

# Performance of Hybrid Photon Detectors and Studies of Two-Body Hadronic $B$ Decays at LHCb

Laurence Carson



University of Glasgow | Department of  
Physics & Astronomy

University of Glasgow  
Department of Physics and Astronomy

*Submitted in fulfilment of the requirements  
for the degree of Doctor of Philosophy*

December 2009

© L. Carson, December 2009

# Performance of Hybrid Photon Detectors and Studies of Two-Body Hadronic $B$ Decays at LHCb

## Abstract

The LHCb experiment at the CERN LHC accelerator will begin physics data taking in late 2009. LHCb aims to discover New Physics processes via precision measurements using heavy flavoured hadrons, such as  $B$  and  $D$  hadrons. This thesis describes studies relevant to measurements of  $B$  decays to hadronic final states at LHCb. The Ring Imaging Cherenkov (RICH) counters of LHCb are crucial to the performance of such measurements. They use arrays of Hybrid Photon Detectors (HPDs) as their photodetection system. Detailed results are presented from the characterisation programme of the entire sample of 557 HPDs that were produced. Their overall performance is found to be outstanding, with only 2.2% of HPDs judged to be unusable for the RICHes. The LHCb requirements and the contractual specifications are met and often exceeded in key areas. The measurement of the single photoelectron detection efficiency,  $\eta$ , of the HPD anode is described in detail. The efficiency was measured as  $\eta = (87.9 \pm 1.4)\%$ . This value exceeds the LHCb-RICH requirement, and is in agreement with previous measurements.

A method to measure the detector proper time resolution for two-body hadronic  $B$  decays from data, making use of the per-event proper time error, is described. A proper time resolution model is proposed and is shown to accurately match the simulated resolution for these decays. The model parameters can be measured on data by fitting the flavour-tagged proper time distribution of the  $B_s \rightarrow K^- \pi^+$  decay. Constraining the proper time resolution model via this method can potentially reduce systematic errors in time-dependent studies.

A study is presented which examines the prospects of LHCb to discover new baryonic  $B$  decay modes, with particular focus on the experimentally most promising mode,  $B_d \rightarrow p\bar{p}$ . It is found that a  $5\sigma$  discovery of  $B_d \rightarrow p\bar{p}$  is possible with only  $0.25 \text{ fb}^{-1}$  of nominal LHCb data, if its true branching fraction is close to the current experimental upper limit.

Finally, the prospects of LHCb to measure the direct and mixing-induced  $\mathcal{CP}$  asymmetries for the decay  $B_d \rightarrow \pi^+ \pi^-$ , via a time-dependent study, are assessed. A fit is made to the invariant mass and proper time distributions of simulated data. The total sensitivities with early data ( $0.3 \text{ fb}^{-1}$ ) are found to be  $0.135(\text{stat}) + 0.012(\text{syst})$  and  $0.093(\text{stat}) + 0.018(\text{syst})$  for the direct and mixing-induced asymmetries respectively. These sensitivities are competitive with current experimental measurements, and indicate that LHCb will come to dominate the world average values for these  $\mathcal{CP}$  asymmetries as more data is collected.

# Acknowledgements

This section of the thesis, which is easily the most fun part to write<sup>1</sup>, gives me the opportunity to say thank you to the people who have helped me on the journey from my state of general ignorance four years ago to the successful completion of this thesis; those who have made various parts of those four years more enjoyable; and a few who managed to do both.

I must begin by thanking my supervisors, Paul Soler and Chris Parkes; firstly for giving me the opportunity to join the Glasgow LHCb group, secondly for imparting many well-placed words of wisdom, and finally for their assiduous reading of this thesis and suggestions for improving it. Any mistakes that remain are my own. Thanks also to PPARC/STFC for sponsoring my studentship.

Apart from my supervisors, many people from the Glasgow group deserve thanks for their day-to-day collaboration and help, and for making the group a pleasure to work in. Firstly, thanks to Andrew for his helpfulness and patience when showing me the ropes of the HPDs and when working on the backpulse measurement, and for proving to me that you can support ManYoo but still be a true Manc. Massive thanks to Franciole for all of his kind help with my basic software troubles. I simply wouldn't have survived the first year of my PhD without it. Thanks also for providing me with the L<sup>A</sup>T<sub>E</sub>X template for this thesis. Thanks to Alison for getting me started with DaVinci and ROOT, and with my proper time work. A big thanks to Marco for many enlightening discussions, on physics and everything else; and for developing the code that was such a useful basis for the work in chapter 6. Thanks to Eduardo, Vava and Tomasz for their help with overcoming problems when I needed it, and in particular to Eduardo for his collaboration on the work in chapter 5. Thanks to Lars for answering my technical questions over lunch, for teaching me so much about the Swedish take on life, and for showing me exactly where my limits are on the piste. Thanks to Njaka for always being ready to offer help on any topic whatsoever, and for the invitation to his wedding. Last but not least, thanks to Michael and Paul for undertaking the continuation of the work in chapter 6 to get it ready for real data.

Of course, I have also benefited from the expertise of many other collaborators from outside Glasgow. Thanks to all the members of the PDTF team for their dedication and hard work in getting all the HPDs tested. Thanks to Gerhard Raven and Peter Vankov for being kind enough to invite me to NIKHEF to spend a week working with them, which kick-started the work that became chapter 4 of this thesis. Thanks to George Hou for inspiring the work in chapter 5 by pointing out the interesting experimental and theoretical status of two-body charmless baryonic  $B$  decays. Thanks to Thierry Gys, Guy Wilkinson and Vincenzo

---

<sup>1</sup>It doesn't have much competition in this respect, but you can only beat what's put in front of you.

Vagnoni for reviewing various LHCb notes that much of the work in this thesis is based on. They helped to improve both the work and the write-ups.

Moving on from my collaborators, I owe a huge debt of gratitude to my family, especially my Mum, for all of the support, encouragement and belief in me that they have shown over the years. I couldn't have done it without you and I really can't thank you enough.

Next, I want to thank everyone I got to know in Geneva who made the 17 months I spent living there one of the best times of my life. Since it would be impossible to mention all of you by name I will refrain from trying, but I hope you all know who you are. Whether it involved lunch, beer o'clock, skiing, fondue (in or out of "season"), pizza, steak, company on the #9 or #56 bus, or just general banter, it was invariably great. Special thanks to those involved in the fantastic trips to Prague, Zürich, the Netherlands and Illinois.

Back in Glasgow, a lot of thanks should go to everyone who at one time or another formed part of the lunch crowd. The lunch hour was often a much-needed oasis in a day of HPD testing, coding or thesis writing. Several of the lunch crowd, apart from the LHCb people already mentioned, deserve a special thanks. Dima, for all the good times in the flat this past year, and latterly for "kicking my ass" to help make sure that this thesis got submitted in time. Kenny Wraight, for having instigated the tradition of the Sunday night pub quiz, and also for his sheer bloody-mindedness in a debate, regardless of the topic. Aaron Mac Raighne, for being even more fun than your typical Irishman, and for his patience in teaching me how to pronounce his surname. Andrew Laing, for providing some great West Coast vs. East Coast banter, and for inviting me to Valencia.

Now, I'd like to leave the workplace behind to thank my other friends who have helped keep me sane and remind me that there is life outside physics! A big thanks to Eilidh for always being available for lunch or coffee, for tolerating my constant criticism of both her musical tastes (completely deserved) and her singing prowess (wholly undeserved), and for generally being an awesome friend. Thanks to Ellen for providing top-quality company over lunch and at gigs, and for never succumbing to the temptation to give me a richly deserved punch in the face for my many disparaging comments about arts students and vegetarians. Thanks also to all those I have shared flats with over the years for putting up with me. Finally, thanks to my old friends from Irvine; Alan, Daniel, Gary, Laura, Andrew, Kenny, and James; for countless fun hours spent at gigs, in the snooker club, up hills and everywhere else. Special thanks to those from Irvine who visited me in Switzerland, even if some of the ensuing antics are best forgotten.

Looking to the future, I will finish by thanking Bernardo Adeva and the Santiago de Compostela group for giving me the opportunity to continue working on LHCb, and to finally get my hands on some real data! The coming year promises to be a very exciting one.

# **Declaration**

The research results presented in this thesis are the product of my own work. Appropriate references are provided when results of third parties are mentioned. The research presented here was not submitted for another degree in any other department or university.

Laurence Carson

# Preface

The LHCb detector is one of four large experiments that are set to begin data taking at the Large Hadron Collider (LHC), a particle accelerator ring located at the European Laboratory for Particle Physics (CERN), near Geneva, Switzerland. LHCb aims to make groundbreaking discoveries of New Physics through precision studies of heavy flavoured hadrons, such as  $B$  and  $D$  hadrons. The studies presented in this thesis aim to contribute towards the ability of LHCb to make such discoveries via particular measurements of  $B$  decays to hadronic final states. A brief overview will now be given of the contents and structure of this thesis.

Chapter 1 gives the theoretical background required to place the studies of this thesis in context. The Standard Model (SM) of particle physics is described, with a focus on the SM description of quark mixing, which is known as the CKM mechanism. The way in which  $CP$  violation, which characterises differences in behaviour between particles and antiparticles, is incorporated into the CKM mechanism is discussed. Several methods to constrain the CKM mechanism and search for New Physics effects using hadronic  $B$  decays are outlined.

Chapter 2 provides a brief overview of the LHC accelerator complex and the purpose of the four main LHC experiments. The LHCb detector is then described in depth. Particular attention is paid to describing the Ring Imaging Cherenkov (RICH) counters of the experiment. These detectors aim to provide excellent particle identification (PID) capability for charged hadrons in a wide momentum range of 1–100GeV. High-quality PID is a requirement for studies of hadronic  $B$  decays, to allow different final states that share the same topology to be separated from each other.

The photon detectors of any RICH system play an essential rôle in ensuring its PID performance. The RICHes of LHCb use arrays of Hybrid Photon Detectors (HPDs) to detect the Cherenkov photons emitted by charged particles as they traverse the RICHes. Chapter 3 gives a detailed description of the comprehensive characterisation programme that was carried out on the entire sample of 557 HPDs that were produced for the LHCb RICHes. One particular measurement, that of the single photoelectron detection efficiency of the HPD anode, is highlighted.

Some physics measurements using hadronic  $B$  decays involve time-dependent studies, where a good understanding of the detector proper time resolution is needed to ensure that systematic errors are not introduced into the analysis. Chapter 4 describes a method to measure the detector proper time resolution for two-body charmless hadronic  $B$  decays (referred to within LHCb as  $B \rightarrow h^+h'^-$  decays) from data, without recourse to information from simulation. After the initial data taking period, this method should be able to provide useful constraints on the proper time resolution model for  $B \rightarrow h^+h'^-$  decays. This can help to

reduce systematic errors in time-dependent studies of these decays.

Some hadronic  $B$  decays that are theoretically allowed to occur have yet to be observed experimentally. One class of such decays are the two-body charmless baryonic decays. Chapter 5 describes a study into the feasibility of discovering new baryonic decay modes of  $B$  mesons at LHCb. The main focus of the chapter is on the decay  $B_d \rightarrow p\bar{p}$ , which is considered to be the most likely candidate for the first observation of a two-body charmless baryonic  $B$  decay. An exclusive selection for this decay is developed, and the prospects for LHCb to observe it in the early stages of data taking are assessed. The likelihood of observing other similar baryonic decays is briefly discussed.

As has been stated already, studies of the time-dependent distributions relating to certain hadronic  $B$  decays can yield interesting physics results. A good example is a method for measuring the parameter of the CKM description known as  $\gamma$ . This fundamental parameter can be measured by studying the time-dependent  $\mathcal{CP}$  asymmetry distributions of two  $B \rightarrow h^+h'^-$  decays,  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^+K^-$ . The value of  $\gamma$  measured using this method is sensitive to New Physics effects. Comparing this value for  $\gamma$  with that measured from other  $B$  decays can reveal inconsistencies in the CKM description, demonstrating the presence of New Physics. Chapter 6 describes a study of the potential for LHCb to measure the  $\mathcal{CP}$  asymmetries in  $B_d \rightarrow \pi^+\pi^-$ . Apart from their rôle in the measurement of  $\gamma$ , there is significant experimental interest in the values of these  $\mathcal{CP}$  asymmetries, as current measurements for one of them are not consistent with each other.

A two-stage fit method to measure the  $\mathcal{CP}$  asymmetries is presented. The method measures the signal and background yields using the mass distribution, then uses these yields as an input to a fit to the proper time distribution. The statistical and systematic errors that are expected with this method using early data are estimated, and compared to the current experimental precision.

Chapter 7 gives a summary of the studies presented in this thesis and their results, and looks forward to the initial data taking run of the LHC, which marks the beginning of an exciting new chapter in the story of High Energy Physics.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>ii</b>
<b>Declaration</b>	<b>iv</b>
<b>Preface</b>	<b>v</b>
<b>Contents</b>	<b>vii</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Figures</b>	<b>xiii</b>
<b>1 Flavour Physics and <math>\mathcal{CP}</math> Violation</b>	<b>1</b>
1.1 The Standard Model . . . . .	1
1.1.1 Status of the Standard Model . . . . .	1
1.1.2 Particle Content of the Standard Model . . . . .	2
1.1.3 The One-Generation Standard Model . . . . .	4
1.1.4 Adding Further Generations . . . . .	9
1.1.5 Discrete Symmetries . . . . .	12
1.2 Flavour Physics . . . . .	14
1.2.1 Time Evolution of Neutral Mesons . . . . .	15
1.2.2 $\mathcal{CP}$ Violation in Neutral Meson Decays . . . . .	17
1.3 Testing the CKM Mechanism with $\mathcal{CP}$ Violation Measurements . . . . .	21
1.3.1 The Unitarity Triangles . . . . .	21
1.3.2 Constraints on the Unitarity Triangles . . . . .	23
1.4 Two-body Charmless Hadronic $B$ Decays . . . . .	26
1.4.1 Overview of $B \rightarrow h^+h'^-$ Decays . . . . .	27
1.4.2 Time Dependent $\mathcal{CP}$ Asymmetries in $B_d \rightarrow \pi^+\pi^-$ . . . . .	29

1.4.3	Measuring $\sin 2\alpha_{\text{eff}}$ Using the Decay $B_d \rightarrow \pi^+\pi^-$ . . . . .	30
1.4.4	Measuring $\gamma$ Using $B_d \rightarrow \pi^+\pi^-$ and $B_s \rightarrow K^-K^+$ . . . . .	34
1.5	Summary . . . . .	37
<b>2</b>	<b>The LHCb Detector</b> . . . . .	<b>38</b>
2.1	The LHC . . . . .	38
2.1.1	The LHC Accelerator . . . . .	38
2.1.2	The LHC Experiments . . . . .	39
2.2	The LHCb Detector . . . . .	40
2.2.1	Vertex Locator . . . . .	41
2.2.2	Ring Imaging Cherenkov Counters . . . . .	44
2.2.3	Magnet . . . . .	53
2.2.4	Tracking System . . . . .	54
2.2.5	Calorimeters . . . . .	56
2.2.6	Muon System . . . . .	59
2.2.7	Trigger . . . . .	61
2.2.8	Online System . . . . .	64
2.3	Summary . . . . .	66
<b>3</b>	<b>Characterisation of RICH Hybrid Photon Detectors</b> . . . . .	<b>67</b>
3.1	Hybrid Photon Detectors in LHCb . . . . .	67
3.1.1	Requirements . . . . .	68
3.1.2	Design and Operation . . . . .	69
3.2	Photon Detector Test Facilities . . . . .	72
3.2.1	PDTF Setup . . . . .	72
3.2.2	Visual Inspection and Mechanical Tests . . . . .	76
3.2.3	Software Tests . . . . .	77
3.3	Quantum Efficiency Measurements . . . . .	89
3.4	Measurement of the Photoelectron Detection Efficiency of the HPD Anode . . . . .	92
3.4.1	Why Measure the Photoelectron Detection Efficiency? . . . . .	92
3.4.2	How to Measure the Photoelectron Detection Efficiency . . . . .	93
3.4.3	Factors Affecting Detection Efficiency . . . . .	94
3.4.4	Backpulse Setup at PDTF . . . . .	95
3.4.5	Procedure for Data Taking and Fitting . . . . .	97
3.4.6	Results . . . . .	104
3.5	PDTF Test Results . . . . .	108
3.5.1	Performance . . . . .	108

3.5.2	Discussion	111
3.6	Conclusions	115
<b>4</b>	<b>Proper Time Resolution for Two-Body Hadronic <math>B</math> Decays from Data</b>	<b>116</b>
4.1	Experimental Effects on Decay Rates	117
4.1.1	Proper Time Resolution Effects	118
4.1.2	Mistag Effects	119
4.1.3	Measuring Experimental Effects on Data	120
4.2	Validation of the Resolution Model on Full Monte Carlo Simulation	122
4.2.1	Form of the Proper Time Resolution Model	123
4.2.2	Method for Validation of Model	123
4.2.3	Selection of $B \rightarrow h^+ h'^-$ Decays	126
4.2.4	Results of Model Validation	129
4.3	Determining the Resolution Model Parameters from Data	133
4.3.1	Construction of Toy Data $\tau_{rec}$ Distribution	134
4.3.2	Construction of $\tau_{rec}$ Distribution for Fit	137
4.3.3	Results from Fit to Toy Data	139
4.4	Conclusions	147
<b>5</b>	<b>Two-Body Charmless Baryonic <math>B</math> Decays</b>	<b>149</b>
5.1	Current Status of Theory and Experiment	149
5.1.1	Theoretical Predictions for Branching Ratios	150
5.1.2	Experimental Limits on Branching Ratios	152
5.2	Selection of $B_d \rightarrow p\bar{p}$ Events	152
5.3	Background studies	153
5.3.1	Sources of Background	153
5.3.2	Background Suppression	154
5.4	LHCb Sensitivity to $B_d \rightarrow p\bar{p}$	158
5.4.1	Selection Performance	158
5.4.2	Signal and Background Yields	159
5.4.3	Signal Significance	162
5.4.4	Prospects for 2010	162
5.5	Measurement of the $B_d \rightarrow p\bar{p}$ Branching Ratio	164
5.6	Trigger Mass Window	165
5.7	Other Two-Body Charmless Baryonic $B$ Decays	166
5.7.1	Prospects for the Observation of $B_s \rightarrow p\bar{p}$	166
5.7.2	Prospects for Decays Involving a $\Lambda$	167

5.8	Conclusions . . . . .	167
<b>6</b>	<b>Direct and Mixing-Induced <math>\mathcal{CP}</math> Asymmetries in <math>B_d \rightarrow \pi^+\pi^-</math></b>	<b>169</b>
6.1	Strategy for Fit to Time-Dependent $\mathcal{CP}$ Asymmetry . . . . .	169
6.2	Event Selection and Yields . . . . .	170
6.3	Construction of Toy Data . . . . .	175
6.3.1	Mass Distributions for Toy Data . . . . .	175
6.3.2	Proper Time Distributions for Toy Data . . . . .	175
6.4	Construction of Distributions for Fit . . . . .	178
6.4.1	Mass Distributions for Fit . . . . .	178
6.4.2	Proper Time Distributions for Fit PDFs . . . . .	179
6.5	Results from Fit to Toy Data . . . . .	182
6.5.1	Validation of Proper Time Fitter . . . . .	182
6.5.2	Mass and Proper Time Fits to Realistic Data . . . . .	187
6.5.3	Sources of Systematic Error . . . . .	192
6.5.4	Sensitivity to $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$ and $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$ . . . . .	197
6.5.5	Sensitivity to $\sin 2\alpha_{\text{eff}}$ . . . . .	198
6.6	Conclusions . . . . .	198
<b>7</b>	<b>Conclusions and Outlook</b>	<b>200</b>
7.1	Summary . . . . .	200
7.2	Outlook . . . . .	203

# List of Tables

1.1	World-average experimental values for the magnitude of each element of the CKM matrix. . . . .	12
1.2	World-average experimental values, from direct measurements, for the angles of the CKM unitarity triangle. . . . .	24
1.3	Values for the angles of the CKM unitarity triangle from a global fit for the apex of the triangle. . . . .	26
2.1	Cherenkov angle resolution and the expected number of detected photoelectrons in the RICHes. . . . .	52
3.1	Selected contract specifications for the manufactured HPDs. . . . .	71
3.2	Quantum efficiency performance of the manufactured HPDs. . . . .	90
3.3	Contributions to the total error on the sensor efficiency. . . . .	107
3.4	Single photoelectron detection efficiency of the manufactured HPDs. . . . .	108
3.5	Performance of the manufactured HPDs. . . . .	109
3.6	Classification criteria for HPDs. . . . .	112
4.1	List of cuts comprising the inclusive $B \rightarrow h^+h'^-$ selection. . . . .	129
4.2	Mass and PID cuts for the exclusive $B \rightarrow h^+h'^-$ selections. . . . .	129
4.3	Comparison of resolution model parameters from fits to proper time residual distributions for $B \rightarrow h^+h'^-$ decays. . . . .	133
4.4	Inputs for the proper time distribution of the signal and background components of the toy data. . . . .	134
4.5	Intervals used to generate the fit PDF and perform the fit. . . . .	138
4.6	Comparison of fit results with $2 \text{ fb}^{-1}$ of data and $10 \text{ fb}^{-1}$ of data. . . . .	145
4.7	Comparison of fit results with incorrect mistag value assumed in the fit PDF. . . . .	146
5.1	Theoretical predictions for the branching ratios of different baryonic two-body $B$ decays. . . . .	151

5.2	Experimental upper limits on the branching ratios of different baryonic two-body $B$ decays. . . . .	152
5.3	List of extra cuts applied in the selection of $B_d \rightarrow p\bar{p}$ events. . . . .	153
5.4	Channel-specific values used to evaluate signal and background yields. . . . .	159
5.5	General constants used to evaluate signal and background yields. . . . .	160
5.6	Values used to evaluate signal yield. . . . .	160
5.7	Upper limits on background yields and resulting background-to-signal (B/S) ratios. . . . .	161
6.1	$B \rightarrow h^+h'^-$ and background yields with $0.3 \text{ fb}^{-1}$ of data, before and after PID and mass cuts. . . . .	174
6.2	Values used in the toy data for the $\mathcal{CP}$ asymmetries. . . . .	176
6.3	Comparison of tagging efficiencies and mistag rates in different tagging strategies. . . . .	177
6.4	Number of signal and background events in a typical toy data sample, corresponding to $0.3 \text{ fb}^{-1}$ of data. . . . .	178
6.5	Summary of results from proper time fits to large simulated data samples containing signal and main backgrounds. . . . .	185
6.6	Summary of results from mass fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds. . . . .	190
6.7	Summary of results from proper time fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds. . . . .	192
6.8	Summary of results from proper time fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds, with incorrect estimation of $\omega_{\text{tag}}$ in the fit PDF. . . . .	194

# List of Figures

1.1	Illustration of the two “non-squashed” CKM unitarity triangles. . . . .	22
1.2	The main Feynman diagram contributing to $B_{d,s} - \bar{B}_{d,s}$ mixing. . . . .	23
1.3	Global fit of the apex of the CKM unitarity triangle. . . . .	25
1.4	Feynman diagrams for the processes contributing to $B_{d,s} \rightarrow \{\pi, K\}^+ \{\pi, K\}^-$ decays. . . . .	28
1.5	Current B-Factory measurements of the direct and mixing-induced $\mathcal{CP}$ asymmetries in the $B_d \rightarrow \pi^+ \pi^-$ decay. . . . .	31
1.6	Isospin triangles relating the amplitudes for $B_d \rightarrow \pi^+ \pi^-$ , $B^+ \rightarrow \pi^+ \pi^0$ and $B_d \rightarrow \pi^0 \pi^0$ . . . . .	33
1.7	Constraint on the apex of the CKM unitarity triangle from direct measurements of $\gamma$ . . . . .	34
2.1	Overview of the layout of the LHC. . . . .	39
2.2	Angular distribution of $b$ and $\bar{b}$ quarks from proton-proton collisions at 14 TeV. . . . .	41
2.3	Layout of the LHCb detector. . . . .	42
2.4	Geometry of the VELO stations along the LHCb $z$ axis. . . . .	43
2.5	Geometry of the VELO sensors (R and $\Phi$ ). . . . .	44
2.6	Invariant mass distributions for $B_d \rightarrow p\bar{p}$ and selected three-body backgrounds, before and after RICH PID information is used. . . . .	45
2.7	Schematic diagram and photograph of RICH1. . . . .	47
2.8	Schematic diagram and photograph of RICH2. . . . .	47
2.9	Distribution in polar angle $\theta$ and momentum $p$ of tracks from $B_d \rightarrow \pi^+ \pi^-$ events. . . . .	48
2.10	Dependence of the Cherenkov angle $\theta_c$ on momentum for different particle types in the three radiators of the LHCb RICHes. . . . .	48
2.11	Effect of longitudinal magnetic field on HPD hit pattern. . . . .	49
2.12	Schematic diagrams of the Laser Alignment Monitoring Systems (LAMS) for the RICHes. . . . .	50

