state includes a $K \mu$ pair originating from a displaced vertex with missing energy carried by a neutrino. The lifetimes of the three mother particles ($D^0$, $D^+$ and $D^{*+}$) are all approximately 500 fs, so for each, we expect the final decay vertex to be approximately the same distance from the primary vertex on average.

A final class of MC background is considered in the ‘peaking’ category, namely fake $D^*$ decays. This is a sample of MC true signal $D^0$ decays which were combined with an incorrect slow pion to wrongly reconstruct a $D^*$. There is some overlap between the $D^0$ candidates in this sample and the signal sample from cases where a true $D^0$ decay is used to make multiple $D^*$ candidates.

The efficiency of each type of background to pass the selection requirements can be evaluated by reproducing the stripping cuts on MC, and counting the number of events that pass the neutrino reconstruction method. At this stage, no requirements are made on triggers. With the exception of minimum bias, each sample was made of approximately equal proportions of simulated data with each magnet polarity\(^1\).

As before, genuine signal events are rejected from the background samples. The numbers of events in the MC samples and the numbers retained after stripping and neutrino reconstruction are shown in Table 7.1.

The $\delta_m$ distributions after neutrino reconstruction for each of these MC types for wrong-sign (when appropriate) and right-sign combinations are shown in Figures 7.11, 7.12, 7.13 and 7.14.

Several characteristics of these distributions are notable. The MC true signal distribution has a much narrower width than the non-reconstructed distribution, comparable to the typical widths seen in $D^*$-tagged $D^0 \rightarrow hh$ decays shown in Figure 7.4. It is clear that

\(^1\)The minimum bias sample was composed of a very large sample of magnet down MC.
Table 7.1: Efficiencies in MC after stripping and neutrino reconstruction, with no requirements on trigger.

<table>
<thead>
<tr>
<th>MC sample</th>
<th># Generated</th>
<th>RS selected (Eff.)</th>
<th>WS selected (Eff.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>4.0 million</td>
<td>30,648 (0.8%)</td>
<td>985 (0.02%)</td>
</tr>
<tr>
<td>Fake $D^*$</td>
<td>7,280 (0.2%)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MB</td>
<td>69.2 million</td>
<td>153 ($2 \times 10^{-6}$)</td>
<td>65 ($9 \times 10^{-7}$)</td>
</tr>
<tr>
<td>$b\bar{b}$</td>
<td>6.1 million</td>
<td>982 ($2 \times 10^{-4}$)</td>
<td>687 ($1 \times 10^{-4}$)</td>
</tr>
<tr>
<td>$c\bar{c}$</td>
<td>3.6 million</td>
<td>156 ($4 \times 10^{-5}$)</td>
<td>70 ($2 \times 10^{-5}$)</td>
</tr>
</tbody>
</table>

$D^0 \rightarrow K^{*+} \mu^+ \nu_\mu$

$D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$

$D^+_s \rightarrow \phi \mu^+ \nu_\mu$

Figure 7.11: Distribution of $\delta_m$ after neutrino reconstruction for signal (left) and fake $D^*$ (right) MC. Error bars are purely statistical.

...a tight cut around the signal peak in $\delta_m$ could reject a large proportion of backgrounds, with minimal loss of signal...

The MC samples all correspond to different luminosities, but nevertheless from the distributions one can expect the largest backgrounds in data to be due to true $D^0$ decays associated with a random pion (Figure 7.11) and from the two peaking backgrounds $D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$ and $D^{*+} \rightarrow \pi^+_s D^0 \rightarrow \pi^+_s K^{*-} \mu^+ \nu_\mu$ (Figure 7.12). The contributions from the generic samples are small. The efficiencies, especially under the sharp signal peak, are low and the wrong- and right-sign distributions are similar. For this reason, one can expect only a small right-sign excess in the signal region from these MC samples after a subtraction of wrong-sign events.
Figure 7.12: Distribution of $\delta_m$ after neutrino reconstruction for MC of the peaking backgrounds $D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$ (left) and $D^{*+} \rightarrow \pi^+ D^0 \rightarrow \pi^+ K^{*-} \mu^+ \nu_\mu$ (right). Right-sign candidates are marked in blue and wrong-sign candidates in red. Error bars are purely statistical.

Figure 7.13: Distribution of $\delta_m$ after neutrino reconstruction for MC of the peaking background $D^+ \rightarrow \phi \mu^+ \nu_\mu$ (left) and minimum bias (right). Right-sign candidates are marked in blue and wrong-sign candidates in red. Error bars are purely statistical.
The peaking background $D_s^+ \rightarrow \phi \mu^+ \nu_{\mu}$ (Figure 7.13) is not as significant as the other two. The presence of muons of both charges in the final state means that almost all decays in this mode are removed by a subtraction of wrong-sign events.

![Figure 7.14: Distribution of $\delta_{m}$ after neutrino reconstruction for $b\bar{b}$ (left) and $c\bar{c}$ (right) MC. Right-sign candidates are marked in blue and wrong-sign candidates in red. Error bars are purely statistical.](image)

### 7.3 Data analysis

The data used for the remainder of this thesis were recorded in July and August 2011 and consist of all Charm microDST events for this time period, under Stripping 15\(^1\) (runs 96214 to 98656). During this period the LHCb magnet had the ‘up’ polarity (MagUp) between 16 July and 26 July, and the ‘down’ polarity (MagDown) between 27 July and 13 August. The total number of events in the microDST under each polarity were 63.1 million (MagUp) and 168.1 million (MagDown). During this period, the protons in each beam had an energy of 3.5 GeV.

The number of stripped candidates was 3.3 million RS and 490,000 WS under MagUp and 8.8 million RS and 1.3 million WS under MagDown. This resulted in a final sample of TOS candidates (on the dedicated HLT2 signal lines) of 708,266 RS and 17,083 WS candidates. The number of candidates in each case passing the neutrino reconstruction was 684,291 (97\%) and 16,325 (96\%). This cannot be directly compared to the 44\% of

---

\(^1\)Brunel v40r0, DaVinci v28r4.
signal MC events which pass the stripping cuts and have real solutions. The requirement on HLT2 TOS for data candidates causes the fraction of candidates with real solutions to increase substantially. With no HLT2 TOS requirement, the fraction of stripped data events (which can come through any HLT2 line) with real solutions is 56%.

There is a factor of five between the right-sign and wrong-sign trigger prescales, so the wrong-sign distribution is multiplied by this factor to make the samples equivalent.

![Figure 7.15: Distribution of $\delta_m$ after neutrino reconstruction for right-sign and wrong-sign candidates in data (left) and the wrong-sign subtracted distribution (right). Right-sign candidates are marked in blue and wrong-sign candidates in red. Error bars are purely statistical.](image)

The $\delta_m$ distributions in data are shown in Figure 7.15, with the wrong-sign subtracted distribution on the right. In this sample, there is still a small contamination from $D^0 \rightarrow hh$ decays, as can be seen in the small peak near the $D^0$ mass in Figure 7.16. These plots shows the $K\mu$ mass distribution for those data events passing the neutrino reconstruction.

### 7.3.1 Background reduction

If a significant proportion of backgrounds in data are due to fake $D^*$ decays, additional cuts to increase $D^0$ purity may not significantly improve the signal to background ratio. Instead, to reduce backgrounds it is necessary to make further requirements on the slow pion and the $D^*$ vertex.

After optimisation with CROP, the five quantities which offer the most separation power are DOCA of the $D^*$ vertex and the transverse momentum, DLL($K - \pi$) and
Figure 7.16: Distribution of $M(K\mu)$ after neutrino reconstruction for right-sign and wrong-sign candidates in data (left) and the wrong-sign subtracted distribution (right). Right-sign candidates are marked in blue and wrong-sign candidates in red. Error bars are purely statistical.

DLL($e - \pi$) of the slow pion. The distributions of these for signal and fake $D^*$ MC\(^1\) are shown in Figure 7.17 and Figure 7.18. In these plots, the trigger and stripping cuts have been applied to the MC.

A notable feature of Figure 7.18 is the sharp peak at zero for the fake DLL($K - \pi$) distribution. The PID value is set to zero when no particle is detected by the PID system. This peak at zero indicates the presence of ghost tracks – for the fake $D^*$ sample, the $D^0$ decay is required to be MC true, but no requirements are made on the $D^*$ decay. In the signal sample, the slow pion track must be associated with a genuine MC particle. The DOCA is defined as the closest distance between the momentum vector of the $K\mu$ system and the slow pion.

A CROP optimisation on these four quantities returns the values shown in Table 7.2. The transverse momentum cut of 270 MeV/c is equivalent to a cut at 2.43 on the $\log_{10}(p_T)$ plot of Figure 7.17.

When applied to MC candidates which pass neutrino reconstruction, these cuts have efficiencies of 80.8% for signal candidates and 50.2% for fake $D^*$ candidates. To make a final selection of signal candidates a tight cut is placed on the $\delta_m$ distribution, $|\delta_m - (M_{D^*} - M(D^0))| < 3$ MeV/c\(^2\). To reject $D^0 \to hh$ decays with a hadron mis-identified as a muon we require $M(K\mu) < 1.84$ GeV/c\(^2\).

\(^1\)Each candidate in the fake $D^*$ MC is a true $D^0$ decay associated with a non-signal pion.
Figure 7.17: Signal (black) and fake $D^*$ (red) distributions, after trigger and stripping, for the $K\mu$ and $\pi_s$ distance of closest approach (left) and the logarithm of pion $p_T$ (right).

Figure 7.18: Signal (black) and fake $D^*$ (red) distributions, after trigger and stripping, for pion DLL($K-\pi$) (left) and pion DLL($e-\pi$) (right).
Table 7.2: List of cuts used for offline $\pi_s$ selection.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Cut type</th>
<th>Cut Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pion transverse momentum</td>
<td>$p_T^\pi$</td>
<td>$&gt;$</td>
<td>270 MeV/c</td>
</tr>
<tr>
<td>Distance of closest approach ($D^0 - \pi$)</td>
<td>DOCA</td>
<td>$&lt;$</td>
<td>0.16 mm</td>
</tr>
<tr>
<td>Pion PID $DLL(K - \pi)$</td>
<td>$\pi_{PIDK}$</td>
<td>$&lt;$</td>
<td>5</td>
</tr>
<tr>
<td>Pion PID $DLL(e - \pi)$</td>
<td>$\pi_{PIDe}$</td>
<td>$&lt;$</td>
<td>1</td>
</tr>
</tbody>
</table>

### 7.3.2 Fit to data and MC

The ratio of signal to background in data can be calculated. The peak of the signal MC distribution of $\delta_m$ is well-modelled by the sum of a core Gaussian and a bifurcated Gaussian to account for the asymmetric tail. However, this distribution appears narrower in data (see Figure 7.22), so a second narrower Gaussian distribution is added at the $D^*$ mean to account for this difference. This is pictured in Figure 7.19. In this section, all MC has the stripping cuts and trigger (TCK 0x002e002a) applied. The trigger requirements are the same as in data, i.e. either L0Hadron or L0Muon must fire, and TOS on Hlt1TrackMuon is required for the muon.