2.13	Efficiency and mis-identification rates, as a function of momentum, for identifying a kaon as “heavy”, i.e. as a kaon or a proton. . . . .	53
2.14	Perspective view of the LHCb dipole magnet. . . . .	54
2.15	Layout of Tracker Turicensis layers. . . . .	55
2.16	Outline of the tracking stations layout, and cross-section of an OT module. .	56
2.17	Segmentation scheme of the calorimeters. . . . .	57
2.18	Photographs of modules for the ECAL, and of the completed ECAL detector.	58
2.19	Schematic diagram of scintillator-absorber layout for the HCAL, and photograph of the completed HCAL detector. . . . .	59
2.20	Layout of chambers in a muon station quadrant, and segmentation of each region of the quadrant. . . . .	60
2.21	Diagram of a four-layer MWPC, and photograph of a two-layer MWPC. . .	61
2.22	Schematic diagram and photograph of a triple-GEM detector. . . . .	61
2.23	Overview of the structure of the LHCb High Level Trigger system. . . . .	64
2.24	Overview of the architecture of the LHCb Online system. . . . .	65
3.1	Schematic diagram of the column mounting scheme for RICH2, and photograph of a single RICH2 column populated with HPDs. . . . .	68
3.2	Schematic diagram and photograph of the pixel hybrid photon detector. . .	69
3.3	Cross-section diagram of the HPD silicon sensor. . . . .	70
3.4	Expected distributions of Cherenkov photons in RICH1 and RICH2 for collisions that contain a $b$ quark. . . . .	72
3.5	Photograph of an HPD test station. . . . .	74
3.6	Schematic diagram of the standard PDTF setup. . . . .	74
3.7	Screenshot showing the main Labview Virtual Instrument used to run the software tests on HPDs at PDTF. . . . .	78
3.8	Screenshot from PDTF showing current-voltage behaviour of HPD pixel chip.	79
3.9	Screenshots from PDTF showing response of the HPD pixel chip with different thresholds. . . . .	80
3.10	Screenshots from PDTF showing response of an HPD pixel chip to a test pulse.	81
3.11	Screenshots from PDTF showing HPD hit rate during high voltage ramp-up, and dark-count rate following high voltage ramp-up. . . . .	82
3.12	Screenshots of strobescans at PDTF. . . . .	84
3.13	Screenshot of bias voltage scan at PDTF. . . . .	85
3.14	Screenshots of long LED runs at PDTF. . . . .	86
3.15	Screenshot of a distortion map run at PDTF. . . . .	87
3.16	Screenshots of dark-count runs at PDTF. . . . .	88

3.17	Schematic of the setup used to measure HPD quantum efficiency at PDTF. .	89
3.18	Distributions of HPD quantum efficiency at 270 nm. . . . .	90
3.19	Comparison of HPD quantum efficiency measurements at PDTF and DEP. .	91
3.20	Diagram showing a Si pixel bump-bonded to a channel of the readout chip.	93
3.21	Schematic diagram of the setup for the measurement of the charge spectrum at the sensor backplane. . . . .	96
3.22	The single photoelectron response of the HPD sensor. . . . .	98
3.23	Analogue data taken from the sensor backplane. . . . .	100
3.24	Typical fit to a backpulse spectrum with the pedestal subtracted. . . . .	101
3.25	Sketch of the backplane signal and analogue gate setup. . . . .	102
3.26	Dependencies of the measured sensor efficiency ( $\eta$ ) on the analogue gate setup and on the light input level. . . . .	103
3.27	Effect of varying the fit region for a spectrum at low analogue $\langle npe \rangle$ . . . . .	106
3.28	Results for $\eta$ using digital gate length of 50 ns. . . . .	107
3.29	Results for $\eta$ using digital gate length of 25 ns. . . . .	108
3.30	Distributions of the number of dead and noisy ALICE pixels. . . . .	109
3.31	Distribution of leakage current at 80 V reverse bias. . . . .	109
3.32	Distributions of average pixel thresholds and noise. . . . .	110
3.33	Distributions of dark-count rate and ion feedback probability. . . . .	110
3.34	Distributions of radius and radial displacement of the photocathode image. .	110
3.35	Final classification of HPDs according to the tests performed at the PDTFs.	111
3.36	Three different classes of leakage current curve for HPDs. . . . .	113
4.1	Illustration of the impact of the proper time resolution $\sigma_\tau$ and the mistag rate $\omega_{\text{tag}}$ on the flavour tagged $\tau_{rec}$ distribution for $B_d \rightarrow K^+\pi^-$ . . . . .	121
4.2	Illustration of the impact of the proper time resolution $\sigma_\tau$ and the mistag rate $\omega_{\text{tag}}$ on the flavour tagged $\tau_{rec}$ distribution for $B_s \rightarrow K^-\pi^+$ . . . . .	121
4.3	Proper time per-event error ( $\sigma_{\tau_{rec}}$ ) distribution for $B_d \rightarrow \pi^+\pi^-$ . . . . .	124
4.4	Single Gaussian fits to proper time residual in two bins of the proper time per-event error. . . . .	124
4.5	Dependencies of the single Gaussian mean and width on the proper time per-event error. . . . .	125
4.6	Resolution model fit to the proper time residual distribution for $B_d \rightarrow \pi^+\pi^-$ events. . . . .	130
4.7	Resolution model fit to the proper time residual distribution for $B_s \rightarrow K^+K^-$ events. . . . .	130

4.8	Resolution model fit to the proper time residual distribution for $B_s \rightarrow K^- \pi^+$ events. . . . .	131
4.9	Resolution model fit to the proper time residual distribution for $B_d \rightarrow K^+ \pi^-$ events. . . . .	131
4.10	Proper time acceptance function used in the toy data. . . . .	135
4.11	Example of a toy dataset including $B_s \rightarrow K^- \pi^+$ signal events, as well as specific background events and combinatoric background events. . . . .	137
4.12	Fit results for the parameter GM with $2 \text{ fb}^{-1}$ of data. . . . .	140
4.13	Fit results for the parameter GS with $2 \text{ fb}^{-1}$ of data. . . . .	141
4.14	Correlation coefficient between GM and GS in the fits with $2 \text{ fb}^{-1}$ of data. . . . .	141
4.15	Fit results for the parameter GM with $10 \text{ fb}^{-1}$ of data. . . . .	143
4.16	Fit results for the parameter GS with $10 \text{ fb}^{-1}$ of data. . . . .	144
4.17	Correlation coefficient between GM and GS in the fit with $10 \text{ fb}^{-1}$ of data. . . . .	144
5.1	Feynman diagrams for the tree-level process contributing to $B_d \rightarrow p\bar{p}$ , and for the gluonic penguin process contributing to $B^+ \rightarrow p\bar{\Lambda}$ . . . . .	151
5.2	Reconstructed mass distributions before and after PID cuts for $B_d \rightarrow p\bar{p}$ , $B_d \rightarrow K^+ \pi^-$ , $B_s \rightarrow K^+ K^-$ and $\Lambda_b \rightarrow pK^-$ events. . . . .	154
5.3	Reconstructed mass distributions before and after PID cuts for $B_d \rightarrow p\bar{p}$ , $B^+ \rightarrow K^+ K^- K^+$ , $B^+ \rightarrow p\bar{p}K^+$ and $B^+ \rightarrow p\bar{p}\pi^+$ events. . . . .	155
5.4	Reconstructed mass distributions before and after PID cuts for $B_d \rightarrow p\bar{p}$ and $B_d \rightarrow \pi^+ \pi^- \pi^0$ events. . . . .	155
5.5	Reconstructed mass distributions before and after PID cuts for $B_d \rightarrow p\bar{p}$ , $B^+ \rightarrow \pi^+ \pi^- K^+$ , $B^+ \rightarrow \pi^+ \pi^- \pi^+$ , $B_d \rightarrow K_S \pi^- \pi^+$ and $B^+ \rightarrow \pi^+ K^- K^+$ events. . . . .	155
5.6	Distributions of $\text{DLL}(p - \pi)$ for $B_d \rightarrow p\bar{p}$ and $B_d \rightarrow \pi^+ \pi^- \pi^0$ events. . . . .	156
5.7	Distributions of $\text{DLL}(p - K)$ for $B_d \rightarrow p\bar{p}$ and $B^+ \rightarrow \pi^+ \pi^- K^+$ events. . . . .	156
5.8	Distributions of track $\chi^2/n_{\text{DoF}}$ for $B_d \rightarrow p\bar{p}$ and ghost tracks from $b\bar{b}$ inclusive events. . . . .	157
5.9	Significance of the $B_d \rightarrow p\bar{p}$ signal as a function of integrated luminosity, assuming a centre-of-mass energy of $E_{\text{CM}} = 14 \text{ TeV}$ . . . . .	162
5.10	Significance of the $B_d \rightarrow p\bar{p}$ signal as a function of integrated luminosity, assuming $E_{\text{CM}} = 10 \text{ TeV}$ . . . . .	163
5.11	Reconstructed mass distribution for offline selected $B_d \rightarrow p\bar{p}$ events taking the pion mass hypothesis for the $B$ -daughters. . . . .	165
5.12	Reconstructed mass distributions for $B_d \rightarrow p\bar{p}$ and $B_s \rightarrow p\bar{p}$ events. . . . .	166

---

6.1	Cumulative mass distribution for $B \rightarrow h^+h'^-$ events following the inclusive selection, with the number of simulated events corresponding to $0.3 \text{ fb}^{-1}$ of data. . . . .	171
6.2	Comparison of $\text{DLL}(K - \pi)$ distributions for selected $B_d \rightarrow \pi^+\pi^-$ and $B_s \rightarrow K^+K^-$ events. . . . .	172
6.3	Comparison of $\text{DLL}(K - \pi)$ distributions for selected $B_d \rightarrow \pi^+\pi^-$ and $\Lambda_b \rightarrow pK^-$ events. . . . .	172
6.4	Cumulative mass distribution for $B \rightarrow h^+h'^-$ events following the inclusive selection and PID cuts, with the number of simulated events corresponding to $0.3 \text{ fb}^{-1}$ of data. . . . .	174
6.5	A typical fit to the proper time distribution of a large simulated data sample containing signal and main backgrounds, and the corresponding fractional asymmetry distribution. . . . .	184
6.6	Distributions of fit results for $\mathcal{CP}$ asymmetries from proper time fits to large simulated data samples containing signal and main backgrounds. . . . .	186
6.7	A typical fit to the mass distribution of a simulated $0.3 \text{ fb}^{-1}$ data sample containing signal and all considered backgrounds. . . . .	188
6.8	Distributions of fit results for decay fractions from mass fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds. . . . .	189
6.9	A typical fit to the proper time distribution of a simulated $0.3 \text{ fb}^{-1}$ data sample containing signal and all considered backgrounds, and the corresponding fractional asymmetry distribution. . . . .	191
6.10	Distributions of fit results for $\mathcal{CP}$ asymmetries from proper time fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds. . . . .	193
6.11	Distributions of fit results for $\mathcal{CP}$ asymmetries from proper time fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds, with $\omega_{\text{tag}}$ underestimated by 0.5% in the fit PDF. . . . .	195
6.12	Distributions of fit results for $\mathcal{CP}$ asymmetries from proper time fits to simulated $0.3 \text{ fb}^{-1}$ data samples containing signal and all considered backgrounds, with $\omega_{\text{tag}}$ overestimated by 0.5% in the fit PDF. . . . .	196

# Chapter 1

## Flavour Physics and $\mathcal{CP}$ Violation

In this chapter the theoretical background to the work presented in this thesis is reviewed. Section 1.1 gives an overview of the Standard Model of particle physics, including its description of quark mixing, which is known as the CKM mechanism. The concept of  $\mathcal{CP}$  violation, corresponding to the asymmetric behaviour of antimatter with respect to matter, is also introduced. Section 1.2 discusses some aspects of flavour physics that are relevant for the work in this thesis, with particular emphasis on the effects of quark mixing and  $\mathcal{CP}$  violation on the behaviour of neutral mesons. Section 1.3 covers in detail the incorporation of  $\mathcal{CP}$  violation into the CKM mechanism, and summarises the current status of  $\mathcal{CP}$  violation measurements in the quark sector. Section 1.4 outlines the theoretical aspects of two-body charmless hadronic  $B$  decays (in particular, the subclass of such decays known as  $B \rightarrow h^+h'^-$  decays), and discusses two physics measurements that LHCb will make by studying such decays. Studies of  $B \rightarrow h^+h'^-$  decays at LHCb will form Chapters 4–6 of this thesis. The chapter is summarised in Sec. 1.5.

### 1.1 The Standard Model

#### 1.1.1 Status of the Standard Model

The Standard Model (SM) of particle physics was developed several decades ago, and successfully predicts the nature of all interactions between all particles that have been observed to date, with the exception of gravitational interactions [1]. However it is widely believed that the SM is not a fundamental theory of nature, rather it is expected that the SM is a low-energy effective theory of a higher energy theory whose nature has yet to be established. This belief arises from several deep theoretical and experimental issues within the SM, which have proved very difficult to resolve.

Two of the main experimental difficulties arise from the field of cosmology. Firstly, there is no candidate particle within the SM for the dark matter which is believed to be abundant in the universe. Secondly, the asymmetry between matter and antimatter in the Universe is observed to be very large, and no mechanism to generate such a large asymmetry exists in the SM. A further experimental difficulty for the SM is that one of its constituent particles, the Higgs boson, has not yet been observed. The Higgs boson is the particle associated with a scalar field called the Higgs field, in the same way that for example the photon is associated with the electromagnetic field. In the SM the Higgs field is responsible for breaking electroweak symmetry and endowing particles with mass, via a process known as the Higgs mechanism.

The main unresolved theoretical issue within the SM is related to the Higgs mechanism, and is known as the Hierarchy Problem. This problem arises because current experimental constraints dictate that the mass of the Higgs boson is less than 1 TeV, which is many orders of magnitude below the maximum possible mass scale, known as the Planck Scale. In the SM there is no mechanism to protect the Higgs mass from being affected by large quantum corrections that could push the mass up to the Planck Scale. Within the SM the Higgs mass can still be “fine tuned” to be below 1 TeV, but such a solution is considered unnatural and not theoretically satisfying.

Solving the above issues, and distinguishing between the many candidates for the more fundamental, higher energy theory, is the main goal of the particle physics experiments that operate at the Large Hadron Collider. One of these experiments is the LHCb detector. It aims to make precision measurements of the characteristics of particles containing charm and beauty quarks. Such measurements can shed light on the reason for the large matter-antimatter asymmetry in the Universe. This is because LHCb will study differences in the way particles and their antiparticles behave, and the existence of such differences (a phenomenon known as  $\mathcal{CP}$  violation) is necessary for a matter-antimatter asymmetry to exist in the Universe [2].

### 1.1.2 Particle Content of the Standard Model

The particle content of the SM consists of a set of elementary particles whose existence, with the notable exception of the Higgs boson, has been confirmed experimentally. The known elementary particles have no structure down to a scale of around  $10^{-19}$  m. Each particle in the SM has a partner, known as its *antiparticle*, which has the same mass but has the sign of its internal quantum numbers, such as charge, reversed. It is possible for a neutral particle to be its own antiparticle. A particle and its antiparticle are related by the combined transformations of parity  $\mathcal{P}$  and charge conjugation  $\mathcal{C}$ , which are described in Sec. 1.1.5. The

---

elementary particles can be classified into two groups: bosons and fermions.

### 1.1.2.1 The Standard Model Fermions

Fermions have half-integer spin values and have wavefunctions that are antisymmetric under the exchange of two identical particles. Hence they obey the Pauli exclusion principle. The fermions can be further subdivided into quarks and leptons.

Quarks carry another charge known as the colour charge, and are subject to the strong force. The colour charge carried by a quark can be one of three states, labelled red, blue and green. Antiquarks carry one of three different colour states, known as antired, antiblue and antigreen. Free quarks are not observed in nature, rather colourless particles called hadrons, consisting of sets of bound quarks, are observed. Due to the requirement that hadrons are colourless, the only allowed combinations of quarks are a quark-antiquark pair (known as a meson) and a set of three quarks or three antiquarks (known as a baryon or antibaryon respectively). There are six flavours of quark in the SM. Three of these — the up ( $u$ ), charm ( $c$ ) and top ( $t$ ) quarks — carry an electric charge of  $2/3$ , with the other three — the down ( $d$ ), strange ( $s$ ) and bottom ( $b$ ) quarks — carrying an electric charge of  $-1/3$ .

The leptons, of which there are again six in the SM, can also be divided into two sets of three based on their electric charge. The electron ( $e$ ), muon ( $\mu$ ) and tau ( $\tau$ ) particles carry an electric charge of  $-1$ , while their associated neutrinos ( $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$ ) carry no electric charge.

The known fermions can be arranged into three “families” or “generations”, reflecting the hierarchy of how they interact with each other (see the following sections). The first generation consists of the  $u$ ,  $d$ ,  $e$  and  $\nu_e$ , the second generation of the  $c$ ,  $s$ ,  $\mu$  and  $\nu_\mu$ , and the third generation of the  $t$ ,  $b$ ,  $\tau$  and  $\nu_\tau$ , with the associated antiparticles being implicitly included in each generation.

### 1.1.2.2 The Standard Model Bosons

The second class of elementary particles is the bosons, which have integer spin values and symmetric wavefunctions. Bosons act as carriers of the different forces that the fermions described above can feel, and hence they mediate the interactions between fermions. The mediator of the electromagnetic force is the photon ( $\gamma$ ), which is massless. The photon couples to electric charge, and so couples to all of the elementary fermions, with the exception of the neutrinos. Since the photon itself does not carry the electric charge, there is no photon self-coupling. This allows the photon to travel freely through space, giving the electromagnetic force an infinite range.

The strong force, which is felt by quarks but not leptons, is mediated by gluons ( $g$ ). While gluons, like the photon, are massless, *unlike* the photon they themselves carry the charge that they mediate, and so they can self-interact. This causes the range of the strong force to be finite and in fact very small ( $\mathcal{O}(10^{-15}$  m), roughly the diameter of a nucleon). Free quarks and gluons are not observed. Given the need to mediate between two quarks or antiquarks, which can carry one of six colour charges, the naïve expectation would be for nine different types of gluon to exist, with each carrying some combination of colour and anticolour charge linearly independent from the combinations of the other gluons. However one of these linearly independent combinations would have to be colourless, so that the gluon could have long range interactions. Since long-range strong interactions are not observed, this combination does not exist, and hence there are only eight types of gluon.

The electroweak force is mediated by three gauge bosons: the  $W^+$ , the  $W^-$  and the  $Z^0$ . Every fermion, including the neutrinos, feels the electroweak force. Each of these gauge bosons is very massive ( $M_W \simeq 80.4$  GeV and  $M_Z \simeq 91.2$  GeV), which causes the range of the electroweak force to be very small ( $\mathcal{O}(10^{-18}$  m)).

The following sections will build up the mathematical description of how, in the SM, the particles described above interact with each other. The description given will assume that neutrinos are massless. Although the existence of non-zero neutrino masses was proven in 1998 through the observation that they could oscillate between flavours [3], the conclusions of the discussions in the following sections would not be changed by their inclusion.

### 1.1.3 The One-Generation Standard Model

Following the approach in [4], the electroweak and strong interactions will first be introduced for a “standard model” containing only the first generation of fermions. Then the experimentally observed second and third generations will be added to build up the full Standard Model.

The theoretical structure of the SM is often considered as two quasi-separate theories — Quantum Chromodynamics (QCD), and Electroweak (EW) theory. QCD describes how particles interact via the strong force, while EW theory deals with the electromagnetic and weak forces. The reason for the theory being split in this way is that the SM is a gauge theory, in which each interaction is a manifestation of the symmetry of a particular gauge group. The choice of gauge groups is made to fit experimental observations; there is no theoretical reason to exclude groups beyond those which are used. In general, a gauge group of the form  $SU(n)$  describes an interaction with  $n^2 - 1$  gauge bosons. Hence  $SU(3)$  can describe the strong interaction with eight gluons, and  $SU(2) \times U(1)_Y$  can describe the electroweak interaction with the photon,  $W^\pm$  and  $Z^0$ . So the gauge group for the SM as a whole is

$SU(3) \times SU(2) \times U(1)_Y$ , with the electromagnetic and weak forces considered as one unified electroweak force. The gauge group  $U(1)$  of the electromagnetic force is related to (but is not equal to) the  $U(1)_Y$  subgroup of the electroweak gauge group (see Sec. 1.1.3.4). This unification of electromagnetic theory and the theory of the weak interaction was achieved in the 1960s by Sheldon Glashow, Abdus Salam and Steven Weinberg [5, 6, 7].

The gauge bosons associated with each of the gauge groups need not be equivalent to the physical gauge bosons described above, and the coupling constants are not in the form normally associated with each interaction (for example the unit electric charge for the electromagnetic interaction). The coupling constant of the Abelian group  $U(1)_Y$  is called  $g'$ , and its gauge boson  $B^\mu$ . For the non-Abelian group  $SU(2)$ , the coupling constant is called  $g$  and the gauge bosons  $W_\mu^1$ ,  $W_\mu^2$ , and  $W_\mu^3$ . It will be seen that the photon,  $W^\pm$  and  $Z^0$  bosons are constructed from linear combinations of  $B^\mu$  and  $W_\mu^{1,2,3}$ . For the non-Abelian group  $SU(3)$ , the coupling constant is called  $g_s$  and the gauge bosons  $G_\mu^a$ , with  $a \in \{1, 2, \dots, 8\}$ . The gluons are constructed from linear combinations of these.

### 1.1.3.1 Lagrangians in Quantum Field Theory

In Quantum Field Theory (QFT), the dynamics of a given system are defined in terms of the Lagrangian  $\mathcal{L}$ , from which the equations of motion are found by evaluating the Euler-Lagrange equations associated with the Lagrangian [8]. In QFT the basic form of the Lagrangian for a given field depends upon the spin of the particle.

For a scalar field  $\phi$  with mass  $m$ , the Lagrangian is:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4!} \phi^4, \quad (1.1)$$

where the last term is the simplest interaction term that can be added while keeping the theory renormalisable.

For a field  $\psi$  with half-integer spin (e.g. a fermion field), the Lagrangian is:

$$\mathcal{L} = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi, \quad (1.2)$$

with the corresponding Euler-Lagrange equation being the Dirac equation.

Finally, for a field  $A_\mu$  with non-zero integer spin (e.g. a gauge boson), the Lagrangian is:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (1.3)$$

defining the field strength  $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ . If  $A_\mu$  represents the photon field, the corresponding Euler-Lagrange equations are the Maxwell equations.

### 1.1.3.2 Overview of the Standard Model Lagrangian

The Lagrangian for the one-generation Standard Model can be divided into parts dealing with different particles, as follows:

$$\mathcal{L} = \mathcal{L}_{\text{gauge bosons kinetic}} + \mathcal{L}_{\text{fermion kinetic}} + \mathcal{L}_{\text{fermion masses}} + \mathcal{L}_{\text{Higgs}}, \quad (1.4)$$

so that there is one kinetic term for the gauge bosons and one for the fermions, a Yukawa coupling term giving mass to the fermions, and a Higgs term giving mass to the gauge bosons. The following sections will consider each of these terms in turn, giving their mathematical form and discussing the physical phenomena that are elucidated.

### 1.1.3.3 Gauge Boson Kinetic Terms

Following the pattern of (1.3), the kinetic terms for the gauge bosons are:

$$\mathcal{L}_{\text{gauge bosons kinetic}} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} - \frac{1}{4}F_{\mu\nu}^A F^{A\mu\nu} + \mathcal{L}_{\text{gauge fixing}} + \mathcal{L}_{\text{ghosts}}, \quad (1.5)$$

where  $B_{\mu\nu}$  is the field strength for  $U(1)_Y$ ,  $F_{\mu\nu}^a$  (with  $a \in \{1, 2, 3\}$ ) the field strengths for  $SU(2)$ , and  $F_{\mu\nu}^A$  (with  $A \in \{1, 2, \dots, 8\}$ ) the field strengths for  $SU(3)$ . The penultimate term allows for gauge fixing, i.e. choosing the gauge for the theory. The gauge fixing has the side-effect of introducing extra particles which can appear in loop processes. The final term is the Lagrangian for these ‘‘Fadeev-Popov ghosts’’, which do not contribute to observable quantities as their effect is cancelled by loops of gauge bosons. The form of these last two terms depends on the chosen gauge.