![Figure 7.19: The sum (blue) of two core Gaussian distributions (red) and a bifurcated Gaussian (green) fitted to the $\delta_m$ MC signal peak over the range $[140, 153]$ GeV/c$^2$. The mean of the core Gaussians are fixed to the PDG $\delta_m$ value, while all other parameters are free during the fit procedure.](image)
The mean of the core Gaussians were set equal to PDG $\delta_m$ value (145.4 MeV/$c^2$) and an un-binned Rooit fit over the peak (140 to 152 MeV/$c^2$) resulted in the values: $\mu_b = 145.25 \pm 0.12$ MeV/$c^2$, $\sigma_{b1} = 2.40 \pm 0.10$ MeV/$c^2$, $\sigma_{b2} = 3.67 \pm 0.13$ MeV/$c^2$, $\sigma_{\text{core}} = 0.98 \pm 0.08$ MeV/$c^2$, and $\sigma_{\text{core2}} = 0.10 \pm 0.05$ MeV/$c^2$ where $\mu_b$, $\sigma_{b1}$, and $\sigma_{b2}$ represent the mean, left-hand width and right-hand width of the bifurcated Gaussian, and $\sigma_{\text{core}}$ and $\sigma_{\text{core2}}$ represent the widths of the core Gaussians. Of the core distribution yield, 98.2 $\pm$ 1.3% is from the wider Gaussian, and 24.2 $\pm$ 2.6% of the total yield is from the core sum.

When fitted against signal MC, the second core Gaussian has an almost negligible effect, but the extra free parameter is necessary to model the small differences between MC and data.

The physics underlying the most significant backgrounds have certain similarities. From what was shown in the previous section, the most significant backgrounds are due to (i) combinatorics, (ii) the peaking backgrounds $D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$ (associated with a random pion), $D^{*+} \rightarrow \pi^+_s D^0 \rightarrow \pi^+_s K^{*-} \mu^+ \nu_\mu$ (associated with the correct slow pion) and (iii) with genuine $D^0 \rightarrow K^- \mu^+ \nu_\mu$ decays associated with a random pion.

After neutrino reconstruction, the surviving RS candidates from categories ii and iii have similar distributions and can be modelled by the same PDF. The $\delta_m$ distributions over the range $[140, 200]$ MeV/$c^2$ are shown in Figure 7.20 and Figure 7.21. In all three cases, an unbinned fit is made using the standard Rooit library PDF named $\text{RooDstD0BG}$ [97]. This PDF was developed to model the background $\delta_m$ distributions of $D^0$ decays matched to an incorrect slow pion.

The $\text{RooDstD0BG}$ is an empirical function with four parameters: a mass threshold $\delta_{m0}$, and three shape parameters $a$, $b$ and $c$. The parameters $a$ and $b$ are dimensionless and $c$ has units of MeV/$c^2$. The mass threshold is set to be the pion mass, $\delta_{m0} = 139.57$ MeV/$c^2$. The PDF takes the form:

$$f(\delta_m) = \left(1 - \exp\left(-\frac{(\delta_m - \delta_{m0})}{c}\right)\right) \left(\frac{\delta_m}{\delta_{m0}}\right)^a + b \left(\frac{\delta_m}{\delta_{m0}} - 1\right),$$

for $\delta_m > \delta_{m0}$.

The results of the unbinned fits for the three backgrounds are given in Table 7.3. The low statistics result in large uncertainties on the fitted values, especially for the shape parameter $a$. For the peaking background $D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$, the statistics are very low and although the $\text{RooDstD0BG}$ model fits the data adequately, it could equally be fitted with a straight line.

The combinatoric background contribution is compensated for in data by subtracting all WS candidates from the RS distributions. Figure 7.22 shows the result of a binned fit of a combination of the signal PDF and a single background $\text{RooDstD0BG}$ PDF to wrong-sign subtracted data. In this fit all parameters except the core signal mean were

152
7.3. DATA ANALYSIS

Figure 7.20: Distribution of $\delta_m$ for the peaking backgrounds $D^+ \rightarrow \bar{K}^{*0} \mu^+ \nu_\mu$ (left) and $D^{*+} \rightarrow \pi^+_s D^0 \rightarrow \pi^+_s K^- \mu^+ \nu_\mu$ (right) after offline selection. An unbinned fit with the RooDstD0BG PDF is made in each case and superimposed. The resulting fitted parameters are given in Table 7.3.

free in the fit.

The fit in Figure 7.22, with eleven free parameters, matches the data well and has a $\chi^2$ value of 539 over 625 bins in the fitted region. We define the signal region as $|\delta_m - (M_{D^*} - M_{D^0})| < 3 \text{ MeV}/c^2$. Signal purity can be determined by a numerical integration of the normalised PDFs. The total number of signal and background events are 79,477 and 216,965 respectively over the full range shown in Figure 7.22, and in the signal region the numbers are 37,497 and 24,210. Thus, in the signal region we expect a signal purity of 60.8%.

7.3.3 Measurement of $q^2$ in data

Finally, the $q^2$ distribution in data can be measured. The $q^2$ distributions are divided into eight bins of equal width between 0 and 1.6 GeV$^2$/c$^2$, and one additional bin for all candidates with $q^2 \geq 1.6$ GeV$^2$/c$^2$. In this section, a ‘candidate’ is one that has passed
CHAPTER 7. DATA ANALYSIS

Figure 7.21: Distribution of $\delta_m$ for the fake $D^*$ background after offline selection. An unbinned fit with the RooDstD0BG PDF is superimposed. The resulting fitted parameters are given in Table 7.3.

Table 7.3: Fitted parameters for the RooDstD0BG PDF for the two most significant peaking backgrounds and for fake $D^*$ decays (i.e. genuine $D^0$ decays associated with a random pion). The results are from an unbinned fit over the range $[140, 200]$ MeV/c$^2$.

<table>
<thead>
<tr>
<th>Type</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$ (MeV/c$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^+ \rightarrow K^{*0} \mu^+ \nu_\mu$</td>
<td>$0.5 \pm 1.2$</td>
<td>$-1.6 \pm 1.3$</td>
<td>$3.6 \pm 1.1$</td>
</tr>
<tr>
<td>$D^{<em>+} \rightarrow \pi_\delta^+ D^0 \rightarrow \pi_\delta^+ K^</em>^- \mu^+ \nu_\mu$</td>
<td>$0.5 \pm 0.4$</td>
<td>$-2.6 \pm 0.4$</td>
<td>$1.3 \pm 0.2$</td>
</tr>
<tr>
<td>Fake $D^*$</td>
<td>$0.5 \pm 0.6$</td>
<td>$-2.2 \pm 0.6$</td>
<td>$2.9 \pm 0.3$</td>
</tr>
</tbody>
</table>

stripping and the offline cuts and satisfies $|\delta_m - (M_{D^*} - M_{D^0})| < 3$ MeV/c$^2$.

The efficiency matrix in signal MC for candidates of a certain reconstructed $q^2$ to be in the same bin as the true value is given in Tables 7.4 and 7.5. The statistical uncertainty on each efficiency value is taken to be binomial. By definition the sum of each row (across the two tables) is 100%, since the candidates are required to be MC true signal, so have a well-defined true $q^2$ value.

The comparison of $q^2$ distributions for generated and reconstructed signal MC is shown in Figure 7.23. In this figure, three distributions of $q^2$ in signal MC are shown: generator-level, MC true values for candidates passing offline selection and reconstructed
Figure 7.22: Distribution of $\delta_m$ for data. A binned fit of the sum of a RooDstD0BG PDF (green) and the signal model (red) is superimposed. All parameters except the core signal Gaussian means were free in the fit. The right-hand plot is a zoom of the signal peak.

values of the same candidates. The generator-level distribution is normalised to have the same integral as the other two distributions. Multiplication by the inverse of the efficiency matrix given in Tables 7.4 and 7.5 transforms the MC true distribution (blue) to the reconstructed distribution (black).

Similarly, the $q^2$ distributions for the three major backgrounds ($D^+ \rightarrow \bar{K}^{*0} \mu^+\nu_\mu$, $D^{*+} \rightarrow \pi_s^+ D^0 \rightarrow \pi_s^+ K^{*-} \mu^+\nu_\mu$ and fake $D^*$) are given shown in Figure 7.24 and Figure 7.25. The physics underlying the category of fake $D^*$ candidates is similar to signal, since the only difference is an incorrect slow pion. For this reason, the candidates which lie close enough in phase space to a genuine signal decay exhibit a $q^2$ distribution similar to that of the signal.

By the random choice inherent in the neutrino reconstruction method for non-signal candidates, approximately 50% of the fake $D^*$ candidates were assigned the correct polynomial root. The additional application of offline cuts results in a $q^2$ distribution (Figure 7.25 similar to the distribution in signal MC (Figure 7.23).
Figure 7.23: Comparison of $q^2$ distributions in nine bins for generated and reconstructed signal MC. The three colours represent generator level-MC truth (red), MC true values (blue) and reconstructed values (black) for candidates passing offline selection. The uncertainties are purely statistical.

Figure 7.24: Reconstructed $q^2$ distributions in nine bins for reconstructed $D^+ \rightarrow \bar{K}^{*0}\mu^+\nu_\mu$ candidates (left) and $D^{*+} \rightarrow \pi^+D^0 \rightarrow \pi^+K^{*+}\mu^+\nu_\mu$ candidates (right). The uncertainties are purely statistical.
Table 7.4: Efficiency matrix $\epsilon_{ij}$ for signal MC candidates, for true $q^2$ values up to 1.0 GeV$^2$/$c^2$. The column gives the true $q^2$ bin $j$, and the row gives the reconstructed $q^2$ value $i$. The statistical uncertainty in the least significant digit $s$ is given in parentheses.

<table>
<thead>
<tr>
<th>True $q^2$ value (GeV$^2$/$c^2$)</th>
<th>Rec $q^2$ (GeV$^2$/$c^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 0.2)</td>
<td>[0.2, 0.4)</td>
</tr>
<tr>
<td>[0, 0.2)</td>
<td>0.724(20)</td>
</tr>
<tr>
<td>[0.2, 0.4)</td>
<td>0.291(17)</td>
</tr>
<tr>
<td>[0.4, 0.6)</td>
<td>0.078(10)</td>
</tr>
<tr>
<td>[0.6, 0.8)</td>
<td>0.022(6)</td>
</tr>
<tr>
<td>[0.8, 1.0)</td>
<td>0.002(2)</td>
</tr>
<tr>
<td>[1.0, 1.2)</td>
<td>0.003(3)</td>
</tr>
<tr>
<td>[1.2, 1.4)</td>
<td>0.005(4)</td>
</tr>
<tr>
<td>[1.4, 1.6)</td>
<td>0.009(8)</td>
</tr>
<tr>
<td>[1.6, ∞)</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 7.5: Efficiency matrix $\epsilon_{ij}$ for signal MC candidates, for true $q^2$ values above 1.0 GeV$^2$/$c^2$. The column gives the true $q^2$ bin $j$, and the row gives the reconstructed $q^2$ value $i$. The statistical uncertainty in the least significant digits is given in parentheses.