### 1.1.3.4 Fermion Kinetic Terms

Before the form of the fermion kinetic terms can be discussed, a formalism to describe the coupling of the fermions to the weak interaction is needed: since the weak interaction maximally violates parity, acting only on left-handed fields, it is convenient to split the fermion wavefunction into its left-handed and right-handed components:

$$\psi = \psi_L + \psi_R = \frac{1 - \gamma^5}{2}\psi + \frac{1 + \gamma^5}{2}\psi, \quad (1.6)$$

where the  $\frac{1 \mp \gamma^5}{2}$  are known as the projection operators ( $P_{L/R}$ ) for the left and right-handed components, respectively. This leads (considering for the moment only the first generation of particles) to the left-handed fields forming doublets under  $SU(2)$ :

$$q_L \equiv \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad \text{and} \quad l_L \equiv \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad (1.7)$$

and the right-handed fields forming singlets under  $SU(2)$ :

$$u_R, \quad d_R, \quad \nu_R \quad \text{and} \quad e_R. \quad (1.8)$$

The  $\nu_R$  will not be considered in the following, as a massless right-handed neutrino has no interactions with any other fields.

With the above formalism, the fermion kinetic terms are (following the pattern of (1.2)):

$$\begin{aligned} \mathcal{L}_{\text{fermion kinetic}} = & i \bar{l}_L^T \gamma^\mu \mathbf{D}_\mu^l l_L + i \bar{e}_R \gamma^\mu D_\mu^e e_R + i \bar{q}_L^T \gamma^\mu \mathbf{D}_\mu^q q_L + i \bar{d}_R \gamma^\mu D_\mu^d d_R \\ & + i \bar{u}_R \gamma^\mu \partial_\mu u_R, \end{aligned} \quad (1.9)$$

where the covariant derivatives include the couplings to the relevant gauge bosons, e.g.

$$\begin{aligned} \mathbf{D}_\mu^l &= \partial_\mu \mathbf{I} + ig \mathbf{T}^a W_\mu^a + ig' \mathbf{Y}(l_L) B_\mu, \\ D_\mu^e &= \partial_\mu + ig' Y(e_R) B_\mu, \\ D_\mu^d &= \partial_\mu + ig_s T_s^a G_\mu^a + ig' Y(d_R) B_\mu, \end{aligned} \quad (1.10)$$

where the previously defined coupling constants and gauge bosons have appeared, along with the gauge group generators:  $\mathbf{T}^a$  and  $\mathbf{Y}$  for  $SU(2) \times U(1)_Y$ , and  $\mathbf{T}_s^a$  for  $SU(3)$ . Furthermore the parameter  $Y(f)$ , the weak hypercharge of fermion  $f$ , has been introduced:

$$Y(f) \equiv 2(Q(f) - I_Z(f)), \quad (1.11)$$

where  $Q(f)$  is the electric charge of  $f$ , and  $I_Z(f)$  is the projection of the weak isospin operator (which is  $+1/2$  for  $u_L$  and  $\nu_L$ ,  $-1/2$  for  $d_L$  and  $e_L$  and  $0$  for right-handed fermions).

So the interaction terms between the left-handed leptons and the gauge bosons are

$$-\frac{g}{2} \begin{pmatrix} \bar{\nu}_L \\ \bar{e}_L \end{pmatrix}^T \gamma^\mu \left[ \begin{pmatrix} W_\mu^3 & W_\mu^1 + iW_\mu^2 \\ W_\mu^1 - iW_\mu^2 & -W_\mu^3 \end{pmatrix} + \frac{g'}{g} B_\mu \right] \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}. \quad (1.12)$$

The physical gauge bosons for  $SU(2) \times U(1)_Y$  are found by requiring that there be two massive charged bosons ( $W^\pm$ ), one massive neutral boson ( $Z^0$ ) and one massless neutral boson ( $\gamma$ ), and that there be no mixing between the gauge bosons. This leads to the definitions

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2) \quad (1.13)$$

for  $W_\mu^\pm$ , and (introducing the Weinberg angle  $\theta_W$ , defined by  $\tan \theta_W \equiv \frac{g'}{g}$ ):

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \equiv \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \quad (1.14)$$

for  $Z_\mu$  and  $A_\mu$  (the photon field). So the interaction terms as a function of the physical gauge bosons are:

$$-\frac{g}{2} \begin{pmatrix} \overline{\nu}_L \\ \overline{e}_L \end{pmatrix}^T \gamma^\mu \begin{pmatrix} Z_\mu / \cos \theta_W & \sqrt{2} W_\mu^- \\ \sqrt{2} W_\mu^+ & Z_\mu \cos \theta_W - 2A_\mu \sin \theta_W \end{pmatrix} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}. \quad (1.15)$$

This demonstrates that  $\nu - \nu$  interactions can only proceed via  $Z^0$  exchange, while for  $e - e$  interactions a photon can also be exchanged. The charged-current interactions between  $e$  and  $\nu$ , mediated by the  $W^\pm$ , also appear. Repeating the exercise for the quarks shows the form of the strong force interaction terms.

### 1.1.3.5 Fermion Mass Terms

Attempting to introduce explicit mass terms for the fermions breaks the gauge invariance of the theory, since the terms mix chirality (handedness), having the form  $m\overline{\psi}\psi = m\overline{\psi}_L\psi_R + m\overline{\psi}_R\psi_L$ . So the mass terms are not  $SU(2)$  invariant. This problem can be solved via the introduction of an  $SU(2)$  doublet  $\Phi$ , which is the Higgs doublet [9]. This allows the generation of gauge-invariant mass terms, via the Yukawa interaction:

$$\mathcal{L}_{\text{Yukawa}} = -Y_f \overline{f}_L \Phi f_R + h.c., \quad (1.16)$$

where  $Y_f$  is the Yukawa coupling constant for fermion  $f$ , and  $h.c.$  stands for the Hermitian conjugate of the first term. The potential for the Higgs takes the form:

$$V(\Phi) = -\mu^2 \Phi^* \Phi + \lambda |\Phi^* \Phi|^2, \quad (1.17)$$

which has a minimum at  $\Phi^* \Phi = \frac{1}{2} \mu^2 / \lambda$ . Since the minimum is not at zero, the Higgs doublet acquires a non-zero vacuum expectation value (or ‘‘vev’’). In the unitary gauge, the vev resides only in one part of the doublet, so that:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \mu / \sqrt{\lambda} \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.18)$$

Given this expectation value,  $\Phi$  can be decomposed as:

$$\Phi = \frac{e^{i(\omega_a \mathbf{T}^a - \omega_3 \mathbf{Y})}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \quad (1.19)$$

where the  $\mathbf{T}^a$  and  $\mathbf{Y}$  are the generators of  $SU(2) \times U(1)_Y$ , for some real constants  $\omega_i$ , whose value depends on the choice of gauge. In the unitary gauge, the exponential is rotated away, so that the leptonic part of (1.16) becomes:

$$\mathcal{L}_{\text{Yukawa}} = -\frac{Y_e}{\sqrt{2}} \begin{pmatrix} \overline{\nu}_L & \overline{e}_L \end{pmatrix} \begin{pmatrix} 0 \\ v + H \end{pmatrix} e_R + h.c., \quad (1.20)$$

so that the electron acquires a mass term of the form

$$\frac{Y_e}{\sqrt{2}}v(\bar{e}_L e_R + \bar{e}_R e_L) = \frac{Y_e}{\sqrt{2}}v\bar{e}e = m_e\bar{e}e, \quad (1.21)$$

with the mass being proportional to both the Yukawa coupling and the vev. Repeating this for the other fermion fields, the fermion mass Lagrangian for the first generation particles is

$$\mathcal{L}_{\text{Yukawa}} = -Y_e \bar{l}_L \Phi e_R - Y_d \bar{q}_L \Phi d_R - Y_u \epsilon_{ij} \bar{q}_L \Phi^* u_R + h.c., \quad (1.22)$$

where  $\epsilon_{ij}$ , the two-dimensional antisymmetric tensor, acts on  $q_L$  to allow the up quark to acquire a mass.

### 1.1.3.6 Higgs Term and Gauge Boson Masses

The final ingredient for the one-generation SM Lagrangian is the Higgs term. Since the Higgs is a scalar, its Lagrangian takes the form of (1.1):

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} &= |\mathbf{D}_\mu \Phi|^2 - \mu^2 \Phi^* \Phi + \lambda (\Phi^* \Phi)^2 \\ &= \frac{1}{2}(\partial_\mu H)^2 + \mu^2 H^2 + \frac{g^2 v^2}{4} W^{+\mu} W_\mu^- + \frac{g^2 v^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu + \text{interaction terms} \\ &= \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}m_H^2 H^2 + m_W^2 W^{+\mu} W_\mu^- + \frac{1}{2}m_Z^2 Z_\mu Z^\mu + \text{interaction terms}, \end{aligned} \quad (1.23)$$

where the ‘‘interaction terms’’ describe interactions between more than two bosons, e.g. two gauge bosons and the Higgs. So the gauge bosons have acquired masses that depend on the coupling constants, and are proportional to the vev.

## 1.1.4 Adding Further Generations

In this section the consequences of adding a second fermion generation will be outlined, before this is extended to the third generation. The third generation is of particular interest for this thesis as it contains the bottom ( $b$ ) quark, which forms the  $B$  hadrons whose study is the main focus of the LHCb physics programme.

### 1.1.4.1 Adding The Second Generation

The main feature that emerges when a second generation of fermions is added is the phenomenon of mixing between generations. Since testing the SM description of quark mixing will be the purpose of the LHCb experiment, the following will focus on the effect of adding a second generation of quarks. Mixing between leptons, which leads to neutrino oscillations, is well established. This mixing is governed by the Pontecorvo-Maki-Nakagawa-Sakata

(PMNS) matrix [10], which is the leptonic equivalent of the CKM quark mixing matrix that will be introduced in the next section. A discussion of lepton mixing is beyond the scope of this thesis.

The second generation of quarks, the charm and strange quarks, transform under  $SU(2)$  in exactly the same way as the first generation, into a left-handed doublet and two right-handed singlets. The only differences between the up and the charm quarks (and between the down and the strange quarks) are that they possess their own generation-specific quantum numbers (charm for the charm quark, strangeness for the strange quark and isospin<sup>1</sup> for the up and down quarks), and that they have different masses.

When adding the second generation, the Lagrangian described in Sec. 1.1.3 changes for the most part only by the addition of exactly equivalent terms describing the interactions of the new particles with the gauge bosons, and mass terms from their Yukawa coupling. The non-trivial change is that extra Yukawa terms can now be written down which allow the masses to mix across generations, so that the part of the Lagrangian that gives masses to the quarks becomes

$$\mathcal{L}_{\text{quark masses}} = -[Y_d]_{ij} \bar{q}_{Li} \Phi d_{Rj} - [Y_u]_{ij} \bar{q}_{Li} \epsilon_{jk} \Phi^* u_{Rk} + h.c., \quad (1.24)$$

where the Yukawa couplings have become matrices, whose indices  $ij$  run over the generations. If the off-diagonal element of  $[Y_d]$  ( $[Y_u]$ ) is non-zero, then mass mixing arises between the  $d$  and  $s$  ( $u$  and  $c$ ) quarks. Since fermion masses are proportional to the Yukawa couplings (see (1.21)), the mass eigenstates of the quarks cannot have off-diagonal terms acting on them. This can be achieved by introducing a mixing matrix  $V_C$ , which relates the mass eigenstates to the states which couple to the gauge bosons (the flavour eigenstates). The interaction term between quarks and the  $W^\pm$  gauge bosons is then (following (1.9)):

$$\frac{g}{2\sqrt{2}} \begin{pmatrix} \bar{u} \\ \bar{c} \end{pmatrix}^T \gamma^\mu (1 - \gamma^5) V_C \begin{pmatrix} d \\ s \end{pmatrix} W_\mu^- + h.c. . \quad (1.25)$$

To conserve particle number,  $V_C$  has to be unitary. So it takes the form

$$V_C = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix}, \quad (1.26)$$

for some angle  $\theta_C$ , which is known as the Cabibbo angle after Nicola Cabibbo, who first introduced the mixing matrix concept [11].

One important consequence of the above mixing structure is that (for the lowest order “tree” interactions) there are no  $Z$  boson interactions between the (physical)  $d$  and  $s$  quarks,

---

<sup>1</sup>The isospin subgroup of  $SU(3)$  referred to here should not be confused with the concept of weak isospin introduced in Sec. 1.1.3.4.

or in other words no Flavour Changing Neutral Currents (FCNCs). This prediction of the SM is known as the GIM mechanism [12], after Glashow, Iliopoulos and Maiani. The experimental observation of the FCNC decay  $K_L \rightarrow \mu^+ \mu^-$ , which proceeds via a second-order “box” interaction, demonstrated the existence of the charm quark before it had been directly observed. Measuring the rate of the decay allowed an accurate prediction of the charm quark mass, as the box amplitudes would cancel if  $m_u = m_c$ .

### 1.1.4.2 Adding The Third Generation: The CKM Matrix

Incorporating the third generation of quarks, the bottom and the top, in an analogous way to the second generation clearly requires a  $3 \times 3$  unitary matrix, in place of the  $2 \times 2$  unitary matrix  $V_C$ . This matrix is known as the Cabibbo-Kobayashi-Maskawa matrix,  $V_{CKM}$ , after Makoto Kobayashi and Toshihide Maskawa, who first proposed the existence of the third quark generation [13] to explain the phenomenon of  $\mathcal{CP}$  violation (see Sec. 1.1.5), and also after Nicola Cabibbo (see above). In 2008 Kobayashi and Maskawa were awarded the Nobel Prize in Physics for their development of this mechanism.

Analogously to  $V_C$  above,  $V_{CKM}$  relates the flavour eigenstates  $(d', s', b')$  to the mass eigenstates  $(d, s, b)$ :

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.27)$$

The elements of  $V_{CKM}$  are named to reflect the transition between quark flavours that is associated with each element:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (1.28)$$

In principle, if each element  $V_{ij}$  is allowed to be a complex number,  $V_{CKM}$  has eighteen free parameters. However, as stated above,  $V_{CKM}$  must be a unitary matrix. This requirement removes nine of the free parameters. Furthermore, because there are six quarks between which transitions can occur, there are five relative phase transformations that do not correspond to physical observables. These can therefore be rotated away, leaving only four free parameters. These can be represented in many different ways, depending on how the relative phases are rotated away. The most frequently used analytical representation has three mixing angles (generalisations of the Cabibbo angle) and one phase:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \quad (1.29)$$

where  $c_{ij}$  and  $s_{ij}$  represent  $\cos \theta_{ij}$  and  $\sin \theta_{ij}$  respectively. If the phase  $\delta_{13}$  is non-zero, then the behaviour for antiquarks (depending on  $V_{CKM}^\dagger$ ) will differ from the behaviour for quarks, which means that  $\mathcal{CP}$  (see Sec. 1.1.5) will be violated. Section 1.2.2 will show how this  $\mathcal{CP}$  violation manifests itself for neutral mesons, and Sec. 1.3 will describe how this can be measured experimentally.

Although Eqn. (1.29) is useful, it gives no indication of the hierarchy in magnitude that is present in the elements of  $V_{CKM}$ . This hierarchy is illustrated in a non-analytical parameterisation by Wolfenstein [14], which expands each element in powers of  $\lambda \equiv \sin \theta_{12} (= \sin \theta_C)$ :

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.30)$$

The three remaining Wolfenstein parameters are defined as  $A \equiv \sin \theta_{23} / \sin^2 \theta_{12}$  and  $\rho - i\eta \equiv (\sin \theta_{13} \cdot e^{-i\delta_{13}}) / (\sin \theta_{12} \cdot \sin \theta_{23})$ , so that each  $V_{ij}$  is expressed in terms of variables of order 1.

The current world-average experimental values [15] for the magnitude of each element of the CKM matrix are given in Table 1.1. Brief descriptions of the key measurement(s) contributing to each value are also given. Current measurements relating to the phase in the CKM matrix will be discussed in Sec. 1.3.

Parameter	Experimental Value	Key Measurement(s)
$ V_{ud} $	$0.97418 \pm 0.00027$	Pure vector transitions in nuclear beta decay
$ V_{us} $	$0.2255 \pm 0.0019$	Semileptonic kaon decays
$ V_{ub} $	$(3.93 \pm 0.36) \times 10^{-3}$	Inclusive and exclusive $B \rightarrow X_u l \bar{\nu}$ decays
$ V_{cd} $	$0.230 \pm 0.011$	Di-muon production from neutrino beams
$ V_{cs} $	$1.04 \pm 0.06$	Semileptonic $D$ decays
$ V_{cb} $	$(41.2 \pm 1.1) \times 10^{-3}$	Inclusive and exclusive $B \rightarrow X_c l \bar{\nu}$ decays
$ V_{td} $	$(8.1 \pm 0.6) \times 10^{-3}$	$B_d$ oscillation frequency
$ V_{ts} $	$(38.7 \pm 2.3) \times 10^{-3}$	$B_s$ oscillation frequency
$ V_{tb} $	$>0.74$ (95% C.L.)	Top quark decays

**Table 1.1:** World-average experimental values for the magnitude of each element of the CKM matrix.

### 1.1.5 Discrete Symmetries

Symmetries play an incredibly important rôle in nature. By Noether's theorem [16] any symmetry present in the action of a physical theory necessarily leads to a conservation law.

For example the invariance of the equations of motion of a system under translations in space and time gives rise to momentum and energy conservation respectively, while their invariance under phase changes of the wavefunction gives rise to conservation of electric charge. The above instances are examples of continuous symmetries. However discrete symmetries also have a significant rôle to play in nature. Discrete transformations of wavefunctions can be defined as those Lorentz transformations which are not obtainable by continuous deformation of the identity transformation. The discrete transformations relevant to the work presented in this thesis are parity ( $\mathcal{P}$ ), charge conjugation ( $\mathcal{C}$ ) and time reversal ( $\mathcal{T}$ ):

- The parity transformation reverses spatial coordinates, which has the effect of reversing the chirality of a particle.
- Charge conjugation transforms a particle into its antiparticle, changing the sign of internal quantum numbers such as electric charge and lepton number.
- Time reversal is the transformation  $t \rightarrow -t$ , which changes the sign of a particle's linear and angular momentum.

Symmetry under the combined transformation  $\mathcal{CPT}$  has to be conserved in any quantum field theory which respects Lorentz invariance and locality. A key implication of  $\mathcal{CPT}$  symmetry is that any antiparticle must have the same mass and lifetime as its associated particle. The best experimental test of  $\mathcal{CPT}$  symmetry is the constraint on the  $K^0 - \bar{K}^0$  mass difference [15], which is currently  $|(m_{\bar{K}^0} - m_{K^0})/m_{K^0}| < 0.8 \times 10^{-18}$  at 90% confidence level (C.L.). Although  $\mathcal{CPT}$  is always respected in a “sensible” QFT, sub-transformations of  $\mathcal{CPT}$  can be violated, and determining the exact nature of such violations is key to understanding the SM and any New Physics which may exist beyond it. All observations to date indicate that the strong, electromagnetic and gravitational interactions all conserve  $\mathcal{C}$ ,  $\mathcal{P}$  and  $\mathcal{T}$  individually. The only interaction that has been seen to violate any of these is the weak interaction, as was discussed in Sec. 1.1.3.4, where the weak interaction (maximally) violated parity by acting only on left-handed fermion fields.

The phenomenon of  $\mathcal{P}$  violation was first experimentally observed in the  $\beta$ -decay of  $^{60}\text{Co}$  atoms in 1957 [17]. Following this discovery, it was still assumed that the combined symmetry of  $\mathcal{CP}$  was always conserved [18]. However soon after, in 1964,  $\mathcal{CP}$  violation in the decays of neutral kaons was observed by Cronin and Fitch [19]. Direct observation of  $\mathcal{T}$  violation (which, as long as  $\mathcal{CPT}$  is conserved, is equivalent to  $\mathcal{CP}$  violation) was not made until 1998, when the CPLEAR experiment observed a difference between the mixing rates for  $K^0 \rightarrow \bar{K}^0$  and  $\bar{K}^0 \rightarrow K^0$  [20]. Recent observations of  $\mathcal{CP}$  violation in the  $B$  meson system will be discussed in the following sections.

With the concept of  $\mathcal{CP}$  violation having now been introduced, its manifestation in the quark sector can now be discussed. The focus will be on the behaviour of neutral mesons, as these have a richer  $\mathcal{CP}$  phenomenology than either charged mesons or baryons. A study concerning the measurement of one example of this phenomenon will be the subject of chapter 6 of this thesis.

## 1.2 Flavour Physics

The term “flavour physics” is used to describe the interactions between flavours both in the lepton sector and the quark sector. This section will look at some of the flavour phenomenology which arises in the quark sector, focusing on neutral mesons. Section 1.2.1 outlines the formalism used to describe the mixing of neutral mesons, and Sec. 1.2.2 discusses how  $\mathcal{CP}$  violation can manifest itself within this formalism.

The CKM matrix governs the physics of quark mixing, which is probed using studies of hadrons containing at least one quark of the second and third generations. The first such hadrons to be discovered were the charged and neutral kaons, in 1947 [21]. Although existence of the charm quark was predicted by the observation of  $K_L \rightarrow \mu^+ \mu^-$  (see Sec. 1.1.4.1), and the third generation by the observation of  $\mathcal{CP}$  violation (see Sec. 1.1.4.2), the discovery of further quarks took some time, due to the large increase in mass (and thus energy required) compared to the strange quark.

The  $c$  and  $b$  quarks were finally discovered in quick succession in the mid-1970s by the observation of quarkonium resonances ( $q\bar{q}$  states) decaying to  $e^+e^-$  or  $\mu^+\mu^-$ . A  $c\bar{c}$  (charmonium) state with a mass of 3.1 GeV was observed in 1974 by two independent experiments, one [22] at the Stanford Linear Accelerator Laboratory (SLAC) in California, and one [23] at Brookhaven National Laboratory (BNL) in New York state. The  $c\bar{c}$  state produced was named as the  $J$  by BNL and the  $\psi$  by SLAC; today it is usually known as the  $J/\psi$ . The production of a  $b\bar{b}$  (bottomonium) state with a mass of 9.5 GeV followed in 1977 [24] at Fermilab, Illinois. This state was named the  $\Upsilon$  (today known as the  $\Upsilon(1S)$ ).

While the Fermilab and BNL experiments had fired a beam of protons at a metallic fixed target, the SLAC experiment had used an  $e^+ - e^-$  collider (the SPEAR ring). The success of the colliding beam experiment led to the construction, at Cornell University in New York state, of a more powerful  $e^+ - e^-$  accelerator, the CESR (Cornell Electron Storage Ring). In 1985 the CLEO-I detector [25] running at the CESR discovered the  $\Upsilon(4S)$  bottomonium resonance [26], with a mass of 10.58 GeV. Since this is just above twice the mass of a  $B_d$  or  $B^+$  meson, the  $\Upsilon(4S)$  almost always decays to a  $B - \bar{B}$  pair (either  $B_d - \bar{B}_d$  or  $B^+ - B^-$ ). Hence the  $\Upsilon(4S)$  discovery opened the door to high-statistics studies of  $B$  mesons.

These studies were carried out in the 1990s by the upgraded CLEO detector, CLEO-II [27], and in the 2000s by two detectors operating at asymmetric  $e^+ - e^-$  colliders (known as B-Factories). These detectors are the Babar detector [28] at the PEP-II accelerator at SLAC, and the Belle detector [29] at the KEKB accelerator at the High Energy Accelerator Research Organization (KEK) laboratory in Tsukuba, Japan. Some of the discoveries made by these experiments will be mentioned in the remainder of this chapter. The phenomenology of mixing in neutral mesons, such as the  $B_d$ , will now be described.