<table>
<thead>
<tr>
<th>True $q^2$ value (GeV$^2$/$c^2$)</th>
<th>Rec $q^2$ (GeV$^2$/$c^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1, 1.2)</td>
<td>[1.2, 1.4)</td>
</tr>
<tr>
<td>[0, 0.2)</td>
<td>–</td>
</tr>
<tr>
<td>[0.2, 0.4)</td>
<td>0.005(3)</td>
</tr>
<tr>
<td>[0.4, 0.6)</td>
<td>0.008(3)</td>
</tr>
<tr>
<td>[0.6, 0.8)</td>
<td>0.032(8)</td>
</tr>
<tr>
<td>[0.8, 1.0)</td>
<td>0.166(18)</td>
</tr>
<tr>
<td>[1.0, 1.2)</td>
<td>0.545(29)</td>
</tr>
<tr>
<td>[1.2, 1.4)</td>
<td>0.194(27)</td>
</tr>
<tr>
<td>[1.4, 1.6)</td>
<td>0.044(19)</td>
</tr>
<tr>
<td>[1.6, ∞)</td>
<td>0.020(20)</td>
</tr>
</tbody>
</table>

The RS and WS $q^2$ distributions in data are shown in Figure 7.26. The offline selection, in particular the tight $\delta_m$ cut means the difference in statistics between RS and WS candidates is more pronounced than before offline selection. The WS distribution is flat across the $q^2$ distribution within statistical uncertainties, with the exception of the first bin.

The time scale and scope of this thesis does not allow for a detailed analysis of backgrounds and systematic uncertainties. These are two areas of potential interest for further studies of this channel. We conclude this thesis by constructing a simple background model to determine the $D^0 \rightarrow K^- \mu^+ \nu_\mu$ form factors in data.
CHAPTER 7. DATA ANALYSIS

Figure 7.25: Reconstructed $q^2$ distributions in nine bins for reconstructed fake $D^*$ candidates. The uncertainties are purely statistical.

Figure 7.26: Reconstructed $q^2$ distributions in nine bins for reconstructed RS (black) and WS (red) candidates in data. The uncertainties are purely statistical.

We make the assumptions that the only non-negligible backgrounds under the data $q^2$ distribution are those shown in Figures 7.25 and 7.26, that the cross-section for $D^+$ and $D^0$ production in LHCb are equal, and that all $D^0$ candidates in the selected sample originate from $D^*$ decays. The three background samples are added in ratios based on an estimation of their luminosity. We take $N_i = L\sigma\epsilon_iB_i$ to be the total number of events in each mode $i$ remaining after reconstruction, where $\epsilon_i$ represents the total reconstruction efficiency of mode $i$, $\sigma$ is the cross-section for $D^+$ or $D^0$ production and $B_i$ is the branching fraction of each background mode.

A comparison of the number of fake (175) and true (3,950) $D^*$ candidates in signal
MC permits an estimate of the branching fraction for fake $D^*$ candidates at 4.4% of the $D^0 \rightarrow K^- \mu^+ \nu_\mu$ branching fraction. The estimated product $\mathcal{L}\sigma\epsilon_i$ for each MC sample is shown in Table 7.6. Based on these numbers, we construct a background model containing 93.5% fake $D^*$ candidates, 5.8% $D^0 \rightarrow K^{*-} \mu^+ \nu_\mu$ candidates and 0.7% $D^+ \rightarrow \overline{K}^{*0} \mu^+ \nu_\mu$ candidates. This ‘cocktail’ is subtracted from wrong-sign subtracted data in the ratio determined in Section 7.3.2.

<table>
<thead>
<tr>
<th>MC sample</th>
<th>Branching fraction</th>
<th># candidates</th>
<th>$N_i/B_i = \mathcal{L}\sigma\epsilon_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fake $D^*$</td>
<td>4.4% × 3.31%</td>
<td>175</td>
<td>1.2 × 10^5</td>
</tr>
<tr>
<td>$D^0 \rightarrow K^{*-} \mu^+ \nu_\mu$</td>
<td>1.98%</td>
<td>147</td>
<td>7.4 × 10^3</td>
</tr>
<tr>
<td>$D^+ \rightarrow \overline{K}^{*0} \mu^+ \nu_\mu$</td>
<td>3.7%</td>
<td>34</td>
<td>9.2 × 10^2</td>
</tr>
</tbody>
</table>

Table 7.6: Luminosity estimates of the three main background MC samples.

To estimate the uncertainties in this simple background model, the fractions of each mode in the cocktail are changed by ±20%. The fractions of the other two modes are adjusted in proportion to keep the total integral unchanged. We take the uncertainty on each bin of the background cocktail to be half of the variation in each $q^2$ bin, defined as the difference between the maximum and minimum value in each bin for the six cocktail variations. The effects of these variations are shown in Figure 7.27.

Figure 7.27: Reconstructed $q^2$ distributions in nine bins for variations in the background cocktail. The cocktail using the ratios based on Table 7.6 is shown in red, and the six varied cocktails are shown in black. The uncertainties on the red points are based on the spread in each bin.

The $q^2$ signal distribution in data is shown in Figure 7.28. To account for the difference

---

1With the exception of the fake $D^*$ mode, which can only be increased by 6.5% until it is equivalent to 100% of the cocktail.
between MC and signal, the entries in each data bin are corrected by the application of the efficiency matrix given in Tables 7.4 and 7.5. A final correction for the difference between reconstructed MC and the underlying generated distribution is made by scaling each bin by the ratio of generated MC to reconstructed MC true MC in Figure 7.23. This rescaling compensates for events which are not reconstructable. The final corrected data distribution is shown in Figure 7.29.

![Figure 7.28: Reconstructed $q^2$ distributions in nine bins for wrong-sign subtracted data (black), background cocktail (red) and estimated signal from data (blue). The signal from data is the difference between the wrong-sign subtracted data and background in each bin.](image)

![Figure 7.29: Corrected signal $q^2$ distribution in data. The assessment of uncertainties is described in the text.](image)

The uncertainties in the final distribution are assessed by adding in quadrature three
7.3. DATA ANALYSIS

sources. The uncertainties due to data and MC differences are assessed by applying the efficiency error matrix\(^1\) to the data. Secondly, an uncertainty is estimated for the correction between reconstructed data and the underlying \(q^2\) distribution, and is set to be equal to the difference between the MC true reconstructed and generator level distribution in Figure 7.23. The third source is the uncertainty due to background cocktail choice, which is taken as half the spread of points in each \(q^2\) bin in Figure 7.27.

The yield in each bin \((i)\) of \(q^2\) is proportional to the partial width in that bin, \(\Delta \Gamma_i\).

From (2.2.21):

\[
\frac{d\Gamma}{dq^2} = \frac{G_F^2 |V_{cs}|^2}{24\pi^3} |f_+(q^2)|^2 |p_K|^3, \tag{7.3.2}
\]

from which we define:

\[
\Delta \Gamma_i \equiv \int_i^* \frac{d\Gamma}{dq^2} dq^2. \tag{7.3.3}
\]

The simple single-pole form of \(f_+(q^2)\) is:

\[
f_+(q^2) = \frac{f_+(0)}{1 - q^2/m_{pole}^2}. \tag{7.3.4}
\]

A binned fit with the above partial width (with the single-pole form factor) is made to the data distribution (Figure 7.29). In the fit, the two free variables are the pole mass and an arbitrary overall normalisation constant. We minimise the sum:

\[
\chi^2_{pole} \equiv \sum_i \left( \frac{\Delta \Gamma_i - N_i}{\sigma_i} \right)^2 \tag{7.3.5}
\]

across the kinematic range of \(q^2\), where \(N_i\) and \(\sigma_i\) are the number events in bin \(i\) and the associated uncertainty. The resulting best-fit value for the pole mass is \(m_{pole} = 2.35\ \text{GeV/}c^2\) (the uncertainty on this number is discussed below), which compares to a theoretical expectation of \(m_{pole} = m(D_s^*+)^2 = 2.1123 \pm 0.0005\ \text{GeV/}c^2\). The value of \(\chi^2_{pole}\) at the best-fit value is 22.4. Explicitly requiring the \(m_{pole} = m(D_s^*)\) results in a \(\chi^2_{pole}\) value of 24.0. The single-pole model with \(m_{pole}\) set at the most precisely measured value of 1.884 GeV/\(c^2\) is shown against the data in Figure 7.30. For comparison, this plot also shows the distribution for \(m_{pole}\) set at the most precisely measured value of 1.884 GeV/\(c^2\).

The uncertainty in this number can be estimated by observing how the quality of the fit depends on the value of \(m_{pole}\). We define a reduced \(\chi^2\):

\[
\chi^2_r(m_{pole}) = \frac{1}{nDOF - 1} \sum_i \left( \frac{\Delta \Gamma_i(m_{pole}) - N_i}{\sigma_i} \right)^2 \tag{7.3.6}
\]

for any allowed value of \(m_{pole}\), where \(nDOF\) represents the number of degrees in the fit, equal in this case to nine. The variation with this quantity as a function of \(m_{pole}\) is shown

---

\(^1\)This is the matrix of errors associated with each entry in Tables 7.4 and 7.5.
Figure 7.30: Corrected signal $q^2$ distribution in data (black) with the best-fit single-pole form factor model, with $m_{\text{pole}} = 2.35$ GeV/$c^2$ (red) and with the most precise measurement of $m_{\text{pole}} = 1.884$ GeV/$c^2$ (blue).

in Figure 7.31.

Figure 7.31: Variation of $\chi^2$ as a function of $m_{\text{pole}}$. The minimum corresponds to the best-fit $m_{\text{pole}}$ value of 2.35 GeV/$c^2$.

We estimate the uncertainty on the best-fit value by the range of $m_{\text{pole}}$ over which the value of $\chi^2$ increases by half of one unit. This yields unsymmetric uncertainties and the result $m_{\text{pole}} = 2.35^{+0.81}_{-0.35}$ GeV/$c^2$. This measurement of can be compared to those by Belle [26] ($m_{\text{pole}} = 1.82 \pm 0.04_{\text{stat}} \pm 0.03_{\text{syst}}$ GeV/$c^2$), BaBar [22] ($m_{\text{pole}} = 1.884 \pm 0.012_{\text{stat}} \pm 0.015_{\text{syst}}$ GeV/$c^2$) and CLEO [21] ($m_{\text{pole}} = 1.93 \pm 0.02_{\text{stat}} \pm 0.01_{\text{syst}}$ GeV/$c^2$).

The scope for future work on this mode is clear. This is the first time the mode $D^0 \rightarrow$
7.3. **DATA ANALYSIS**

$K^−\mu^+\nu_\mu$ has been studied at LHCb, and there is potential for future analysis of this decay channel. The main objective of the work presented in this thesis was the development of the trigger to allow this mode to be studied. The most likely future application will be a search for $D$ mixing in this mode, although there is potential that additionally LHCb could make a precise measurement of the form factors.

The analysis presented here was based on around one month’s data. The most promising direction for future study would use a sample of $D^0$ mesons originating in $B$ decays. For this class of candidate the displaced vertices and $B$ mass constraint allow for a cleaner data sample.

A further necessity in future study of this mode is a precise assessment of systematic uncertainties. The uncertainties presented in the above analysis are estimates only, and it is clear that to publish a result a more in-depth analysis is required. These will include an understanding of PID performance, the effects of tracking inefficiencies, and the origin of MC/data differences.