### 1.2.1 Time Evolution of Neutral Mesons

Neutral mesons containing quarks and antiquarks of different flavour ( $q\bar{q}'$  states), such as neutral  $K$ ,  $D$  and  $B$  mesons, can oscillate between their particle and antiparticle states. This oscillation is not possible for neutral mesons which are (superpositions of)  $q\bar{q}$  states, such as the  $\pi^0$ ,  $\eta$ ,  $\rho$ ,  $\phi$ ,  $J/\psi$  and  $\Upsilon$ , because they are their own antiparticles.

The discovery of  $B_d$  mixing [30] was made by the ARGUS detector [31], which took data (often at the  $\Upsilon(4S)$  resonance) between 1982 and 1992 at the DORIS-II  $e^+ - e^-$  collider at the DESY laboratory near Hamburg. ARGUS was able to infer that the  $B_d$  must undergo oscillations by observing two like-sign leptons from the decay of a  $B_d - \bar{B}_d$  pair, which meant that one of the mesons must have changed flavour between production and decay.

The discovery of  $B_s$  mixing was made by the TeVatron experiments D0 [32] and CDF [33] (the B-Factories do not create large quantities of  $B_s$  mesons<sup>2</sup>) in 2006 [34, 35], with the oscillation frequency being far higher than that for the  $B_d$ .

The state of a neutral meson that is observed experimentally is some linear combination of the particle and antiparticle states, say

$$a|N^0\rangle + b|\bar{N}^0\rangle \equiv \begin{pmatrix} a \\ b \end{pmatrix}. \quad (1.31)$$

The time evolution of this state is described by:

$$i\frac{d}{dt} \left[ \begin{pmatrix} a \\ b \end{pmatrix} \right] = H \begin{pmatrix} a \\ b \end{pmatrix}. \quad (1.32)$$

The matrix  $H$ , which represents the effective Hamiltonian, is not Hermitian (if it were, the mesons would continue oscillating indefinitely and not decay). However, as with any complex matrix, it can be written in terms of two Hermitian matrices:

$$H = M - \frac{i}{2}\Gamma = \begin{pmatrix} m_{11} - \frac{i}{2}\Gamma_{11} & m_{12} - \frac{i}{2}\Gamma_{12} \\ m_{12}^* - \frac{i}{2}\Gamma_{12}^* & m_{11} - \frac{i}{2}\Gamma_{11} \end{pmatrix}, \quad (1.33)$$

---

<sup>2</sup>Belle has carried out several short runs at the  $\Upsilon(5S)$  resonance, which can decay to a  $B_s - \bar{B}_s$  pair, however the statistics are not competitive with those at the TeVatron.

with  $M$  being the mass matrix and  $\Gamma$  the decay matrix. Note that the Hermiticity of  $M$  and  $\Gamma$  ensures that, as required by  $\mathcal{CPT}$  invariance, the masses and lifetimes of the particle and antiparticle states are equal, i.e.  $m_{11} = m_{22}$  and  $\Gamma_{11} = \Gamma_{22}$ . The off-diagonal terms  $m_{12}$  and  $\Gamma_{12}$  arise from flavour-changing transitions with  $|\Delta F| = 2$  (where  $F$  is the flavour quantum number, e.g. beauty). An example of such a transition is  $N^0 - \bar{N}^0$  mixing, where  $N^0$  is a neutral meson. The  $m_{12}$  ( $\Gamma_{12}$ ) term corresponds to virtual (real) intermediate states.

Diagonalising  $H$  yields the eigenvalues

$$\begin{aligned}\lambda_H &= m_H - \frac{i}{2}\Gamma_H, \\ \lambda_L &= m_L - \frac{i}{2}\Gamma_L,\end{aligned}\tag{1.34}$$

where the labels  $H$  for heavy and  $L$  for light have been introduced. The corresponding mass eigenstates, in terms of the particle and antiparticle states, are

$$\begin{aligned}|N_H\rangle &= p|N^0\rangle - q|\bar{N}^0\rangle, \\ |N_L\rangle &= p|N^0\rangle + q|\bar{N}^0\rangle,\end{aligned}\tag{1.35}$$

where conservation of particle number imposes the constraint  $|p|^2 + |q|^2 = 1$ .

The masses of these eigenstates (which are the observable states) are such that the mass difference  $\Delta m \equiv m_H - m_L$  is positive (hence the labels ‘‘heavy’’ and ‘‘light’’), while the lifetime difference  $\Delta\Gamma \equiv \Gamma_H - \Gamma_L$  can be positive or negative. The mass and lifetime differences can also be expressed relative to the average lifetime  $\Gamma \equiv \Gamma_{11} = (\Gamma_H + \Gamma_L)/2$  by defining  $x \equiv \Delta m/\Gamma$  and  $y \equiv \Delta\Gamma/2\Gamma$  respectively.

By solving the eigenvalue equation (1.34), the coefficients  $p$  and  $q$  can be expressed in terms of the elements of  $\Gamma$  and  $M$ :

$$\left(\frac{q}{p}\right)^2 = \frac{m_{12}^* - (i/2)\Gamma_{12}^*}{m_{12} - (i/2)\Gamma_{12}}.\tag{1.36}$$

From Eqn. (1.32), the mass eigenstates evolve according to:

$$\begin{aligned}|N_H(t)\rangle &= e^{-im_H t - \frac{1}{2}\Gamma_H t}|N_H(0)\rangle, \\ |N_L(t)\rangle &= e^{-im_L t - \frac{1}{2}\Gamma_L t}|N_L(0)\rangle.\end{aligned}\tag{1.37}$$

Now, Eqn. (1.35) can be inverted to obtain an expression for a pure  $N^0$  state:

$$|N^0\rangle = \frac{1}{2p}(|N_H\rangle + |N_L\rangle).\tag{1.38}$$

So, using Eqn. (1.37), the time evolution of a state which is pure  $N^0$  at  $t = 0$  will be:

$$|N^0(t)\rangle = \frac{1}{2} \left( \left( e^{-im_H t - \frac{1}{2}\Gamma_H t} + e^{-im_L t - \frac{1}{2}\Gamma_L t} \right) N^0 - \frac{q}{p} \left( e^{-im_H t - \frac{1}{2}\Gamma_H t} - e^{-im_L t - \frac{1}{2}\Gamma_L t} \right) \bar{N}^0 \right).\tag{1.39}$$

The formalism for neutral meson mixing has now been given, allowing a discussion of  $\mathcal{CP}$  violation in these mesons.

## 1.2.2 $\mathcal{CP}$ Violation in Neutral Meson Decays

The instantaneous decay amplitudes of a pseudoscalar meson  $N^0$  and its  $\mathcal{CP}$  conjugate  $\bar{N}^0$  into a multi-particle final state  $f$  and its  $\mathcal{CP}$  conjugate  $\bar{f}$  are given by [36]:

$$\begin{aligned} A_f &= \langle f | \mathcal{H} | N^0 \rangle, & A_{\bar{f}} &= \langle \bar{f} | \mathcal{H} | N^0 \rangle, \\ \bar{A}_f &= \langle f | \mathcal{H} | \bar{N}^0 \rangle, & \bar{A}_{\bar{f}} &= \langle \bar{f} | \mathcal{H} | \bar{N}^0 \rangle, \end{aligned} \quad (1.40)$$

where  $\mathcal{H}$  is the Hamiltonian governing weak interactions.

### 1.2.2.1 Classification of $\mathcal{CP}$ Violating Effects

There are three distinct types of  $\mathcal{CP}$  violation that are possible for mesons. The first type is known as direct  $\mathcal{CP}$  violation, and the second and third types as indirect  $\mathcal{CP}$  violation. The three types are defined in terms of the above amplitudes and (for neutral mesons) the mixing parameters  $p/q$  as follows:

- $\mathcal{CP}$  violation in the decay: this occurs if the instantaneous amplitudes for a decay and its  $\mathcal{CP}$  conjugate differ in magnitude, i.e. if

$$\left| \frac{\bar{A}_{\bar{f}}}{A_f} \right| \neq 1. \quad (1.41)$$

This is the only source of  $\mathcal{CP}$  violation for charged mesons such as the  $B^+$ . The first observation of direct  $\mathcal{CP}$  violation was made in 1999 by the NA48 experiment at CERN and the KTeV experiment at Fermilab, using neutral kaon decays to two pions [37, 38]. Direct  $\mathcal{CP}$  violation in the  $B$  sector was observed for the first time by the B-Factories [39, 40] in 2004, in the decay  $B_d \rightarrow K^+ \pi^-$ .

- $\mathcal{CP}$  violation in mixing: this occurs if the mass eigenstates of a neutral meson are not  $\mathcal{CP}$  eigenstates, causing the rates  $N^0 \rightarrow \bar{N}^0$  and  $\bar{N}^0 \rightarrow N^0$  to be unequal. This corresponds to the condition

$$\left| \frac{q}{p} \right| \neq 1. \quad (1.42)$$

This was the first type of  $\mathcal{CP}$  violation ever observed (see sec. 1.1.5), with the discovery of the decay  $K_L \rightarrow \pi^+ \pi^-$ . Such an observation, of a decay that is known to come from a particular mass eigenstate, is possible in the neutral kaon sector because the mass eigenstates,  $K_L$  and  $K_S$ , have very different lifetimes:  $\tau_{K_L} = 5 \times 10^{-8}$  s and  $\tau_{K_S} = 9 \times 10^{-11}$  s (hence ‘‘L’’ standing for long, and ‘‘S’’ for short). So a neutral kaon that is observed to survive much longer than, say,  $10^{-9}$  s must be a  $K_L$  and not a  $K_S$ . Under the assumption of  $\mathcal{CP}$  conservation, these mass eigenstates would also

be  $\mathcal{CP}$  eigenstates, so that for example  $K_L$  would be a purely  $\mathcal{CP}$ -odd state, and so could decay to the  $\mathcal{CP}$ -odd final states  $\pi^+\pi^-\pi^0$  and  $\pi^0\pi^0\pi^0$ , but not to the  $\mathcal{CP}$ -even final states  $\pi^+\pi^-$  and  $\pi^0\pi^0$ . However the decay  $K_L \rightarrow \pi^+\pi^-$  was observed in 1964 by Cronin and Fitch [19], demonstrating that  $\mathcal{CP}$  could be violated. The branching ratio was small,  $\text{B.R.}(K_L \rightarrow \pi^+\pi^-) = 2 \times 10^{-3}$ , indicating that  $\mathcal{CP}$  violation was a small effect in the kaon sector.

Much larger  $\mathcal{CP}$  violating effects are possible in the  $B$  sector, as will be seen in Sec. 1.3. However,  $\mathcal{CP}$  violation *in mixing* for neutral  $B$  mesons is predicted in the SM to be proportional to  $(m_c/m_t)^2$ , and therefore to be very small. It can be measured experimentally in semileptonic decays of neutral  $B$  mesons (see below), as such decays are flavour-specific and no direct  $\mathcal{CP}$  violation is expected. A combination of measurements from the B-factories [41, 42] gives [43]

$$\left| \frac{q}{p} \right|_{B_d} = 1.0002 \pm 0.0028, \quad (1.43)$$

which is in agreement with the SM prediction. No precise corresponding measurement has yet been made in the  $B_s$  sector.

- $\mathcal{CP}$  violation in interference between mixing and decay: even if neither of the above two types of  $\mathcal{CP}$  violation occur, there can (for neutral mesons) still be  $\mathcal{CP}$  violation if the final state  $f$  is accessible to both  $M$  and  $\bar{M}$ . This is because the processes  $M \rightarrow f$  and  $M \rightarrow \bar{M} \rightarrow f$  share the same initial and final states, and so can interfere quantum mechanically. This interference is described by the quantity

$$\lambda_f \equiv \frac{q \bar{A}_f}{p A_f}. \quad (1.44)$$

$\mathcal{CP}$  violation arises from this interference if the condition

$$\Im(\lambda_f) \neq 0 \quad (1.45)$$

holds. This type of  $\mathcal{CP}$  violation is often referred to as mixing-induced  $\mathcal{CP}$  violation. A strategy to extract both the direct and mixing-induced  $\mathcal{CP}$  asymmetries ( $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ ) in the decay  $B_d \rightarrow \pi^+\pi^-$  at LHCb will be the subject of Chapter 6 of this thesis.

### 1.2.2.2 Time-Dependent $\mathcal{CP}$ Violating Asymmetries

The appearance of all three of the above  $\mathcal{CP}$  effects can occur only for the case where a neutral meson decays into a final state which is accessible to both the meson and the antimeson. Here the case of a generic final state will be considered at first, with the resulting relations later being simplified for the case of a flavour-specific final state.

Using (1.39), the time-dependent decay amplitudes of a pure initial  $N^0$  or  $\bar{N}^0$  state into a final state  $f$  can be found in terms of the instantaneous amplitudes (1.40):

$$\begin{aligned}\langle f|\mathcal{H}|N^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}(A_f \cosh((ix+y)\Gamma t/2) - \frac{q}{p}\bar{A}_f \sinh((ix+y)\Gamma t/2)), \\ \langle f|\mathcal{H}|\bar{N}^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}(\bar{A}_f \cosh((ix+y)\Gamma t/2) - \frac{p}{q}A_f \sinh((ix+y)\Gamma t/2)),\end{aligned}\quad (1.46)$$

where  $m \equiv m_{11} = (m_H + m_L)/2$  is the average mass. The observable decay rates are proportional to the squared magnitude of these amplitudes. The decay rates are normally written by introducing the asymmetry observables defined by:

$$\mathcal{A}_{\mathcal{CP}}^{\text{dir}} = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \quad \mathcal{A}_{\mathcal{CP}}^{\text{mix}} = \frac{2\Im(\lambda_f)}{1 + |\lambda_f|^2}, \quad A_{\Delta\Gamma} = \frac{2\Re(\lambda_f)}{1 + |\lambda_f|^2}, \quad (1.47)$$

which can be seen to obey the relation

$$(\mathcal{A}_{\mathcal{CP}}^{\text{dir}})^2 + (\mathcal{A}_{\mathcal{CP}}^{\text{mix}})^2 + (A_{\Delta\Gamma})^2 = 1. \quad (1.48)$$

With these definitions, the decay rates are:

$$\begin{aligned}\Gamma(N^0(t) \rightarrow f) &\propto |A_f|^2(1 + |\lambda_f|^2)e^{-\Gamma t} [\cosh(\Delta\Gamma t/2) - A_{\Delta\Gamma} \sinh(\Delta\Gamma t/2) \\ &\quad + \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta mt) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta mt)], \\ \Gamma(\bar{N}^0(t) \rightarrow f) &\propto |A_f|^2 \left| \frac{p}{q} \right|^2 (1 + |\lambda_f|^2)e^{-\Gamma t} [\cosh(\Delta\Gamma t/2) - A_{\Delta\Gamma} \sinh(\Delta\Gamma t/2) \\ &\quad - \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta mt) + \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta mt)],\end{aligned}\quad (1.49)$$

and the  $\mathcal{CP}$  asymmetry is given by (assuming  $\left| \frac{p}{q} \right| = 1$ ):

$$\mathcal{A}_{\mathcal{CP}}(t) \equiv \frac{\Gamma(N^0 \rightarrow f) - \Gamma(\bar{N}^0 \rightarrow f)}{\Gamma(N^0 \rightarrow f) + \Gamma(\bar{N}^0 \rightarrow f)} = \frac{\mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta mt) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta mt)}{\cosh(\Delta\Gamma t/2) - A_{\Delta\Gamma} \sinh(\Delta\Gamma t/2)}. \quad (1.50)$$

If  $\Delta\Gamma$  is assumed to be small (this is an excellent approximation for  $B_d$ , but less so for  $B_s$ ), then the decay rate simplifies to:

$$\begin{aligned}\Gamma(N^0(t) \rightarrow f) &\propto |A_f|^2(1 + |\lambda_f|^2)e^{-\Gamma t} [1 + \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta mt) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta mt)], \\ \Gamma(\bar{N}^0(t) \rightarrow f) &\propto |A_f|^2 \left| \frac{p}{q} \right|^2 (1 + |\lambda_f|^2)e^{-\Gamma t} [1 - \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta mt) + \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta mt)],\end{aligned}\quad (1.51)$$

and the  $\mathcal{CP}$  asymmetry to (again, assuming  $\left| \frac{p}{q} \right| = 1$ ):

$$\mathcal{A}_{\mathcal{CP}}(t) = \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta mt) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta mt). \quad (1.52)$$

The above relations simplify considerably for the case where  $f$  is flavour-specific, so that the instantaneous amplitude for  $\bar{N}^0$  is  $\bar{A}_f = 0$ . This implies  $\lambda_f = 0$ , so the decay rates (1.49) become

$$\begin{aligned}\Gamma(N^0(t) \rightarrow f) &\propto |A_f|^2 e^{-\Gamma t} [\cosh(\Delta\Gamma t/2) + \cos(\Delta m t)], \\ \Gamma(\bar{N}^0(t) \rightarrow f) &\propto |A_f|^2 \left| \frac{p}{q} \right|^2 e^{-\Gamma t} [\cosh(\Delta\Gamma t/2) - \cos(\Delta m t)],\end{aligned}\quad (1.53)$$

with the corresponding rates for  $\bar{f}$  (with  $A_{\bar{f}} = 0$ ) being:

$$\begin{aligned}\Gamma(N^0(t) \rightarrow \bar{f}) &\propto |\bar{A}_{\bar{f}}|^2 \left| \frac{q}{p} \right|^2 e^{-\Gamma t} [\cosh(\Delta\Gamma t/2) - \cos(\Delta m t)], \\ \Gamma(\bar{N}^0(t) \rightarrow \bar{f}) &\propto |\bar{A}_{\bar{f}}|^2 e^{-\Gamma t} [\cosh(\Delta\Gamma t/2) + \cos(\Delta m t)].\end{aligned}\quad (1.54)$$

So the rates for  $N^0(t) \rightarrow f$  and  $\bar{N}^0(t) \rightarrow \bar{f}$  differ only by the time-independent factor  $|\bar{A}_{\bar{f}}/A_f|$ . If  $\left| \frac{p}{q} \right| = 1$ , the same is true for  $\bar{N}^0(t) \rightarrow f$  and  $N^0(t) \rightarrow \bar{f}$ .

The above decay rate expressions for flavour-specific final states can be combined in different ways to measure different physics parameters. For example, semileptonic decays of neutral  $B$  mesons (e.g.  $B_d \rightarrow D^{(*)-} \mu^+ \nu_\mu$  or  $B_s \rightarrow D_s^- \mu^+ \nu_\mu$ ) are predicted to have  $\bar{A}_{\bar{f}} = A_f$ , not only in the SM but in most extensions of it (as the decay proceeds via a tree amplitude only). Hence  $\left| \frac{p}{q} \right|$  can be measured from these decays, using the so-called semileptonic asymmetry  $\mathcal{A}_{\text{SL}}$ :

$$\mathcal{A}_{\text{SL}} \equiv \frac{\Gamma(N^0 \rightarrow \bar{f}) - \Gamma(\bar{N}^0 \rightarrow f)}{\Gamma(N^0 \rightarrow \bar{f}) + \Gamma(\bar{N}^0 \rightarrow f)} = \frac{1 - |A_f/\bar{A}_{\bar{f}}|^2 |p/q|^4}{1 + |A_f/\bar{A}_{\bar{f}}|^2 |p/q|^4} = \frac{1 - |p/q|^4}{1 + |p/q|^4}.\quad (1.55)$$

Measurements of  $\mathcal{A}_{\text{SL}}$  using  $B_d \rightarrow D^{(*)-} \mu^+ \nu_\mu$  decays provide the result for  $\left| \frac{p}{q} \right|_{B_d}$  given in the previous section.

Another combination of the flavour-specific decay rates is the charge asymmetry  $\mathcal{A}_{\mathcal{CP}}$ . If  $\left| \frac{p}{q} \right|$  is taken to be 1, then  $|A_f/\bar{A}_{\bar{f}}|$  for a flavour-specific decay can be measured:

$$\mathcal{A}_{\mathcal{CP}} \equiv \frac{(\Gamma(N^0 \rightarrow f) + \Gamma(\bar{N}^0 \rightarrow f)) - (\Gamma(N^0 \rightarrow \bar{f}) + \Gamma(\bar{N}^0 \rightarrow \bar{f}))}{(\Gamma(N^0 \rightarrow f) + \Gamma(\bar{N}^0 \rightarrow f)) + (\Gamma(N^0 \rightarrow \bar{f}) + \Gamma(\bar{N}^0 \rightarrow \bar{f}))} = \frac{|A_f/\bar{A}_{\bar{f}}|^2 - 1}{|A_f/\bar{A}_{\bar{f}}|^2 + 1}.\quad (1.56)$$

Note that  $\mathcal{A}_{\mathcal{CP}}$  can (unlike  $\mathcal{A}_{\text{SL}}$ ) be measured without initial state tagging, by simply counting the number of decays seen to each final state<sup>3</sup>. Measurements of  $\mathcal{A}_{\mathcal{CP}}$  in the channel  $B_d \rightarrow K^+ \pi^-$  gave the first observation of direct  $\mathcal{CP}$  violation in  $B$  decays.

<sup>3</sup>Of course in a real experiment there will be factors such as imperfect tagging of the initial state flavour, detection asymmetries, etc., which have to be taken into account.

## 1.3 Testing the CKM Mechanism with $\mathcal{CP}$ Violation Measurements

As discussed in Sec. 1.1.4.2,  $\mathcal{CP}$  violation in the quark sector is only generated in the presence of a non-zero phase in the CKM matrix. The  $\mathcal{CP}$  violating effects are normally visualised using so-called unitarity triangles, which will now be described.

### 1.3.1 The Unitarity Triangles

The unitarity of the CKM matrix gives rises to nine relations between the different elements:

$$\sum_i V_{ij} V_{ik}^* = \delta_{jk} \quad \text{and} \quad \sum_j V_{ij} V_{kj}^* = \delta_{ik}. \quad (1.57)$$

The six cases where the Kronecker delta evaluates to zero can be represented as triangles in the complex plane. Each triangle has the same area, which parameterises the amount of  $\mathcal{CP}$  violation that exists in the SM. Four of these triangles have sides which differ in order of magnitude, and are often known as “squashed”. The remaining two triangles have sides which are the same order of magnitude ( $\propto \lambda^3$ ) as each other. They are defined by the following unitarity relations:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (1.58)$$

and

$$V_{td}V_{ud}^* + V_{ts}V_{us}^* + V_{tb}V_{ub}^* = 0. \quad (1.59)$$

The relation (1.58) defines the more commonly used triangle, which is often referred to simply as “the” unitarity triangle. It is normally illustrated with each side divided by  $V_{cd}V_{cb}^*$ , so that its vertices lie at  $(0, 0)$ ,  $(1, 0)$  and  $(\bar{\rho}, \bar{\eta})$ , where  $\bar{\rho} \equiv \rho(1 - \lambda^2/2)$  and  $\bar{\eta} \equiv \eta(1 - \lambda^2/2)$ . The triangle defined by (1.59) is identical to “the” unitarity triangle up to corrections of order  $\lambda^3$ . These two “non-squashed” unitarity triangles are illustrated in Fig 1.1.