A key source of statistical uncertainty in this study is due to the small numbers of background MC events passing all selection cuts. This is due to the necessarily low efficiency for background events to pass selection, but as can be seen in the fits of MC to data, 32% of events in the signal region are backgrounds. An increase in MC statistics by a factor of 10 would allow for a more complete study of this effect.

One future area of work would be to remove the reliance on MC and extract backgrounds from data. For example, the stripping line also searches for the non-allowed (in the absence of mixing) decay $D^{∗+} → \pi^+D^0(K^−\mu^+\nu_\mu)$. From this class of candidates, it may be possible to extract the number of ‘fake’ $D^*$ candidates purely from data. The result of a naïve subtraction of fake candidates (where the slow pion and muon have opposite signs) from signal candidates in data for the $\delta_m$ distribution is shown in Figure 7.32. The tail of the subtracted distribution is substantially reduced compared to 7.22.

The power of LHCb as a tool for gathering signal events is evident. With less than two months of data, operating with a trigger prescale of 1/20, hundreds of thousands of signal events were recorded. As is the case with many high branching fraction charm analyses at LHCb, one of the largest difficulties comes from ‘drowning in signal’.

The challenges for future analyses of this mode will be in improving the separation between signal and background, and constructing background models with higher MC statistics. A parallel study of the decay $D^{∗+} → \pi^+D^0 → \pi^+h^-h^+$ would enable a measurement of the ratio

$$R_0 = \frac{B(D^0 → K^−\mu^+\nu_\mu)}{B(D^0 → h^-h^+)}.$$  \hspace{1cm} (7.3.7)

This method of normalising to a chosen hadronic mode was used by CLEO [98], and enables the reduction of certain systematic errors on reconstruction inefficiencies. This would enable a precise measurement of $B(D^0 → K^−\mu^+\nu_\mu)$. 

163
Figure 7.32: Distribution of $\delta_m$ after neutrino reconstruction for signal (blue) and wrong-sign $\pi_s$ candidates (red) in data. The resulting subtracted distribution is shown in black. Error bars are purely statistical.
LHCb has met or exceeded almost all performance expectations since the LHC began producing data in 2010. The VELO has performed well and during 2011 and has been able to consistently achieve a signal to noise ratio greater than 17. The VELO adapted well to unexpected problems, including higher than anticipated occupancies. This was due in part to the effectiveness of the offline monitoring, including monitoring of gain.

One of the most notable performance achievements is the motion system. The VELO is one of very few movable components in an LHC experiment, and the VELO has now been completely closed safely more than a hundred times, a process which typically takes less than five minutes.

The method developed for gain calibration has proved effective and stable. Experience from 2011 data collection shows that recalibration is only generally necessary after several months of data taking, or if a new TELL1 board is installed after a failure.

At the beginning of 2011, the VELO was monitored 24 hours a day by shift workers in the LHCb control room. Now, the VELO monitoring is now performed remotely by on-call shift workers. This required that all monitoring tasks were streamlined and allowed non-expert users to understand any problems. The process for recalibrating the VELO and uploading parameters to the TELL1 boards, as performed during a gain scan and calibration, has been optimised and documented. A complete recalibration of the VELO can now be performed within one day.

Although the VELO has potential for performing $dE/dx$ particle identification, the study presented in this thesis suggests that for geometrical reasons it is difficult for the forward-oriented LHCb detector to provide similar performance to the more $4\pi$ detectors ATLAS and CMS.

Finally, the trigger and stripping lines described here allow scope for much further work. The data collected are minimally biased in $q^2$, and will potentially allow LHCb to make precision measurements of the form factors for $D^0 \rightarrow K^-\mu^+\nu_\mu$ and $D^0 \rightarrow \pi^-\mu^+\nu_\mu$. After several more months of data taking, the trigger will have selected a large sample of
$D^0$ mesons originating in $B$ decays, which we expect will be the cleanest sample for this study.

The second use for this sample of $D^0$ decays is to search for $D$ mixing in this system. A decay exhibiting mixing is one where the flavour of a $D^0$ or $\overline{D}^0$ meson is tagged by the sign of the slow pion ($\pi_s$) in the decay of a $D^*$ or $\overline{D}^*$, and the semileptonic decay proceeds with a muon of opposite sign to $\pi_s$.

Semileptonic decays at LHCb have only been studied in depth since 2011, and provide an interesting area of study with many potential precision measurements to be made. A disadvantage that LHCb has when compared to many of the other world leading experiments in this domain is the busy hadronic environment from which we must extract signal. Experiments such as BELLE and BaBar collected data from lepton colliders which can yield far cleaner data samples. However, LHCb has its own advantages in the high statistics collected and precise vertexing and track measurement.

With large amounts of data expected in the next few months and years, LHCb is in an excellent position to contribute to our understanding of semileptonic charm decays.
In total there are around $10^6$ TELL1 configuration parameters. The gain settings are stored along with the other parameters in the VELOCOND database using one XML object per TELL1. The XML files are read by the LHCb PVSS control system and used to configure the TELL1 boards.

The XML files for gain settings are generated `MakeGainParams.py`, a script which is kept in the `Velo/VetraScripts` package. The gain settings are written to a ROOT file by `drawFHSPlots` and then subsequently read by this python script.