Constraints are placed on the unitarity triangle via measurements of the lengths of its sides and the size of its angles. The length of one side is 1 by definition, with the lengths of the other two sides (referred to as  $R_b$  and  $R_t$ ) being given by

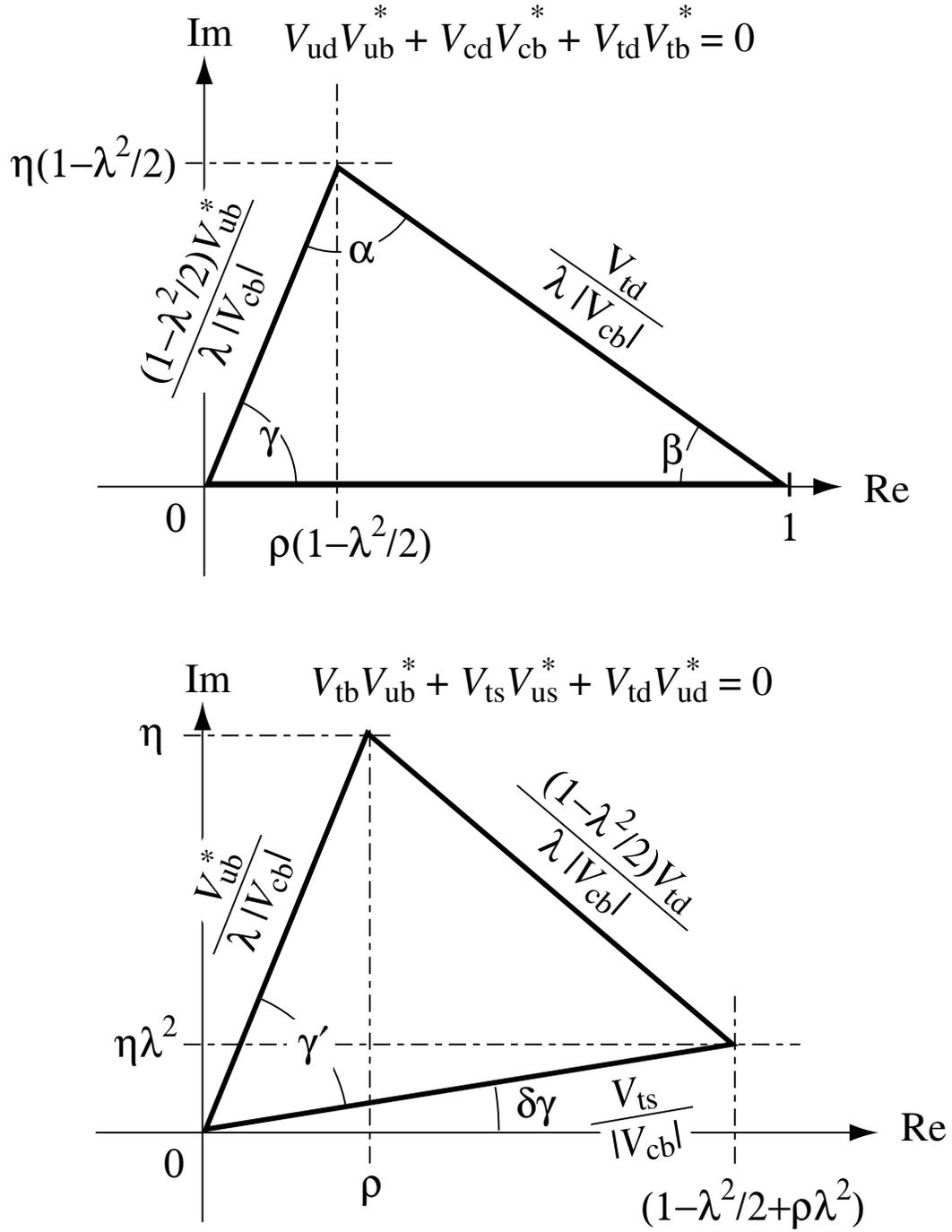
$$R_b = \frac{(1 - \lambda^2/2)|V_{ub}|}{\lambda|V_{cb}|}, \quad R_t = \frac{|V_{td}|}{\lambda|V_{cb}|}. \quad (1.60)$$

The angles are referred to<sup>4</sup> as  $\alpha$ ,  $\beta$  and  $\gamma$ , with the following definitions:

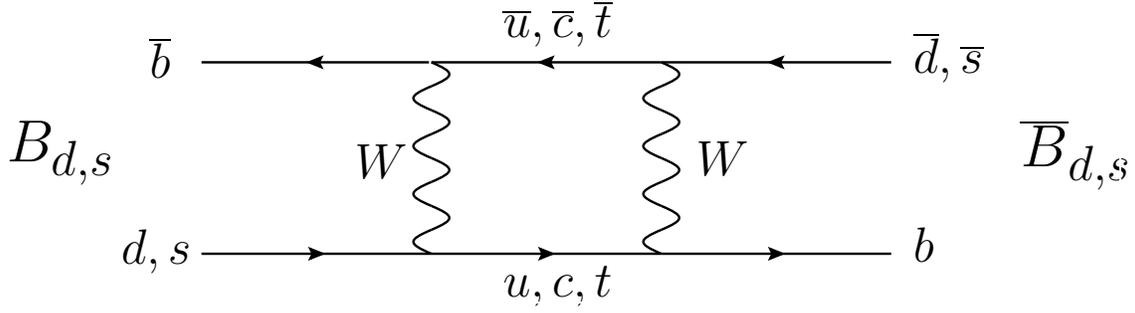
$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \quad (1.61)$$

---

<sup>4</sup>This is the notation used by the Babar experiment, as well as by the TeVatron and LHC experiments. However the Belle experiment uses a different notation  $\phi_{1,2,3}$ , where  $\phi_1 = \beta$ ,  $\phi_2 = \alpha$  and  $\phi_3 = \gamma$ .



**Figure 1.1:** Illustration of the two “non-squashed” CKM unitarity triangles. Reproduced from [44].



**Figure 1.2:** The main Feynman diagram contributing to  $B_{d,s} - \bar{B}_{d,s}$  mixing.

The fact that these angles are large, while those from the squashed unitarity triangles are very small, indicates that, in the SM,  $\mathcal{CP}$  violation can be an  $\mathcal{O}(1)$  effect for  $B$  mesons, but should be a much smaller effect for other mesons, such as  $K$  and  $D$ . Each of the above angles has been measured, to differing degrees of precision, via studies of  $B$  hadrons. Examples of such studies, and the current constraints on the unitarity triangles, will now be given.

### 1.3.2 Constraints on the Unitarity Triangles

An illustrative example of how studies of  $B$  mesons can constrain the unitarity triangles is the measurement of the  $B_d$  and  $B_s$  mixing phases  $\phi_d$  and  $\phi_s$ , which are defined by:

$$\phi_d \equiv \arg \left( \frac{q}{p} \right)_{B_d}, \quad \phi_s \equiv \arg \left( \frac{q}{p} \right)_{B_s}. \quad (1.62)$$

To see the link between these phases and the CKM angles, consider the main Feynman diagram contributing to  $B_{d,s} - \bar{B}_{d,s}$  mixing, which is shown in Fig. 1.2<sup>5</sup>. For both  $B_d$  and  $B_s$ , the mixing phase can be found using the approximation  $\Delta\Gamma \ll \Delta m$ , so that (from (1.36))

$$\frac{p}{q} = \sqrt{\frac{m_{12}^*}{m_{12}}} + \mathcal{O} \left( \frac{\Gamma_{12}}{m_{12}} \right) \simeq \sqrt{\frac{m_{12}^*}{m_{12}}}. \quad (1.63)$$

The dominant contribution to the amplitude for the diagram in Fig.1.2 comes from the loop with an internal top quark, since the contributions are proportional to  $m_q^2$  [46]. So for  $B_d$  the mixing phase is approximately  $-2\beta$ :

$$\left( \frac{p}{q} \right)_{B_d} \simeq \frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*} = e^{-2i\beta}, \quad (1.64)$$

and the  $B_s$  mixing phase is approximately  $-2\delta\gamma$ :

$$\left( \frac{p}{q} \right)_{B_s} \simeq \frac{V_{ts}V_{tb}^*}{V_{tb}V_{ts}^*} = e^{-2i\delta\gamma}, \quad (1.65)$$

<sup>5</sup>All Feynman diagrams in this thesis were typeset by the author, using the JaxoDraw package [45].

where

$$\delta\gamma \simeq \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \quad (1.66)$$

is the angle<sup>6</sup> between the bottom side of the second unitarity triangle and the real axis (see Fig. 1.1). This indicates that measuring mixing-induced  $\mathcal{CP}$  violation in  $B_d$  ( $B_s$ ) decays can allow  $\beta$  ( $\delta\gamma$ ) to be measured. Indeed the measurement of the mixing-induced  $\mathcal{CP}$  violation term in the decay  $B_d \rightarrow J/\psi K_S$  by the B-Factories marked the discovery of the first  $\mathcal{CP}$  violation seen in  $B$  decays [47, 48]. Analogously,  $\delta\gamma$  can be measured via the mixing-induced  $\mathcal{CP}$  violation in  $B_s \rightarrow J/\psi\phi$ . Recent studies of this decay at the TeVatron [49, 50] indicate that  $\phi_s$ , which is very small in the SM, may in fact be sizable<sup>7</sup>. If confirmed this would constitute an unambiguous signal for New Physics. Improving the measurement of  $\phi_s$  is a priority for the LHCb physics programme [51].

The current world-average values for each angle [52, 53], obtained without using constraints from measurements of the other angles or of the sides of the unitarity triangle<sup>8</sup>, are given in Table 1.2.

Parameter	Value	Key Measurement(s)
$\alpha$	$(89.0^{+4.5}_{-4.2})^\circ$	Time-dependent $\mathcal{CP}$ asymmetries in $B_d \rightarrow \rho\pi$
$\beta$	$(21.1 \pm 0.9)^\circ$	Time-dependent $\mathcal{CP}$ asymmetries in $B_d \rightarrow J/\psi K_S$
$\gamma$	$(70^{+27}_{-30})^\circ$	GLW, ADS and Dalitz analyses of $B^+ \rightarrow D^{(*)0} K^{(*)+}$

**Table 1.2:** World-average experimental values, from direct measurements, for the angles of the CKM unitarity triangle.

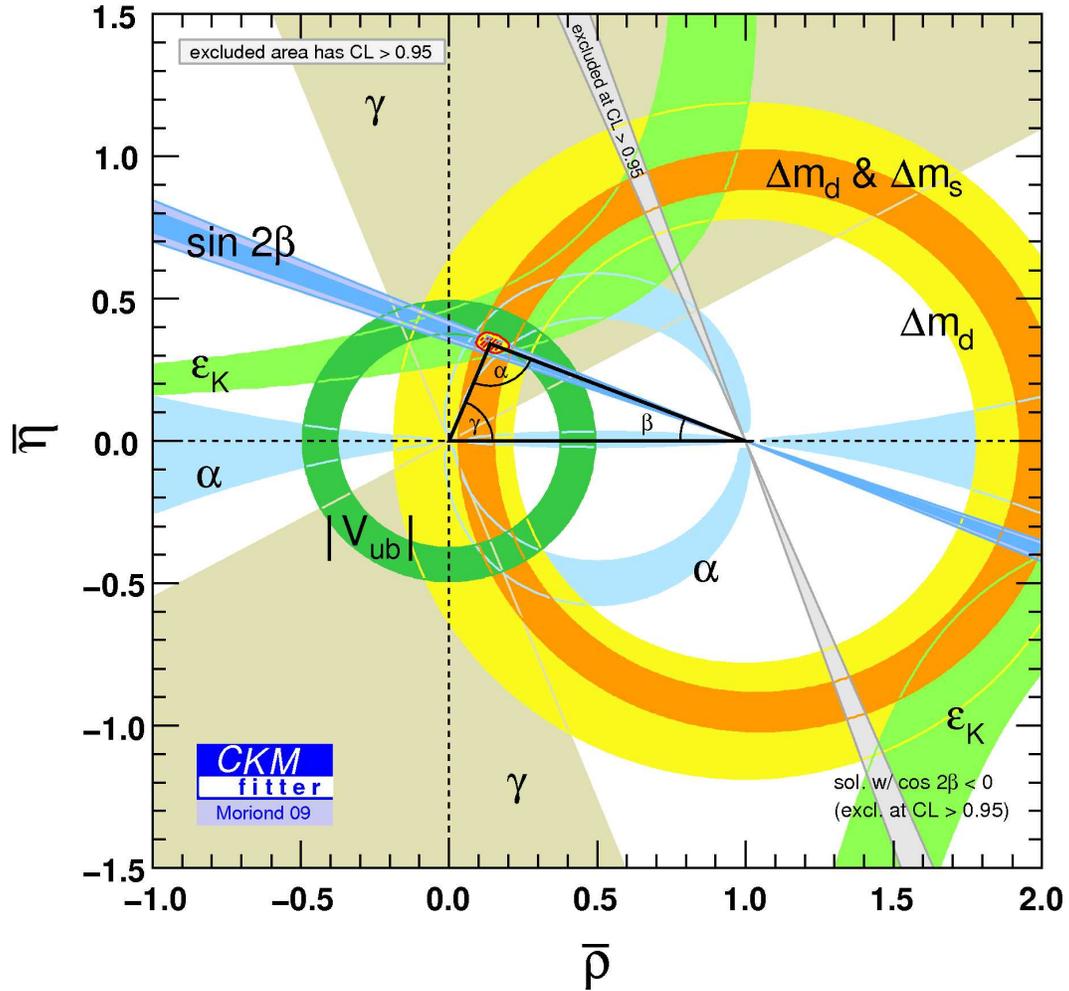
A global fit of the unitarity triangle can be made [53] that combines all available information to constrain the position of the vertex at  $(\bar{\rho}, \bar{\eta})$ . An illustration of the current status of this global fit is given in Fig. 1.3. The resulting constraints on the angles from this global fit are given in Table 1.3. The main difference between these values and those from direct measurements is a drastic decrease in the uncertainty on  $\gamma$ , by a factor of around 6.

The consistency of all of the measurements shown in Fig. 1.3 demonstrates that the many new measurements made by the B-Factories have built up a consistent description of the unitarity triangle. This indicates that the CKM mechanism is the dominant source of  $\mathcal{CP}$  violation in the quark sector, but currently there is still room for a contribution from New Physics,

<sup>6</sup>In the literature the notation  $\beta_s$  is sometimes used in place of  $\delta\gamma$ .

<sup>7</sup>There is inconsistency in the literature concerning the quantity referred to as  $\phi_s$ . To avoid ambiguities, LHCb has adopted [51] a notation  $\Phi_{J/\psi\phi} \equiv -\arg(\eta_{J/\psi\phi}\lambda_{J/\psi\phi})$ , where  $\eta_{J/\psi\phi} = \pm 1$  is the  $\mathcal{CP}$  eigenvalue of the final state. In the SM, the observable  $\Phi_{J/\psi\phi}$  is equal to  $-2\delta\gamma$ .

<sup>8</sup>However ambiguities in translating from trigonometric functions to angles have been removed following the assumption that the sum of the angles should be close to  $180^\circ$ .



**Figure 1.3:** Global fit of the apex of the CKM unitarity triangle. The red hashed region around the apex corresponds to 68% C.L. Reproduced from [53].

Parameter	Value from Global Fit
$\alpha$	$(90.6^{+3.8}_{-4.2})^\circ$
$\beta$	$(21.6^{+0.9}_{-0.8})^\circ$
$\gamma$	$(67.8^{+4.2}_{-3.9})^\circ$

**Table 1.3:** Values for the angles of the CKM unitarity triangle from a global fit for the apex of the triangle.

especially in the poorly-constrained  $B_s$  sector, where mixing-induced  $\mathcal{CP}$  violation in  $B_s$  mixing can cause small (in the SM) differences between the two “non-squashed” unitarity triangles. One such difference arises from  $\delta\gamma$ , which can be measured using decays which proceed dominantly or wholly via  $b \rightarrow c\bar{c}s$  tree transitions. Since these are not expected to be affected by New Physics, any deviation from the SM can be traced to  $B_s$  mixing. Recent interesting studies of one such decay,  $B_s \rightarrow J/\psi\phi$ , were mentioned above.

Overconstraining the parameters of the CKM matrix to search for inconsistencies in the SM predictions is a key physics goal of the LHCb experiment. The LHCb strategy to search for New Physics via direct measurements of the most poorly constrained angle of the “main” unitarity triangle,  $\gamma$ , will be discussed in Sec. 1.4.4.

## 1.4 Two-body Charmless Hadronic $B$ Decays

In this section some of the theoretical aspects underlying two-body charmless hadronic  $B$  decays will be outlined, and two of the physics measurements that can be carried out by studying such decays will be described in detail. The expected performance of LHCb for each of these measurements will be summarised.

The description “two-body charmless hadronic  $B$  decay” can in principle apply to any decay of a  $B$  hadron into two hadrons which do not contain charm quarks. Within LHCb however, it is usually used to refer only to cases where the daughter hadrons are charged pions, charged kaons, and protons. These are the only hadrons that can be tracked in particle detectors, because they are “stable” over the typical time required to pass through the detector, and are charged. Decays of this type are referred to within LHCb as  $B \rightarrow h^+h'^-$  decays. There are ten such decays in total<sup>9</sup>:  $B_d \rightarrow K^+\pi^-$ ,  $B_d \rightarrow \pi^+\pi^-$ ,  $B_d \rightarrow K^-K^+$ ,  $B_d \rightarrow p\bar{p}$ ,  $B_s \rightarrow K^-K^+$ ,  $B_s \rightarrow K^-\pi^+$ ,  $B_s \rightarrow \pi^+\pi^-$ ,  $B_s \rightarrow p\bar{p}$ ,  $\Lambda_b \rightarrow pK^-$  and  $\Lambda_b \rightarrow p\pi^-$ . The first of these to be observed was  $B_d \rightarrow K^+\pi^-$  [54], in 1993 by the CLEO-II experiment [27]. Since then, further modes have been observed, and currently the only remaining unobserved  $B \rightarrow h^+h'^-$  modes are  $B_d \rightarrow K^-K^+$ ,  $B_s \rightarrow \pi^+\pi^-$ ,  $B_d \rightarrow p\bar{p}$  and  $B_s \rightarrow p\bar{p}$ .

<sup>9</sup>Not distinguishing between a decay and its charge conjugate.

Of course many two-body charmless hadronic  $B$  decays exist that are not  $B \rightarrow h^+h'^-$  decays (for example  $B_s \rightarrow \phi\phi$ ), but the physics measurements that can be made using them differ from those made using  $B \rightarrow h^+h'^-$  decays, and they are also very different from an experimental point of view, as they have more than two charged tracks, can feature intermediate resonances, and so on. They are not considered further here.

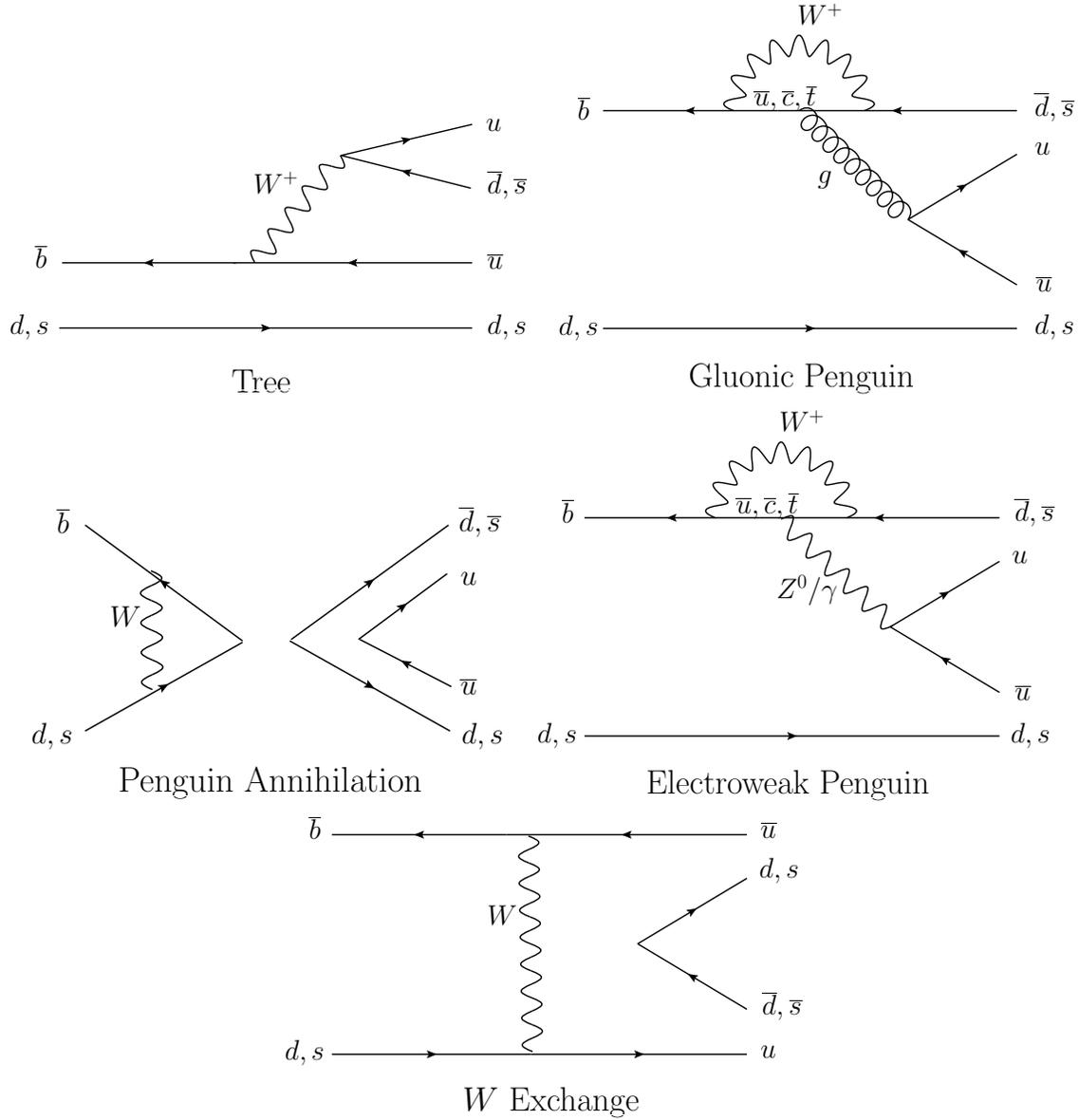
The LHCb collaboration has recently published a detailed document [55] describing how six of its key measurements physics will be carried out. One of these key measurements involves  $B \rightarrow h^+h'^-$  decays [56]. Parts of the  $B \rightarrow h^+h'^-$  section of this document summarise the work described in chapters 4 (proper time resolution model for  $B \rightarrow h^+h'^-$  decays) and 5 (search for undiscovered modes  $B_{\{d,s\}} \rightarrow p\bar{p}$ ) of this thesis. The detailed theoretical aspects relating to the work in those chapters will be outlined at the beginning of those respective chapters. Chapter 6 of this thesis describes a method to make a physics measurement with one particular  $B \rightarrow h^+h'^-$  decay,  $B_d \rightarrow \pi^+\pi^-$ , using an approach which is complementary to that used in [56]. The theory of this measurement is described in Sec. 1.4.3, and that of a related measurement also involving  $B_d \rightarrow \pi^+\pi^-$ , in Sec. 1.4.4.

### 1.4.1 Overview of $B \rightarrow h^+h'^-$ Decays

Of the ten  $B \rightarrow h^+h'^-$  decays listed above, the six of the form  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  offer a particularly rich array of possibilities for physics measurements. The physics measurements that can be made using a particular one of these decays largely depends on the nature of the final state, with  $\pi^+\pi^-$  and  $K^-K^+$  being  $\mathcal{CP}$  eigenstates with eigenvalue  $+1$  (i.e.  $\mathcal{CP}|\pi^+\pi^-\rangle = +|\pi^+\pi^-\rangle$  and  $\mathcal{CP}|K^-K^+\rangle = +|K^-K^+\rangle$ ), and  $\pi^\pm K^\mp$  being flavour-specific. The Feynman diagrams for the processes contributing to the  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  decays are shown in Fig. 1.4.

The tree and gluonic penguin<sup>10</sup> processes are the dominant contributors to the four  $B \rightarrow h^+h'^-$  decays with the highest branching ratios (referred to here as the four ‘‘main’’  $B \rightarrow h^+h'^-$  decays), which are  $B_d \rightarrow K^+\pi^-$ ,  $B_d \rightarrow \pi^+\pi^-$ ,  $B_s \rightarrow K^-K^+$  and  $B_s \rightarrow K^-\pi^+$ . Their branching ratios are of the order of  $10^{-5}$ . These ‘‘main’’ decays can also receive contributions from the other three diagrams shown in Fig. 1.4, with the colour-suppressed electroweak penguin, penguin annihilation and  $W$  exchange all contributing to  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^-K^+$ , and the colour-suppressed electroweak penguin contributing to  $B_d \rightarrow K^+\pi^-$  and  $B_s \rightarrow K^-\pi^+$ . However in the SM the size of these contributions is negligible compared to the two dominant digrams (see e.g. [57]).

<sup>10</sup>A process involving particle emission from an internal loop is referred to as a penguin process.



**Figure 1.4:** Feynman diagrams for the processes contributing to  $B_{d,s} \rightarrow \{\pi, K\}^+ \{\pi, K\}^-$  decays.

However for the modes  $B_d \rightarrow K^- K^+$  and  $B_s \rightarrow \pi^+ \pi^-$ , only the penguin annihilation and  $W$  exchange diagrams are possible, causing the SM branching ratios for these decays to be at least an order of magnitude lower than those of the main  $B \rightarrow h^+ h'^-$  decays [57]. LHCb will search for these undiscovered rare modes [58], as a measurement of their branching ratios will test the SM prediction that the annihilation and exchange amplitudes are small. A significant enhancement of their branching ratios above the SM predictions would indicate the presence of New Physics in the penguin annihilation and  $W$  exchange diagrams.

An important concept relating certain  $B \rightarrow h^+ h'^-$  decays to each other is that of *U-spin symmetry* [46]. Just as isospin symmetry relates to a subgroup of  $SU(3)$  corresponding to the interchange of  $u$  and  $d$  quarks, U-spin symmetry relates to a different subgroup of  $SU(3)$ , which corresponds to the interchange of  $s$  and  $d$  quarks. The decays  $B_d \rightarrow \pi^+ \pi^-$  and  $B_s \rightarrow K^- K^+$  are full U-spin conjugates, as the interchange of  $s$  quarks for  $d$  quarks transforms  $B_d \rightarrow \pi^+ \pi^-$  into  $B_s \rightarrow K^- K^+$ , and each amplitude contributing to  $B_d \rightarrow \pi^+ \pi^-$  also contributes to  $B_s \rightarrow K^- K^+$  (and vice versa). So if U-spin symmetry is respected,  $B_d \rightarrow \pi^+ \pi^-$  and  $B_s \rightarrow K^- K^+$  share the same strong interaction dynamics.

Note that the decays  $B_d \rightarrow \pi^+ \pi^-$  and  $B_s \rightarrow K^- \pi^+$  are *not* full U-spin conjugates: although they are related via the interchange  $d \leftrightarrow s$ , there are amplitudes (from penguin annihilation and  $W$  exchange) which contribute to  $B_d \rightarrow \pi^+ \pi^-$  but not  $B_s \rightarrow K^- \pi^+$ . The situation is exactly the same for  $B_s \rightarrow K^- K^+$  and  $B_d \rightarrow K^+ \pi^-$ . Although they are not full U-spin conjugates, these two pairs of decays still offer a method for testing U-spin symmetry, which predicts  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(B_d \rightarrow \pi^+ \pi^-) = \mathcal{A}_{\mathcal{CP}}(B_s \rightarrow K^- \pi^+)$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(B_s \rightarrow K^- K^+) = \mathcal{A}_{\mathcal{CP}}(B_d \rightarrow K^+ \pi^-)$ , up to small corrections arising from the penguin annihilation and  $W$  exchange topologies. A further test of U-spin symmetry can be made by comparing the branching ratios of the full U-spin conjugate pair  $B_d \rightarrow K^- K^+$  and  $B_s \rightarrow \pi^+ \pi^-$ .

While there are many interesting physics measurements that can be made using  $B \rightarrow h^+ h'^-$  decays, the remainder of this chapter will focus on two particular measurements, which are linked in that they both rely on  $\mathcal{CP}$  asymmetry measurements in the decay  $B_d \rightarrow \pi^+ \pi^-$ .

### 1.4.2 Time Dependent $\mathcal{CP}$ Asymmetries in $B_d \rightarrow \pi^+ \pi^-$

A procedure to extract direct and mixing-induced  $\mathcal{CP}$  asymmetries from a fit to the time-dependent  $\mathcal{CP}$  asymmetry in the decay  $B_d \rightarrow \pi^+ \pi^-$  is described in Chapter 6 of this thesis. This section describes the current experimental status of these asymmetries. The following two sections each describe a physics measurement that LHCb plans to make which requires a precise measurement of these asymmetries.

Since

- $\pi^+\pi^-$  is a  $\mathcal{CP}$  eigenstate,
- The width difference  $\Delta\Gamma$  in the  $B_d$  system is known to be very small compared to the mass difference  $\Delta m$ , and
- $|p/q|_{B_d}$  is known to be equal to 1 to a very good approximation,

the relevant equation for the time-dependent  $\mathcal{CP}$  asymmetry for  $B_d \rightarrow \pi^+\pi^-$  is (1.52):

$$\begin{aligned} \mathcal{A}_{\mathcal{CP}}(B_d(t) \rightarrow \pi^+\pi^-) &\equiv \frac{\Gamma(B_d \rightarrow \pi^+\pi^-) - \Gamma(\bar{B}_d \rightarrow \pi^+\pi^-)}{\Gamma(B_d \rightarrow \pi^+\pi^-) + \Gamma(\bar{B}_d \rightarrow \pi^+\pi^-)} \\ &= \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}} \cos(\Delta mt) - \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}} \sin(\Delta mt). \end{aligned} \quad (1.67)$$

The current B-Factory measurements of the direct and mixing-induced  $\mathcal{CP}$  asymmetries in  $B_d \rightarrow \pi^+\pi^-$  are [59, 60]:

$$\begin{aligned} \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}} &= 0.25 \pm 0.08 \pm 0.02, \quad \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}} = 0.68 \pm 0.10 \pm 0.03 \quad (\text{Babar}), \\ \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}} &= 0.55 \pm 0.08 \pm 0.05, \quad \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}} = 0.61 \pm 0.10 \pm 0.04 \quad (\text{Belle}), \end{aligned} \quad (1.68)$$

where the first error is statistical and the second systematic. While there is agreement between the B-Factories on the value of the mixing-induced asymmetry, there is significant disagreement on the value of the direct asymmetry, leading to a  $1.9\sigma$  discrepancy between the two experiments in the  $\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}} - \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}}$  plane. The situation is illustrated in Fig. 1.5, which also displays the average of the two measurements (as calculated by HFAG):

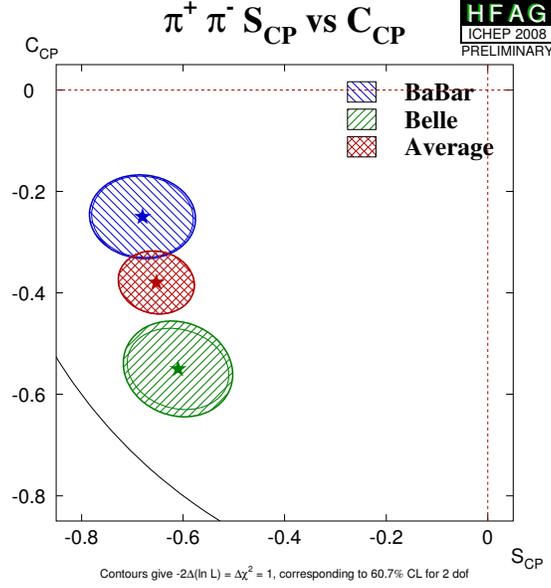
$$\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}} = 0.38 \pm 0.06, \quad \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}} = 0.65 \pm 0.07 \quad (\text{HFAG average}), \quad (1.69)$$

where only the total error is quoted.

LHCb will resolve the conflict between the B-Factories by making a higher-precision measurement of these asymmetries. Aside from the experimental issues, an improved measurement of these asymmetries will also allow better determinations of some important CKM parameters, via the methods described in the following sections.

### 1.4.3 Measuring $\sin 2\alpha_{\text{eff}}$ Using the Decay $B_d \rightarrow \pi^+\pi^-$

The only way to directly measure the CKM angle  $\alpha$ , which is the phase between  $V_{td}V_{tb}^*$  and  $V_{ud}V_{ub}^*$ , is through measurements of time-dependent  $\mathcal{CP}$  asymmetries affecting  $b \rightarrow u\bar{u}d$  transitions. The decay  $B_d \rightarrow \pi^+\pi^-$  features such a transition, and can be used to measure the effective parameter  $\sin 2\alpha_{\text{eff}}$ . This measurement can be combined with information from



**Figure 1.5:** Current B-Factory measurements of the direct and mixing-induced  $\mathcal{CP}$  asymmetries in the  $B_d \rightarrow \pi^+\pi^-$  decay. The definitions  $S_{\mathcal{CP}} \equiv -\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  and  $C_{\mathcal{CP}} \equiv -\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  are used. The average of the two measurements (as calculated by HFAG) is also shown. Figure reproduced from [52].

other decays (as detailed below) to constrain  $\sin 2\alpha$ . The method used for this measurement will now be described.

A general form for the amplitude  $A_{\pi^+\pi^-}$  for  $B_d \rightarrow \pi^+\pi^-$  (following the notation of [61]) is:

$$A_{\pi^+\pi^-} = V_{ud}V_{ub}^*(A_{\text{tree}} + A_{\text{peng}}^u + A_{\text{exch}}) + V_{cd}V_{cb}^*A_{\text{peng}}^c + V_{td}V_{tb}^*A_{\text{peng}}^t, \quad (1.70)$$

where  $A_{\text{tree(exch)}}$  is the tree ( $W$  exchange) amplitude with the weak phase factorised out, and  $A_{\text{peng}}^j$  (for  $j \in \{u, c, t\}$ ) is the part of penguin amplitude corresponding to an internal quark of flavour  $j$ , again with the relevant weak phase factorised out (each  $A_{\text{peng}}^j$  includes the electroweak, gluonic and annihilation penguin amplitudes). The  $\mathcal{CP}$ -conjugated amplitude  $\overline{A}_{\pi^+\pi^-}$  is obtained from  $A_{\pi^+\pi^-}$  by conjugating the CKM elements.

To examine the relationship between these amplitudes and  $\alpha$ , it is convenient to factorise the magnitudes (but not the phases) of the CKM elements back into the amplitudes, and to use the unitarity of the CKM matrix to rewrite  $A_{\pi^+\pi^-}$  and  $\overline{A}_{\pi^+\pi^-}$  as [46]:

$$\begin{aligned} A_{\pi^+\pi^-} &= e^{i\gamma}T + e^{-i\beta}P \equiv e^{i\gamma}|T|e^{i\delta_T} + e^{-i\beta}|P|e^{i\delta_P}, \\ \overline{A}_{\pi^+\pi^-} &= e^{-i\gamma}T + e^{i\beta}P \equiv e^{-i\gamma}|T|e^{i\delta_T} + e^{i\beta}|P|e^{i\delta_P}, \end{aligned} \quad (1.71)$$

where  $T \equiv |V_{ud}||V_{ub}|(A_{\text{tree}} + A_{\text{peng}}^u + A_{\text{exch}} - A_{\text{peng}}^c)$  and  $P \equiv |V_{td}||V_{tb}|(A_{\text{peng}}^t - A_{\text{peng}}^c)$  are

(under the assumptions outlined in Sec. 1.4.1) dominated by the tree and gluonic penguin contributions respectively, with magnitudes  $|T|$  and  $|P|$  and strong phases  $\delta_T$  and  $\delta_P$ .

From these amplitudes,  $\lambda_{\pi^+\pi^-}$  (see definition (1.44)) can be found:

$$\begin{aligned}\lambda_{\pi^+\pi^-} &\equiv \left(\frac{q}{p}\right) \left(\frac{\bar{A}_{\pi^+\pi^-}}{A_{\pi^+\pi^-}}\right) = e^{-2i\beta} \frac{|T|e^{-i\gamma}e^{i\delta_T} + |P|e^{i\beta}e^{i\delta_P}}{|T|e^{i\gamma}e^{i\delta_T} + |P|e^{-i\beta}e^{i\delta_P}} \\ &= e^{-2i(\beta+\gamma)} \frac{1 + \left|\frac{P}{T}\right| e^{i(\beta+\gamma)} e^{i(\delta_P-\delta_T)}}{1 + \left|\frac{P}{T}\right| e^{-i(\beta+\gamma)} e^{i(\delta_P-\delta_T)}} \\ &= e^{2i\alpha} \left(1 + 2i \left|\frac{P}{T}\right| \sin \alpha e^{i(\delta_P-\delta_T)} + \mathcal{O}\left|\frac{P}{T}\right|^2\right),\end{aligned}\quad (1.72)$$

where it has been assumed that  $\alpha + \beta + \gamma = \pi$  and  $|P/T|^2 \ll 1$  (the relative sizes of  $|P|$  and  $|T|$  will be discussed below). Using the definitions (1.47), the direct and mixing-induced  $\mathcal{CP}$  asymmetries can now be related to the physics parameters as follows:

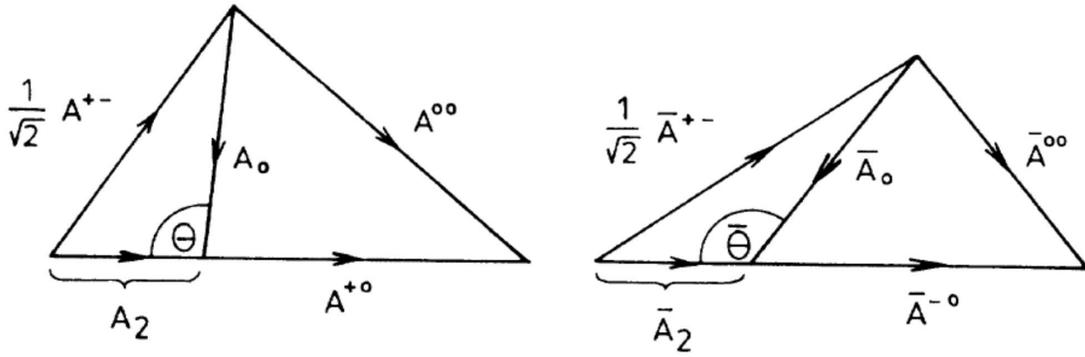
$$\begin{aligned}\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}} &= 2 \left|\frac{P}{T}\right| \sin \alpha \sin(\delta_P - \delta_T) + \mathcal{O}\left|\frac{P}{T}\right|^2, \\ \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}} &= \sin 2\alpha + 2 \left|\frac{P}{T}\right| \sin \alpha \cos 2\alpha \cos(\delta_P - \delta_T) + \mathcal{O}\left|\frac{P}{T}\right|^2.\end{aligned}\quad (1.73)$$

So if the contribution to the amplitude from the penguin topology were negligible, the mixing-induced  $\mathcal{CP}$  asymmetry would equal  $\sin 2\alpha$ , and the direct  $\mathcal{CP}$  asymmetry would be 0. The analogous condition (negligible penguin amplitude) is satisfied for the decay  $B_d \rightarrow J/\psi K_S$ , where the mixing-induced  $\mathcal{CP}$  asymmetry is equal to  $\sin 2\beta$  to a very good approximation. However for  $B_d \rightarrow \pi^+\pi^-$ , the penguin topology contribution is *not* expected to be negligible in the SM, with theoretical calculations pointing to a value of around 0.2 for  $|P/T|$  (see e.g. [62]). Hence the value of the measured quantity  $\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}}$  is significantly different from the physics parameter  $\sin 2\alpha$ . Instead it is related to an effective physics parameter, known as  $\sin 2\alpha_{\text{eff}}$ , through the definition

$$\sin 2\alpha_{\text{eff}} \equiv \frac{\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}}}{\sqrt{1 - \mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}2}}}. \quad (1.74)$$

Clearly the relative magnitudes and strong phases of the tree and penguin amplitudes —  $|P/T|$  and  $(\delta_P - \delta_T)$  — need to be calculated if a precise measurement of  $\sin 2\alpha_{\text{eff}}$  is to be translated into a precise measurement of  $\sin 2\alpha$ . One method for calculating these involves combining measurements of the branching ratio and  $\mathcal{CP}$  asymmetries for  $B_d \rightarrow \pi^+\pi^-$  with those for the decays  $B^+ \rightarrow \pi^+\pi^0$  and  $B_d \rightarrow \pi^0\pi^0$  [63]. These three decays are related via isospin transformations, with their amplitudes forming two so-called “isospin triangles” in the complex plane, defined by

$$\sqrt{2}A(B^+ \rightarrow \pi^+\pi^0) = A(B_d \rightarrow \pi^+\pi^-) + \sqrt{2}A(B_d \rightarrow \pi^0\pi^0) \quad (1.75)$$



**Figure 1.6:** Isospin triangles relating the amplitudes for  $B_d \rightarrow \pi^+\pi^-$ ,  $B^+ \rightarrow \pi^+\pi^0$  and  $B_d \rightarrow \pi^0\pi^0$ . The triangle on the left (right) represents Eqn. 1.75 (1.76). The amplitudes with superscripts are labelled using the charge of the final state pions; e.g.  $A^{+0} \equiv A(B^+ \rightarrow \pi^+\pi^0)$  and  $\bar{A}^{+-} \equiv A(\bar{B}_d \rightarrow \pi^+\pi^-)$ . The amplitudes with subscripts are the amplitudes for a  $B$  or  $\bar{B}$  to decay to a  $\pi\pi$  pair with isospin 0 or 2 respectively. Reproduced from [63].

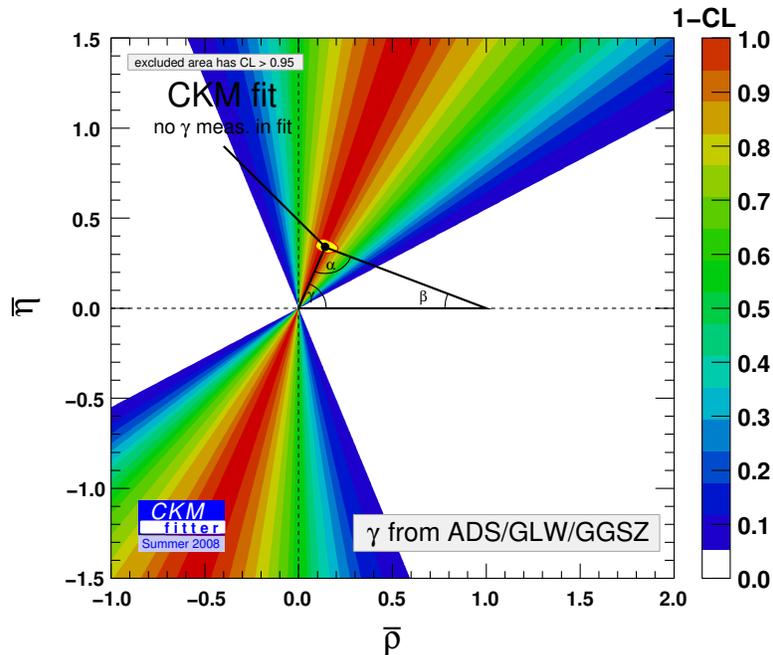
and

$$\sqrt{2}A(B^- \rightarrow \pi^-\pi^0) = A(\bar{B}_d \rightarrow \pi^+\pi^-) + \sqrt{2}A(\bar{B}_d \rightarrow \pi^0\pi^0). \quad (1.76)$$

These isospin triangles are illustrated in Fig. 1.6. The fact that  $\pi\pi$  final states arising from tree processes can have isospin  $I = 0$  or  $I = 2$ , while those arising from penguin processes can only have  $I = 0$ , allows the extraction of  $|P/T|$  and  $(\delta_P - \delta_T)$ . Measuring the branching ratios of the  $B \rightarrow \pi\pi$  modes (i.e. the  $\mathcal{CP}$ -averaged) lengths of the triangle sides) gives  $|P/T|$ , and measuring the  $\mathcal{CP}$  asymmetries (i.e. the relative phase between the triangles) gives  $(\delta_P - \delta_T)$ . The current measurements of  $\alpha$  using the above method are  $\alpha = (96_{-6}^{+10})^\circ$  from Babar [59] and  $\alpha = (97 \pm 11)^\circ$  from Belle [60].

Clearly, improving on these constraints using the decay  $B_d \rightarrow \pi^+\pi^-$  also requires improved measurements in the decays  $B^+ \rightarrow \pi^+\pi^0$  and  $B_d \rightarrow \pi^0\pi^0$ . However, the presence of the neutral particle  $\pi^0$  in the final state will make it difficult for LHCb to improve on the current B-Factor measurements in  $B^+ \rightarrow \pi^+\pi^0$  and  $B_d \rightarrow \pi^0\pi^0$ . This is especially true for  $B_d \rightarrow \pi^0\pi^0$ , which will be extremely problematic for the LHCb trigger (see Sec. 2.2.7) to select. A further experimental difficulty is that the  $B_d \rightarrow \pi^0\pi^0$  tree amplitude is colour-suppressed, leading to a branching ratio of  $\mathcal{O}(10^{-6})$ , an order of magnitude lower than those for  $B_d \rightarrow \pi^+\pi^-$  and  $B^+ \rightarrow \pi^+\pi^0$ .

Despite the obstacles to an accurate determination of  $\alpha$  at LHCb using the above method, it will still provide a valuable cross-check to measurements of  $\alpha$  in other decays at LHCb, for example  $B \rightarrow \rho\pi$  and  $B \rightarrow \rho\rho$ , which are expected (with  $2 \text{ fb}^{-1}$  of data) to individually achieve statistical uncertainties on  $\alpha$  of around  $10^\circ$  ( $B \rightarrow \rho\pi$ ) and  $15^\circ$  ( $B \rightarrow \rho\rho$ )



**Figure 1.7:** Constraint on the apex of the CKM unitarity triangle from direct measurements of  $\gamma$ . The SM solution has  $\gamma > 0$ . Reproduced from [53].

respectively [64]. While (most of) the  $\rho\pi$ ,  $\rho\rho$  decay modes feature at least one  $\pi^0$  in the final state, the additional presence of at least two charged particles means that such modes will have a far better trigger efficiency at LHCb than  $B^+ \rightarrow \pi^+\pi^0$  and (especially)  $B_d \rightarrow \pi^0\pi^0$ .

In addition to contributing to the constraint on  $\alpha$ , a measurement of  $\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}(\pi^+\pi^-)}^{\text{mix}}$  is an important stepping stone to an accurate and theoretically clean measurement of  $\gamma$ , as described in the following section.

#### 1.4.4 Measuring $\gamma$ Using $B_d \rightarrow \pi^+\pi^-$ and $B_s \rightarrow K^-K^+$

##### 1.4.4.1 LHCb Strategy for Direct Measurements of $\gamma$

New Physics contributions to quark mixing could manifest themselves by causing the observed value of a particular CKM parameter to change significantly when measured using different decays. There is scope for this to happen for the CKM angle  $\gamma$ , as current constraints from direct measurements are very poor, as illustrated in Fig. 1.7.

LHCb will make direct measurements of  $\gamma$  using two different classes of  $B$  decay: those that proceed via tree amplitudes only, and those in which both tree and penguin amplitudes play a significant rôle. The aim is not only to tighten the constraint on  $\gamma$ , but also to search for inconsistencies in the measured values of  $\gamma$  in different decays, which would reveal the

presence of New Physics. Such inconsistencies can appear because the value of  $\gamma$  measured from “tree-only” decays is unlikely to be affected by New Physics, as in most NP models the new particles do not enter at tree level. However for decays where penguin amplitudes play a significant rôle, new virtual particles can appear in the loop, affecting the penguin amplitude and thus the measured value of  $\gamma$  for such decays.

Examples of decays that are sensitive to  $\gamma$  and proceed only via tree amplitudes are  $B^+ \rightarrow \bar{D}^0 K^+$ ,  $B_d \rightarrow \bar{D}^0 K^{*0}$ ,  $B_s \rightarrow D_s^\mp K^\pm$  and  $B_d \rightarrow D^\pm \pi^\mp$  [65]. For  $B^+ \rightarrow \bar{D}^0 K^+$  and  $B_d \rightarrow \bar{D}^0 K^{*0}$ ,  $\gamma$  can be extracted via a time-integrated analysis. For  $B_s \rightarrow D_s^\mp K^\pm$  and  $B_d \rightarrow D^\pm \pi^\mp$ , a time-dependent analysis is required. Combining measurements from these four decays, LHCb expects to achieve an uncertainty on the direct measurement of  $\gamma$  of  $\simeq 4^\circ$  after one nominal year of data taking [65].

Two decays that are sensitive to  $\gamma$  and proceed by both tree and penguin amplitudes are the  $B \rightarrow h^+ h'^-$  decays  $B_d \rightarrow \pi^+ \pi^-$  and  $B_s \rightarrow K^- K^+$ . The method that LHCb will use to measure  $\gamma$  using these decays, and the expected sensitivity of the method, will now be described.

#### 1.4.4.2 Extracting $\gamma$ Using the U-Spin Method

A method was proposed by Fleischer in [61] to extract  $\gamma$  from measurements of the direct and mixing-induced  $\mathcal{CP}$  asymmetries in  $B_d \rightarrow \pi^+ \pi^-$  and  $B_s \rightarrow K^- K^+$ . The method is an attractive one because it does not neglect any of the different topologies contributing to each decay, and so does not rely on any model-dependent assumptions about their relative sizes. The only assumption made, using the fact that these two decays are exact U-spin conjugates of each other, is that the final-state strong interaction effects are the same for each decay. Violations of this assumption due to possible U-spin breaking effects are the only source of theoretical error in the method.