After the fitting routine, the optimum gain settings are non-integer numbers, so the algorithm rounds these to their nearest integer values and converts them into the required hexadecimal format for uploading to the TELL1 boards. The gain XML files have a syntax as follows:

```xml
<?XML version='1.0' encoding='UTF-8'?>
<DDDB>
  <condition name="VeloTELL1Board123">
    <paramVector name="gain" type="string">
      0xf72e 0xfa3a 0xf77b 0xf7dc 0xf83f 0xf5bc
      0xf4bf 0xf6af 0xf2ba 0xf5ae 0xf47c 0xf6a4
      0xf82b 0xf5ef 0xf699 0xf76c 0xf892 0xf9dc
      0xf6be 0xfb67 0xf931 0xf758 0xf3c3 0xf2c3
      0xf4e9 0xf838 0xf401 0xf73d 0xf41f 0xf80d
      0xf64e 0xf6ae 0xf709 0xfad9 0xf640 0xf978
      0xf8da 0xf4d9 0xf2c1 0xf67e 0xf5f3 0xf689
      0xf56d 0xf6e1 0xf75b 0xf93d 0xf64e 0xf59b
      0xf3f7 0xf21f 0xf16f 0xf56b 0xf680 0xf6e5
      0xf4c2 0xf6c3 0xf3e0 0xf3e3 0xf259 0xf3a8
      0xf498 0xf534 0xf477 0xf3a2
    </paramVector>
  </condition>
</DDDB>
```
</condition>
</DDDB>
DIFFERENTIAL DECAY RATE

From (2.2.21):
\[
\frac{d\Gamma}{dq^2} = \frac{G_F^2 |V_{cs}|^2}{24\pi^3} |f_+(q^2)|^2 |p_K|^3. \tag{B.0.1}
\]

The simple single-pole form of \( f_+(q^2) \) is:
\[
f_+(q^2) = \frac{f_+(0)}{1 - q^2/m_{\text{pole}}^2}. \tag{B.0.2}
\]

To measure the value of \( m_{\text{pole}} \), the number of signal events seen in each bin of \( q^2 \) is compared to the expectation from an integration of (B.0.1) across the bin. This is equivalent to an integration of \( |f_+(q^2)|^2 |p_K|^3 \) over the bin range, where \( |p_K| \) is the magnitude of the kaon’s 3-momentum in the rest frame of the \( D^0 \). Thus, it is necessary to express \( p_K \) as a function of \( q^2 \). For clarity, in this section all 3-vectors are written in boldface.

In the \( D^0 \) rest frame:
\[
p_D = p_K + q \tag{B.0.3}
\]

where \( q \) represents the 4-momentum of the lepton pair.

\[
q^2 = (p_D - p_K)^2 \tag{B.0.4}
\]
\[
= p_D^2 + p_K^2 - 2p_D p_K \tag{B.0.5}
\]
\[
= m_D^2 + m_K^2 - 2(E_D E_K - p_D \cdot p_K) \tag{B.0.6}
\]
\[
= m_D^2 + m_K^2 - 2m_D E_K \tag{B.0.7}
\]

since in the \( D^0 \) rest frame, \( p_D \) is zero and \( E_D = m_D \). Thus,
\[
E_K = \frac{m_D^2 + m_K^2 - q^2}{2m_D} \tag{B.0.8}
\]
APPENDIX B. DIFFERENTIAL DECAY RATE

and

$$|\mathbf{p}_K(q^2)|^3 = \left( E_k^2 - m_K^2 \right)^{\frac{3}{2}}$$  \hspace{1cm} (B.0.9)

$$= \left( \frac{m_D^2 + m_K^2 - q^2}{4m_D^2} \right)^{\frac{3}{2}} - m_K^2 \right)^{\frac{3}{2}}. \hspace{1cm} (B.0.10)$$

The partial decay rate in bin $i$ is then (ignoring multiplication constants) given by:

$$\Delta \Gamma_i \equiv \int_1^i \frac{d\Gamma}{dq^2} dq^2 \propto \int_1^i \left( \frac{(c_1 - y)^2}{c_2} - c_3 \right)^{\frac{3}{2}} \frac{1}{(1 - y/m_{\text{pole}}^2)^2} dy, \hspace{1cm} (B.0.11)$$

where for simplicity $y = q^2$, $c_1 = (m_D^2 + m_K^2)$, $c_2 = 4m_D^2$ and $c_3 = m_K^2$.

The kinematically allowed range of $q^2$ is given by:

$$m_{\mu}^2 < q^2 < (m_D - m_K)^2. \hspace{1cm} (B.0.12)$$

For common choices of $m_{\text{pole}}$ (such as $m_{\text{pole}} = M(D_s^0)$), $m_{\text{pole}}^2 > q^2$ in the kinematically allowed region and (B.0.11) can be integrated analytically.

To extract the value of $|f_+(q^2)|^2$ at the centre of bin $i$ ($q = q_i$), it is necessary to account for the width of the bin and the variation in $|\mathbf{p}_K|$ across the bin. We can take:

$$f_+(q_i^2) = \frac{1}{|V_{cs}|} \sqrt{\frac{24\pi^3 \Delta \Gamma_i}{G_F^2 |p_i^3| \Delta q_i^2}} \hspace{1cm} (B.0.13)$$

where $\Delta q_i^2$ is the width of bin $i$ and $p_i^3$ is the effective value of $|\mathbf{p}_K|^3$ across the bin:

$$p_i^3 = \frac{\int_i |f_+(q^2)|^2 |\mathbf{p}_K|^3 dq^2}{\int_i |f_+(q_i^2)|^2 |\mathbf{p}_K|^3 dq_i^2}. \hspace{1cm} (B.0.14)$$


[38] J.-L. Caron, The Four Main LHC Experiments, 1991, 
http://cdsweb.cern.ch/record/841493.


http://cdsweb.cern.ch/record/841573.


[91] M. Williams, HLT2 topological and charm lines, Presented at the 41st LHCb analysis and software week at CERN, 28 September, 2010.