The amplitude for  $B_d \rightarrow \pi^+ \pi^-$ , as given in (1.70), can be rewritten in terms of the Wolfenstein parameters (see (1.30)) to give [61]:

$$A_{\pi^+ \pi^-} = \mathcal{C}(e^{i\gamma} - de^{i\theta}), \quad (1.77)$$

where

$$\mathcal{C} \equiv A\lambda^3 R_b (A_{\text{tree}} + A_{\text{peng}}^u - A_{\text{peng}}^t + A_{\text{exch}}) \quad (1.78)$$

and

$$de^{i\theta} \equiv \frac{1}{R_b} \frac{A_{\text{peng}}^c - A_{\text{peng}}^t}{A_{\text{tree}} + A_{\text{peng}}^u - A_{\text{peng}}^t + A_{\text{exch}}}. \quad (1.79)$$

Starting from a parameterisation analogous to (1.70), the amplitude for  $B_s \rightarrow K^- K^+$  can be written in a similar way:

$$A_{K^+K^-} = \frac{\lambda}{1 - \lambda^2/2} \mathcal{C}' (e^{i\gamma} + ((1 - \lambda^2)/\lambda^2) d' e^{i\theta'}), \quad (1.80)$$

where

$$\mathcal{C}' \equiv A\lambda^3 R_b (A'_{\text{tree}} + A'_{\text{peng}} - A'_{\text{peng}} + A'_{\text{exch}}) \quad (1.81)$$

and

$$d' e^{i\theta'} \equiv \frac{1}{R_b} \frac{A'_{\text{peng}} - A'_{\text{peng}}}{A'_{\text{tree}} + A'_{\text{peng}} - A'_{\text{peng}} + A'_{\text{exch}}}. \quad (1.82)$$

The primes in the expressions for  $B_s \rightarrow K^- K^+$  indicate that the transition concerned is  $\bar{b} \rightarrow u \bar{u} \bar{s}$  rather than  $\bar{b} \rightarrow u \bar{u} \bar{d}$ . The  $\mathcal{CP}$ -conjugated amplitudes  $\bar{A}_{\pi^+\pi^-}$  and  $\bar{A}_{K^+K^-}$  are obtained from  $A_{\pi^+\pi^-}$  and  $A_{K^+K^-}$  respectively by conjugating the weak phase  $\gamma$ , but leaving the strong phases unchanged. There is also a multiplicative factor of +1 in each case, because both final states are  $\mathcal{CP}$ -even. From these amplitudes,  $\lambda_{\pi^+\pi^-}$  and  $\lambda_{K^+K^-}$  can be calculated, allowing the direct and mixing-induced  $\mathcal{CP}$  asymmetries (see the definitions (1.47)) to be expressed in terms of  $\gamma$ ,  $\phi_d$ ,  $\phi_s$ ,  $d$ ,  $d'$ ,  $\theta$  and  $\theta'$ :

$$\begin{aligned} \mathcal{A}_{\mathcal{CP}}^{\text{dir}}(B_d \rightarrow \pi^+\pi^-) &= \frac{-2d \sin \theta \sin \gamma}{1 - 2d \cos \theta \cos \gamma + d^2}, \\ \mathcal{A}_{\mathcal{CP}}^{\text{mix}}(B_d \rightarrow \pi^+\pi^-) &= \frac{2d \cos \theta \sin(\phi_d + \gamma) - \sin(\phi_d + 2\gamma) - d^2 \sin \phi_d}{1 - 2d \cos \theta \cos \gamma + d^2}, \\ \mathcal{A}_{\mathcal{CP}}^{\text{dir}}(B_s \rightarrow K^-K^+) &= \frac{2\tilde{d}' \sin \theta' \sin \gamma}{1 + 2\tilde{d}' \cos \theta' \cos \gamma + \tilde{d}'^2}, \\ \mathcal{A}_{\mathcal{CP}}^{\text{mix}}(B_s \rightarrow K^-K^+) &= -\frac{2\tilde{d}' \cos \theta' \sin(\phi_s + \gamma) + \sin(\phi_s + 2\gamma) + \tilde{d}'^2 \sin \phi_s}{1 - 2\tilde{d}' \cos \theta' \cos \gamma + \tilde{d}'^2}, \end{aligned} \quad (1.83)$$

where the abbreviation  $\tilde{d}' \equiv ((1 - \lambda^2)/\lambda^2)d'$  has been used. Note that  $A_{\Delta\Gamma}$  is not considered here as the method assumes the constraint (1.48).

The  $B_d$  mixing phase,  $\phi_d \simeq -2\beta$ , is already well measured by the B-Factories, via the measurement of the mixing-induced  $\mathcal{CP}$  asymmetry in  $B_d \rightarrow J/\psi K_S$ . The  $B_s$  mixing phase,  $\phi_s \simeq -2\delta\gamma$ , is currently only loosely constrained by D0 and CDF using the decay  $B_s \rightarrow J/\psi\phi$ , but LHCb will make a precise measurement of  $\phi_s$  with relatively little data, using this same decay [51]. Hence both  $\phi_d$  and  $\phi_s$  can be considered as known quantities. This however still leaves the system of equations underconstrained, as there are four equations with five unknowns:  $d$ ,  $d'$ ,  $\theta$ ,  $\theta'$  and  $\gamma$ .

The application of U-spin symmetry can solve this problem:  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^-K^+$  share the same strong interaction dynamics under U-spin symmetry, so that  $d = d'$  and  $\theta = \theta'$  (even if U-spin symmetry is broken, the only corrections to these two

equalities will be non-factorisable corrections, which are expected to be small [61]). Hence applying the assumption of U-spin symmetry reduces the number of unknowns to three, leaving the system of equations overconstrained. Either of the conditions  $d = d'$  or  $\theta = \theta'$  can then be released when fitting to the experimental data, providing a test of U-spin symmetry (however this degrades the sensitivity to  $\gamma$ ). The sizes of U-spin symmetry breaking effects are usually parameterised in terms of the variables  $\xi \equiv d'/d$  and  $\Delta\theta \equiv \theta' - \theta$ .

In [56], the LHCb sensitivity to  $\gamma$  was estimated, using a strategy where the  $\mathcal{CP}$  asymmetries for all eight  $B \rightarrow h^+h'^-$  decays with mass peaks near  $m_{B_d}$  under the  $\pi\pi$  mass hypothesis<sup>11</sup> are fitted simultaneously. Possible U-spin breaking effects were taken into account, with the allowed ranges for the U-spin breaking variables being  $\xi \in [0.8, 1.2]$  and  $|\Delta\theta| < 20^\circ$ . The total uncertainty on  $\gamma$  with one nominal year of LHCb data ( $2 \text{ fb}^{-1}$ ) using this strategy was found to be  $\simeq 7^\circ$ , which includes statistical, systematic and theoretical (U-spin breaking) uncertainties. This is a factor of  $\simeq 4$  better than the uncertainty from existing direct measurements (see Table 1.2).

The value of  $\gamma$  found using  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^-K^+$  will be compared with the value measured in the “tree-only” decays, as described in the previous section. A significant discrepancy between these values would signal the presence of New Physics in the loop processes contributing to  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^-K^+$ .

## 1.5 Summary

This chapter has presented an overview of the Standard Model of particle physics, and has described in detail the CKM mechanism, which in the Standard Model is responsible for  $\mathcal{CP}$  violation in the quark sector. The consequences of the CKM mechanism for the behaviour of neutral mesons have been explored. Some ways in which the LHCb experiment can study neutral and charged mesons to improve on current knowledge of the CKM mechanism, and possibly identify inconsistencies therein, have been summarised. The physics of two-body charmless hadronic  $B$  decays ( $B \rightarrow h^+h'^-$  decays) have been outlined, and two physics measurements arising from the study of such decays at LHCb have been described in detail. These descriptions should facilitate understanding of the relevance and impact of the work described in Chapters 4–6 of this thesis.

---

<sup>11</sup>This is all of the  $B \rightarrow h^+h'^-$  decays except  $B_d \rightarrow p\bar{p}$  and  $B_s \rightarrow p\bar{p}$ .

# Chapter 2

## The LHCb Detector

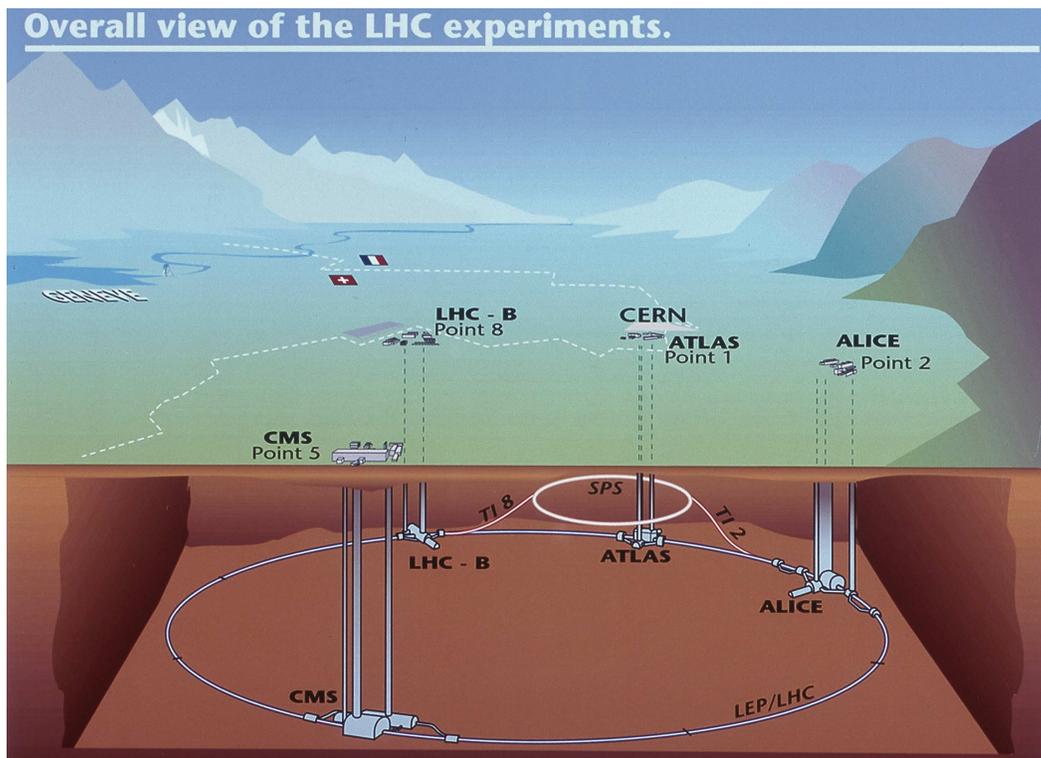
In this chapter a discussion of the design and function of the LHCb detector will be presented. Section 2.1 will briefly describe the LHC accelerator complex and the aims of the four main LHC experiments. Section 2.2 will describe the LHCb detector in detail, by considering each component of the detector's hardware and software systems in turn. Particular attention will be paid to the RICH subdetector, as work carried out on the photon detectors for the RICH will be described in Chapter 3.

### 2.1 The LHC

#### 2.1.1 The LHC Accelerator

The LHC [66], located at CERN [67] near Geneva, is the world's largest and most powerful particle accelerator. It is designed to collide together bunches of protons with a centre-of-mass energy of 14 TeV, and at a frequency of 40 MHz. The LHC accelerator complex consists of several linked accelerators and makes use of previously existing facilities. The protons are obtained from hydrogen gas, and undergo an initial acceleration to an energy of 50 MeV in the Linear Accelerator (LINAC2). They are then fed into the Proton Synchrotron Booster (PSB), where they reach 1.4 GeV, followed by the Proton Synchrotron (PS), where they reach 26 GeV. Next the protons are delivered to the Super Proton Synchrotron (SPS), which accelerates them to 450 GeV. Finally, the protons are then injected into the main LHC ring, where they are accelerated to 7 TeV.

The main LHC accelerator ring is 100 m underground, and is 27 km in circumference. Two independent pipes are used to allow beams of protons to circulate in both directions without interfering with each other. The protons are kept within an approximately circular path within the pipes by 1700 dipole magnets. Each dipole magnet produces a field of 8.33 T.



**Figure 2.1:** Overview of the layout of the LHC. The main accelerator ring is shown in blue and white, with the SPS and the injection system in red and white. Reproduced from [67].

As well as reaching a centre-of-mass energy around 7 times higher than any other particle accelerator, the LHC will also produce much larger event rates than previous hadron accelerators. This rate is known as the luminosity  $\mathcal{L}$ . The LHC maximises its luminosity by colliding dense high-population bunches of protons at a very high rate. Each proton bunch contains around  $10^{11}$  protons, but has (at the LHCb collision point) a transverse radius of only  $71 \mu\text{m}$  and a length of  $7.5 \text{ cm}$ .

The layout of the LHC is shown in Figure 2.1. The location of the four main LHC experiments (see Sec. 2.1.2) are shown, along with the SPS and the injection system.

## 2.1.2 The LHC Experiments

There are four main experiments that will take data from studying LHC collisions. They are:

- A Toroidal LHC Apparatus (ATLAS) [68]. ATLAS is one of two general purpose detectors (GPDs) that will operate at the LHC. The physics goals of ATLAS include finding the mechanism responsible for Electroweak Symmetry Breaking (possibly via observation of the Higgs boson) and searching for evidence of a number of possible

New Physics (NP) scenarios (via direct production of new particles). Such scenarios include supersymmetry and large extra dimensions.

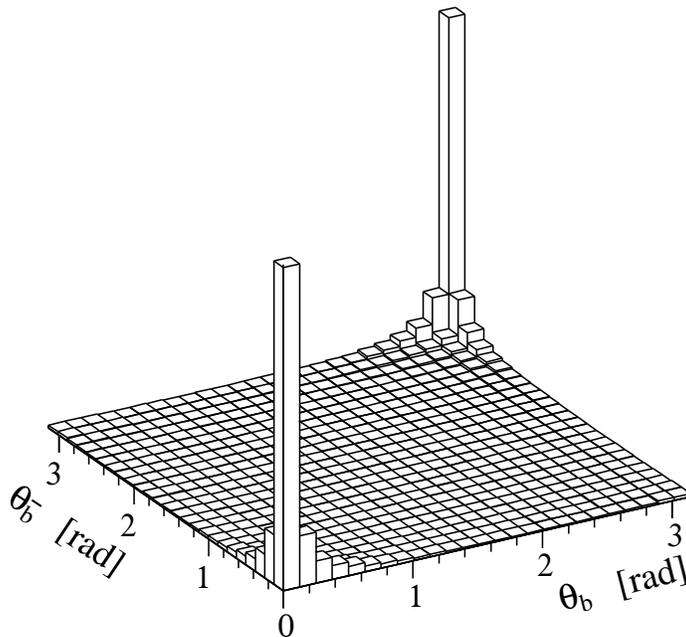
- Compact Muon Solenoid (CMS) [69]. CMS is the other GPD. It has similar physics goals to ATLAS, although the detector layout and choice of detector technologies differ in some aspects. The purpose of having two such GPDs is to provide complementarity in the search for NP, and to allow any discoveries to be independently confirmed.
- Large Hadron Collider Beauty (LHCb) [70]. LHCb is a dedicated heavy flavour physics experiment that aims to carry out stringent tests of the Standard Model and probe New Physics effects by making precision measurements in the B sector. The detector is described in detail in Sec. 2.2, and its physics aims in Sec. 1.3.
- A Large Ion Collider Experiment (ALICE) [71]. The LHC is also capable of acting as a collider of lead ions. ALICE is primarily designed to study such collisions. Its main physics goal is to obtain a better understanding of an exotic state of matter known as the quark-gluon plasma. Improved understanding of this has the potential to shed light on conditions in the early universe.

## 2.2 The LHCb Detector

The LHCb detector [70, 44, 72] is a forward arm spectrometer. It is designed to carry out stringent tests of the Standard Model and probe New Physics effects by making precision measurements in the B sector.

The LHC provides an excellent laboratory for the study of  $B$  hadrons, since the production cross-section of  $b\bar{b}$  pairs in 14 TeV proton-proton collisions is very high,  $\simeq 500 \mu\text{b}$ . This will enable the LHC to quickly overtake previous colliders in terms of the number of  $B$  hadrons produced. The LHCb detector has an angular acceptance starting at 10 mrad, and extending to 300 mrad in the bending plane, and 250 mrad in the non-bending plane. This geometry has been chosen in order to accept the majority of  $B$  hadrons in the forward cone, as they tend to be produced at very low polar angles (see Fig. 2.2). These kinematics arise because the partons (quarks or gluons) inside the two colliding protons are likely to have very different momenta from each other, so that the produced  $b\bar{b}$  quarks subtend a small angle with respect to the beam line.

The LHCb detector consists of a number of subdetectors, each of which plays its own rôle in the reconstruction of interesting events and in the separation of signal and background. The layout of LHCb is shown in Fig. 2.3. From left to right, the following subdetectors can be seen: the Vertex Locator (VELO), the first Ring Imaging Cherenkov counter



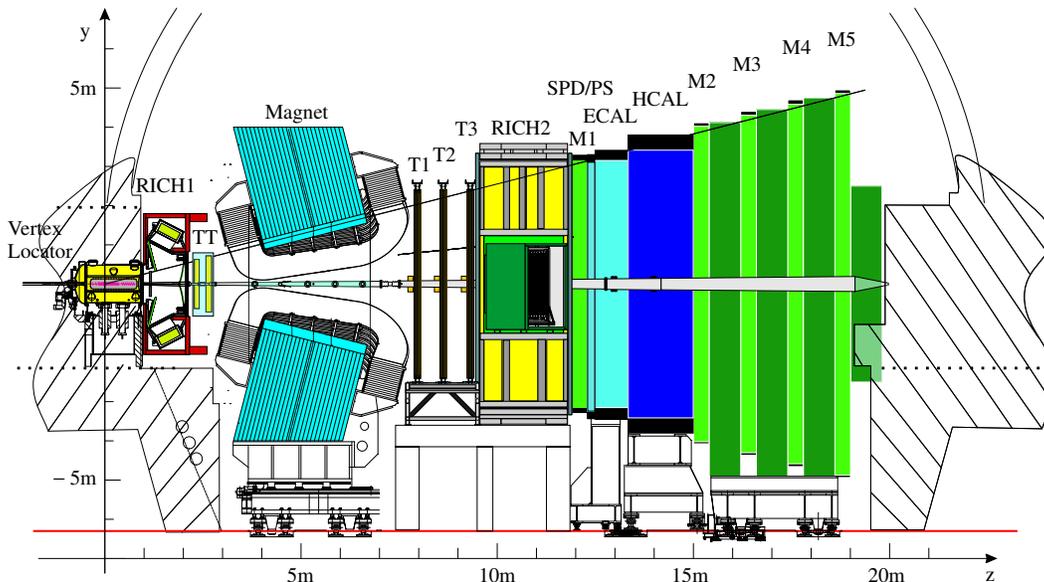
**Figure 2.2:** Angular distribution of  $b$  and  $\bar{b}$  quarks from proton-proton collisions at 14 TeV.  $\theta_b$  ( $\theta_{\bar{b}}$ ) is the production angle of the  $b$  ( $\bar{b}$ ) quark with respect to the proton beam. Reproduced from [44].

(RICH1), the Tracker Turicensis (TT), the magnet, the Tracking Stations (T1,T2,T3), the second Ring Imaging Cherenkov counter (RICH2), the first Muon Station (M1), the Electromagnetic Calorimeter (ECAL), the Hadronic Calorimeter (HCAL), and the remaining Muon Stations (M2,M3,M4,M5). Each of these systems is described in Secs. 2.2.1 to 2.2.6.

LHCb is designed to operate at a nominal luminosity of  $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ , which is around 50 times below the nominal luminosity of ATLAS and CMS. The lower luminosity is achieved by defocusing the beam just before the LHCb collision point. At this lower luminosity the mean number of inelastic proton-proton collisions per bunch crossing is around 0.6. Having a low number of such hard interactions per bunch crossing is desirable for LHCb as it greatly aids the selection of events containing  $B$  hadrons, and also aids the rejection of background events [44, 72]. Furthermore the lower luminosity also reduces the amount of radiation damage suffered by the components of the detector.

### 2.2.1 Vertex Locator

One of the key properties of  $B$  hadrons, which is used to distinguish them from light flavour events, is their large lifetime, of the order of picoseconds. This large lifetime means that  $B$  hadrons can fly a significant distance between their creation point (which at the LHC is the proton-proton collision point, known as the primary vertex or PV) and their decay

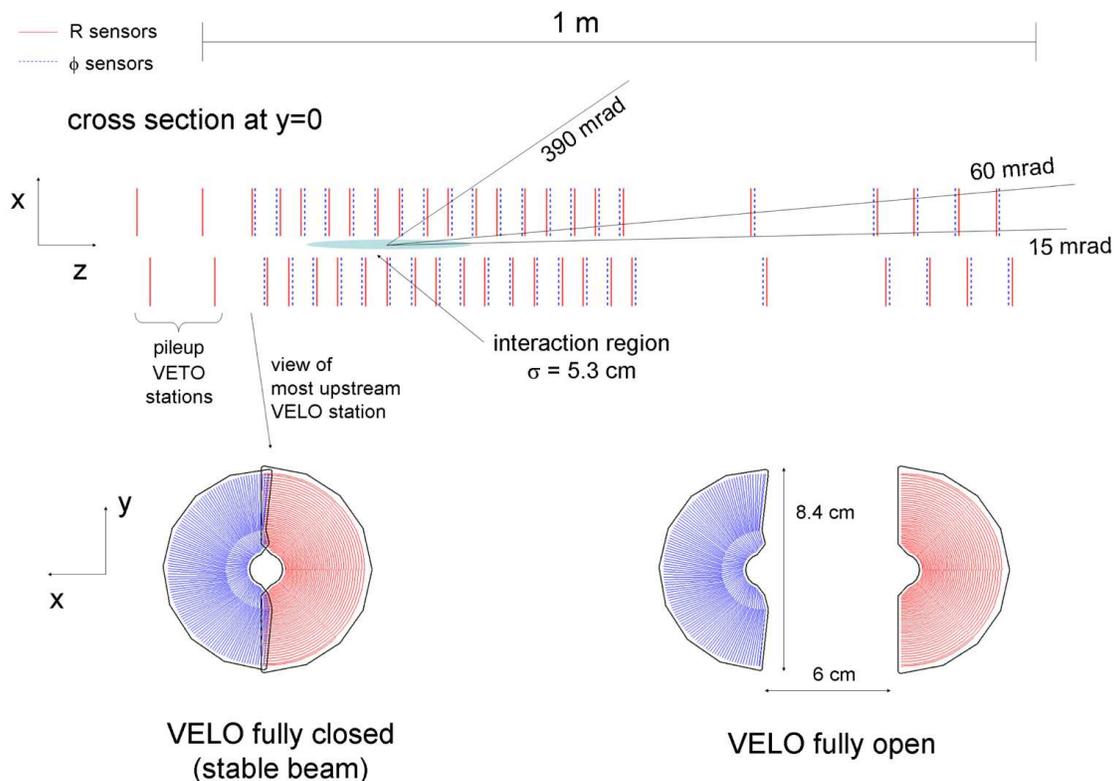


**Figure 2.3:** Layout of the LHCb detector. The beamline is along the  $z$  axis. Reproduced from [72].

point (known as the secondary vertex or SV). To take advantage of this, LHCb uses the Vertex Locator (VELO) [70, 73] to make very precise measurements of the positions of the locations of both the primary vertex and of the secondary vertices of any  $B$  hadrons produced. The VELO can measure the position of primary vertices with a resolution of  $10 \mu\text{m}$  in the transverse ( $xy$ ) plane, and about  $42 \mu\text{m}$  in the beam ( $z$ ) direction. The resolution on the decay length of a  $B$  hadron depends on the decay mode considered, ranging from  $220 \mu\text{m}$  to  $330 \mu\text{m}$ . Along with decay length, another variable that plays an important rôle in separating  $B$  hadron decays from background events is the impact parameter of a track, which is the distance of closest approach of the track to the PV. For tracks with transverse momentum ( $p_T$ ) higher than  $10 \text{ GeV}$ , the VELO resolution on the impact parameter is around  $20 \mu\text{m}$ . In order to achieve the above resolutions, the VELO requires a signal to noise ratio (S/N) greater than 14 and a single channel efficiency greater than 99%. In addition, excellent single hit resolution is required (the best resolution achieved is  $\simeq 4 \mu\text{m}$ , for tracks in the fine pitch region (see Sec. 2.2.1.2) and with an optimal opening angle of  $\simeq 100 \text{ mrad}$ ).

### 2.2.1.1 VELO Layout

The VELO consists of 21 stations, arranged along the  $z$  axis as shown in Fig. 2.4. Note the presence at negative  $z$  (left hand side of Fig. 2.4) of two stations which are known as the pile-up system. Information relating to the number of charged tracks seen in the pile-up system is used by the Level 0 trigger (see Sec. 2.2.7.1) to reject events with many  $p-p$  interactions. Tracks found in the pile-up stations are also used in the reconstruction of primary vertices.

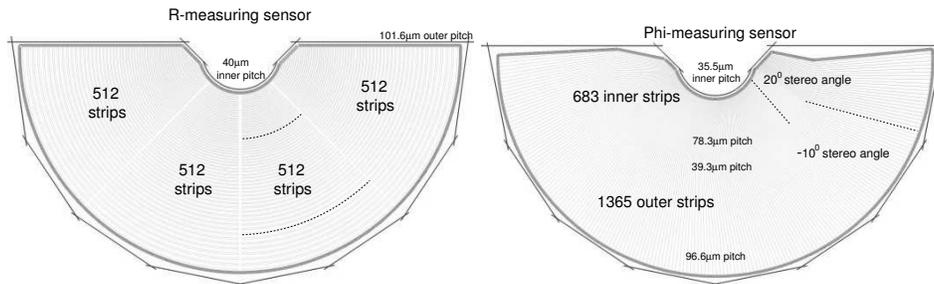


**Figure 2.4:** Geometry of the VELO stations along the LHCb  $z$  axis. The closed and open positions of the modules are also shown. Refer to the text for details. Reproduced from [70].

However information from the pile-up system is not used to reconstruct  $B$  hadrons, as LHCb can only reconstruct  $B$  hadrons that are produced in the forward direction.

### 2.2.1.2 VELO Sensors

Each station consists of two modules, one either side of the beam pipe, with each module covering  $182^\circ$  of the azimuthal angle  $\Phi$ . In turn each module comprises two silicon sensors, an  $R$  sensor and a  $\Phi$  sensor, placed back-to-back. These sensors consist of a series of strips that record hits from particles passing through them. The  $R$  and  $\Phi$  sensors are both n-on-n type silicon and both have a thickness of  $300 \mu\text{m}$ , but the geometries of their strips are very different. The sensor layouts are shown in Fig. 2.5. The strips on an  $R$  sensor are annular, and give information on how far a hit is from the beam axis. To reduce the occupancy, the  $R$  strips are divided into 4 sectors of  $45^\circ$  each. Each sector contains 512 strips. The pitch size of  $R$  strips varies between  $38 \mu\text{m}$  and  $102 \mu\text{m}$ , with the average strip occupancy being around 0.4%. The strips on a  $\Phi$  sensor are radial, and give information on the azimuthal angle of



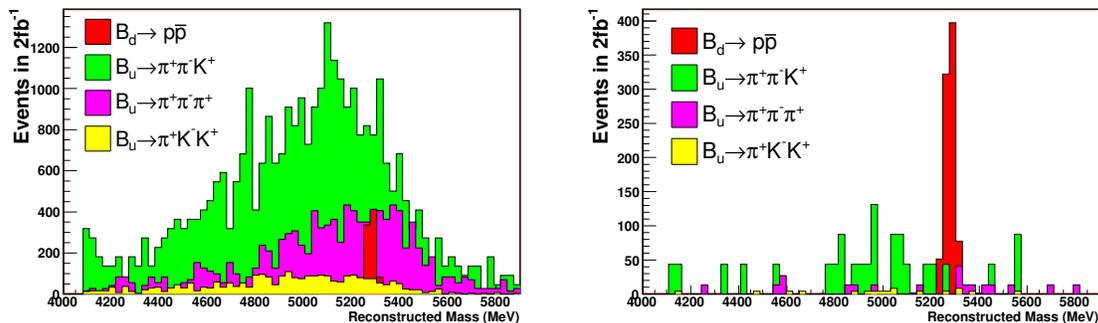
**Figure 2.5:** Geometry of the VELO sensors ( $R$  (left) and  $\Phi$  (right)). See text for details. Reproduced from [72].

the hit with respect to the beam axis. The  $\Phi$  strips are divided into inner and outer strips, in order to reduce the occupancy. There are 683 inner strips and 1365 outer strips. The pitch size of  $\Phi$  strips varies between  $36 \mu\text{m}$  and  $97 \mu\text{m}$ , with the average strip occupancy being around 0.4%. One unusual feature of the  $\Phi$  strips is that they have a non-zero stereo angle with respect to the perfectly radial direction. The inner strips have an initial stereo angle of  $20^\circ$ , while the outer strips have an initial stereo angle of  $-10^\circ$ . The  $\Phi$  sensors in adjacent modules along the beam axis are then flipped with respect to each other, so that they have opposite stereo angles. This improves the pattern recognition. Combining information from the  $R$  and  $\Phi$  sensors allows a 3D point for the hit to be found. Hits along the VELO stations are then used to reconstruct tracks, which correspond to charged particles passing through the VELO.

The nominal position of the sensors places them closer to the beam (the innermost end of an inner  $\Phi$  strip is only 7 mm from the beam axis) than will be safe during LHC injection, when the beam aperture is considerably larger than during stable running. Hence the sensors must be retractable. The change in sensor position this entails is shown in Fig. 2.4. Note that the fact that the sensors must be retracted before every injection and then closed again once the beam is stable places stringent requirements on the VELO alignment system, which do not apply to the alignment systems of the other LHCb subdetectors.

## 2.2.2 Ring Imaging Cherenkov Counters

Particle Identification (PID) is a fundamental requirement for the LHCb physics programme. Particle identification for hadrons (protons, charged pions and charged kaons) in the momentum range 1–100 GeV is provided by two Ring Imaging Cherenkov (RICH) counters [70, 74]. One class of channels whose study requires PID is the  $B \rightarrow h^+ h'^-$  class of channels, where  $h$  stands for a charged pion, kaon or proton. One such channel, which will be studied in Chapter 5, is the rare decay  $B_d \rightarrow p\bar{p}$ . The importance of the RICH to the study of  $B_d \rightarrow p\bar{p}$



**Figure 2.6:** Invariant mass distributions for  $B_d \rightarrow p\bar{p}$  and selected three-body backgrounds, before (left) and after (right) RICH PID information is used.

is illustrated in Fig. 2.6, where the signal is clearly visible above the partially-reconstructed three-body backgrounds only after the RICH PID information is used. In addition to its ability to distinguish decays with identical topological signals, the RICH is also crucial in identifying the flavour of neutral  $B$ -hadrons at production. This process is known as flavour tagging and is necessary in many  $\mathcal{CP}$  violation studies [75]. One important flavour tagging method at LHCb is that of same-side kaon tagging. The aim is to identify the charged kaon that can be produced in association with a  $B_s$  meson. The initial flavour of the  $B_s$  can be inferred from the charge of the kaon. For this method to work it is necessary to reduce background from charged pions, relying on the RICH PID. This section discusses in detail how the RICH detectors of LHCb provide these important inputs to LHCb physics analyses.

### 2.2.2.1 Cherenkov Radiation

The RICH detectors of LHCb rely upon the phenomenon of Cherenkov Radiation. This phenomenon will be briefly described here; a more detailed discussion is available in [76]. The Cherenkov effect causes a charged particle travelling through a medium (a *radiator*) at a velocity that is greater than the velocity of light in that medium, to emit photons. These Cherenkov photons are emitted in a cone, and the opening angle  $\theta_c$  of the cone is related to the velocity  $v$  of the particle:

$$\cos \theta_c = \frac{c}{nv}, \quad (2.1)$$

where  $c$  is the speed of light in vacuum, and  $n$  is the refractive index of the radiator. Hence a measurement of  $\theta_c$  can be used to measure the particle's velocity. When combined with momentum information this allows a measurement of the particle's mass (the details of how the particle identification is actually performed in LHCb are given in Sec. 2.2.2.6).

The number of photons radiated by a particle with charge  $z$  (in units of the elementary charge) traversing a radiator of length  $L$  is:

$$N_\gamma = L \frac{\alpha z^2}{\hbar c} \int \sin^2 \theta_c(E) dE, \quad (2.2)$$

where  $\alpha \simeq 1/137$  is the fine structure constant and  $E$  is the energy of the Cherenkov photon.

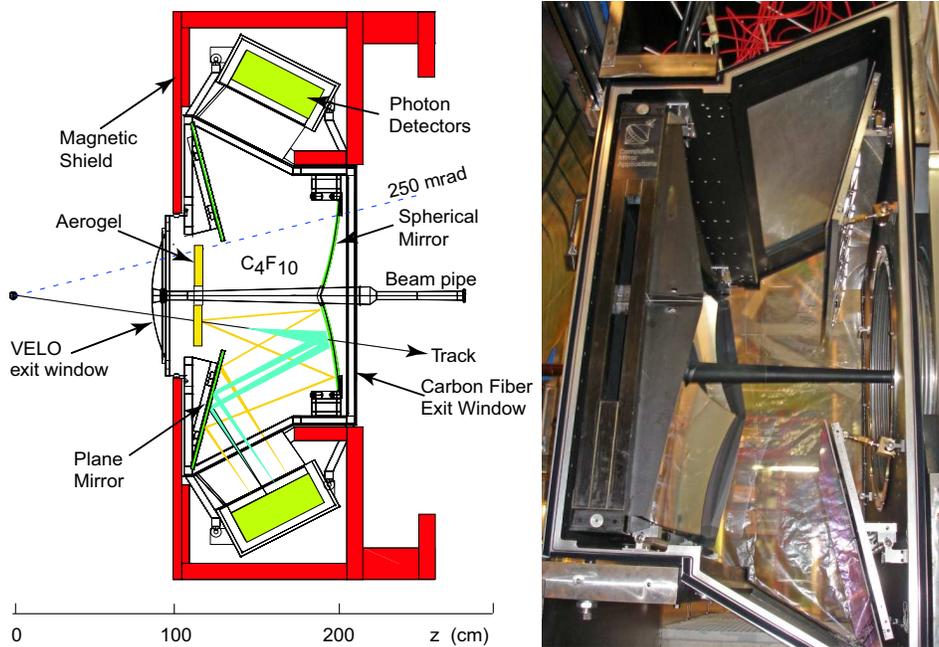
### 2.2.2.2 The RICHes of LHCb

In LHCb there are two separate RICH systems used: RICH1 [77] and RICH2 [78]. Figures 2.7 and 2.8 show schematic diagrams and photographs of RICH1 and RICH2 respectively. Two separate systems are necessary in order to maximise the number of tracks on which particle identification can be reliably performed, as tracks from  $B$  decays vary widely in their polar angle and momentum. A scatterplot illustrating how tracks from  $B_d \rightarrow \pi^+\pi^-$  events are distributed in these variables is shown in Fig. 2.9. The RICHes cover different ranges of polar angle and momentum, as follows: RICH1 is placed upstream of the LHCb magnet and uses two radiators; aerogel ( $n \simeq 1.03$ ) and  $C_4F_{10}$  ( $n \simeq 1.0014$ ). It provides momentum coverage ranging from around 1 GeV to around 70 GeV, and angular coverage starting at 25 mrad and extending to 300 mrad in the horizontal plane and 250 mrad in the vertical plane. RICH2 is placed downstream of the magnet, and uses a single radiator of  $CF_4$  ( $n \simeq 1.0005$ ). It provides momentum coverage ranging from around 15 GeV to beyond 100 GeV, and angular coverage starting at 15 mrad and extending to 120 mrad in the horizontal plane and 100 mrad in the vertical plane. The angular and momentum coverage of each RICH is superimposed onto Fig. 2.9.

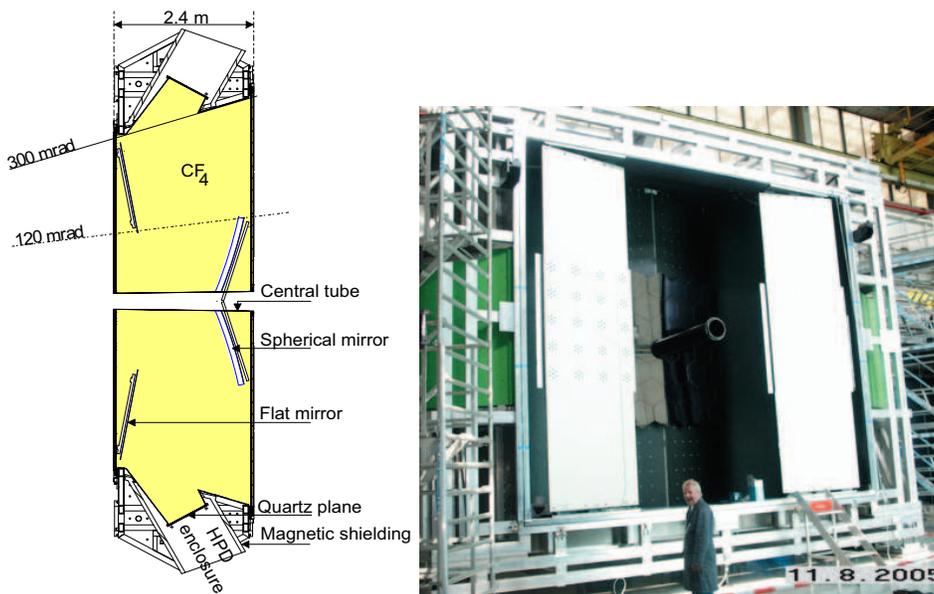
The dependence of  $\theta_c$  on momentum for different particle types in each of the three radiators is shown in Fig. 2.10. At high momentum, where  $\theta_c$  is almost independent of the particle type, the track is said to be *saturated*.

### 2.2.2.3 Detection of Cherenkov Photons

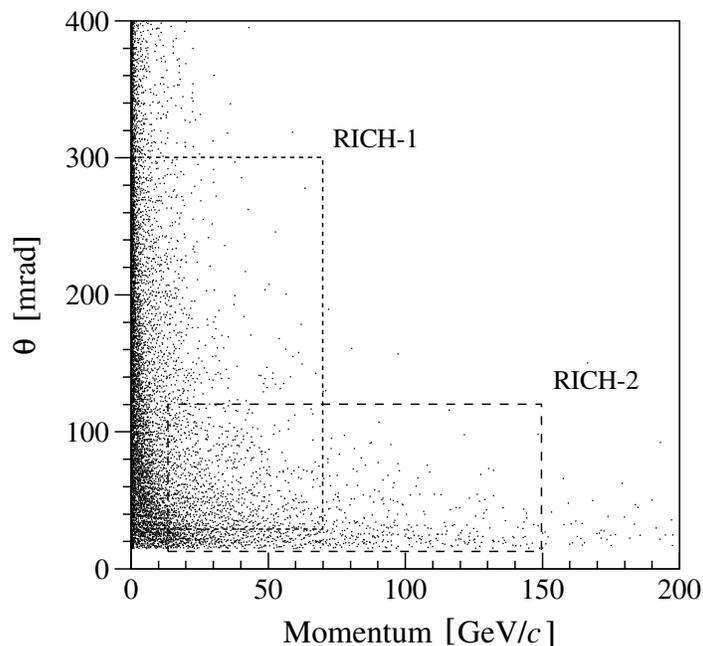
In both RICHes the Cherenkov photons are reflected off a series of mirrors onto a photodetector plane, where they give hits that form rings. These hits are used to identify the particle type, as described in Sec. 2.2.2.6. The photodetector planes are populated by arrays of Hybrid Photon Detectors (HPDs) [79], which detect the Cherenkov photons. This is done by making use of the photoelectric effect: the Cherenkov photons strike a quartz window on which a photocathode material is deposited, and produce photoelectrons. These photoelectrons are then accelerated toward a pixellated silicon sensor, where they create many electron-hole pairs. It is this signal from the photoelectron that is actually detected, and



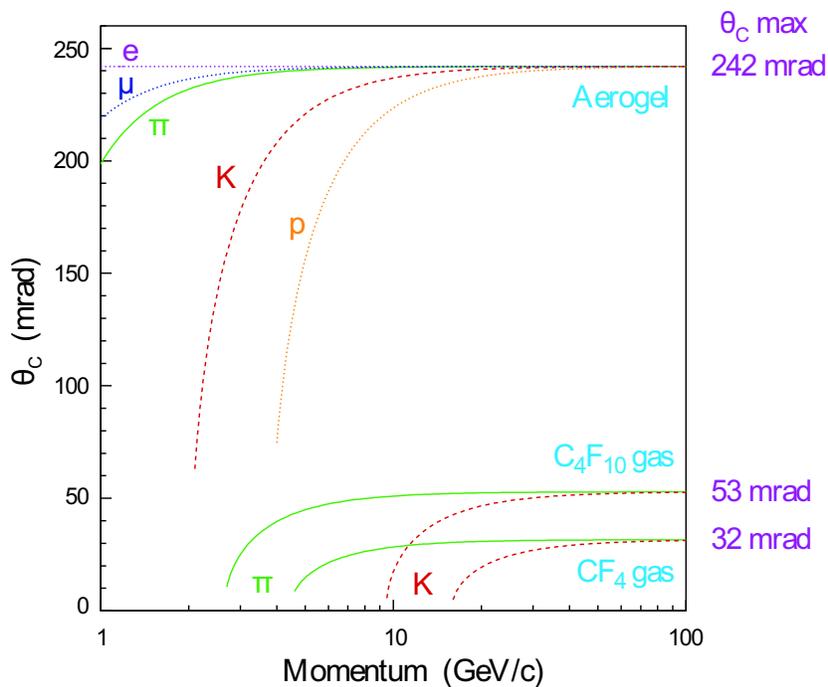
**Figure 2.7:** Schematic diagram (left) and photograph (right) of RICH1. Note that the LHCb interaction point is to the left in the diagram, but to the right in the photograph. Both figures reproduced from [70].



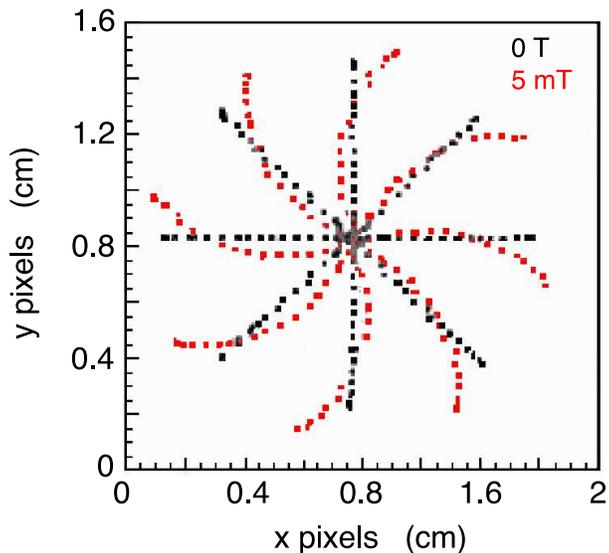
**Figure 2.8:** Top-view schematic diagram (left) and photograph (right) of RICH2. Note that the RICH2 entrance window is not present in the photograph. Both figures reproduced from [70].



**Figure 2.9:** Distribution in polar angle  $\theta$  and momentum  $p$  of tracks from  $B_d \rightarrow \pi^+ \pi^-$  events. The angular and momentum coverage of each RICH is superimposed. Reproduced from [74].



**Figure 2.10:** Dependence of the Cherenkov angle  $\theta_c$  on momentum for different particle types in the three radiators of the LHCb RICHes. The maximum possible value of  $\theta_c$  for each radiator is also given. Reproduced from [70].



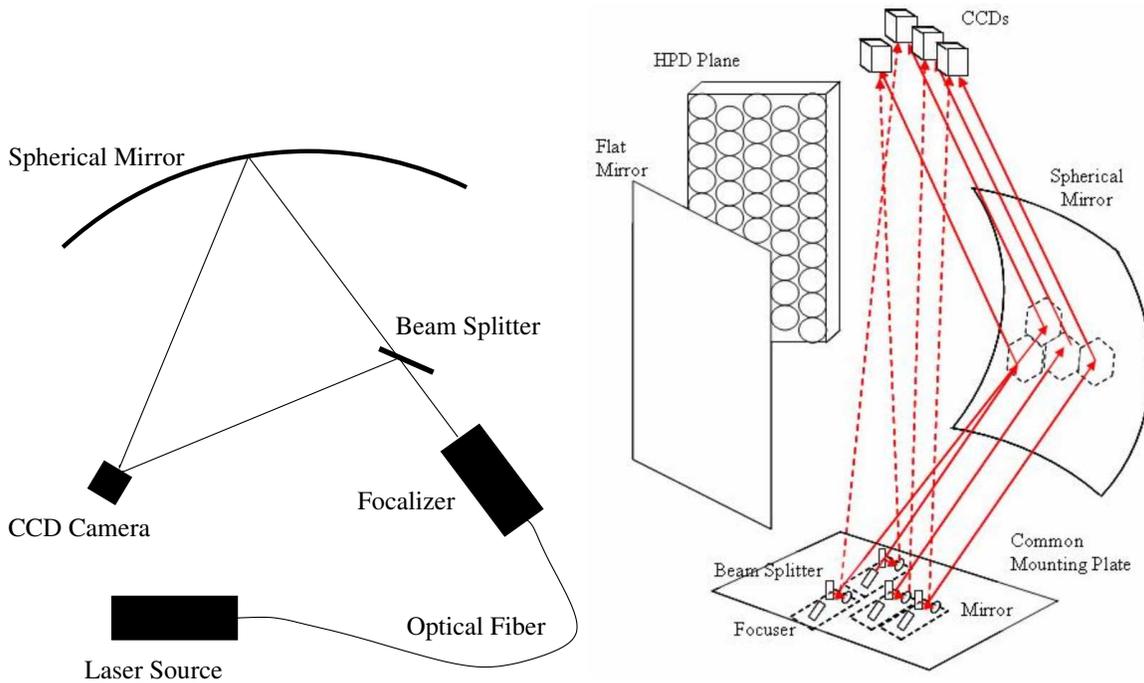
**Figure 2.11:** Effect of longitudinal magnetic field on HPD hit pattern. Reproduced from [70].

the location of the photon on the quartz window is then inferred. Note that the HPDs are described in detail in Chapter 3, and hence are not discussed further in this chapter.

The path of the photoelectrons within the HPDs can be affected by the presence of magnetic fields. To protect the HPDs from the fringe field of the LHCb magnet, iron shielding is present around the superstructure of both RICHes. Also each HPD is surrounded by a cylindrical MuMetal shield, to combat the residual field existing inside the RICH [80]. However despite these measures, the remaining field inside the HPDs can still be large enough to have a significant effect on the photoelectron path, with the maximum expected residual field in RICH1 being around 2.4 mT. The effect of a 5 mT magnetic field, applied parallel to the HPD axis, on a hit pattern on the HPD pixel chip is shown in Fig. 2.11. To measure the effect of the residual magnetic field and account for it, both RICHes use magnetic distortion monitoring systems, which shine known light patterns onto the HPD plane and then monitor how the pattern on each HPD changes as the magnetic field is changed.

#### 2.2.2.4 RICH Mirrors

Within each RICH there are two sets of mirrors [70]; one set of spherical mirrors and a set of flat mirrors [77, 78]. The spherical mirrors of RICH1 lie within the LHCb acceptance, and are composed of a carbon fibre reinforced polymer in order to help reduce the amount of material seen by particles before they reach the calorimeter. Indeed the spherical mirrors represent only 1.5% of a radiation length ( $X_0$ ). The flat mirrors of RICH1 and both sets of RICH2 mirrors lie outside of the LHCb acceptance, and so can be constructed from glass as



**Figure 2.12:** Schematic diagrams of the Laser Alignment Monitoring Systems (LAMS) for RICH1 (left) and RICH2 (right). Reproduced from [70] and [81] respectively.

the material budget is not an issue. Each set of mirrors is made up of segments: in RICH1 the spherical mirrors consist of two spherical segments in each plane (above and below the beam pipe), while the flat mirrors consist of eight rectangular mirrors in each plane. In RICH2 the spherical mirrors consist of 26 hexagonal mirrors in each plane (left and right of the acceptance), while the flat mirrors consist of 20 rectangular mirrors in each plane. The reflectivity of all RICH mirrors has been studied, and was found to have a typical value of around 90% for incoming photons in the relevant wavelength region.

As any deviation of the mirrors from their ideal position will affect where Cherenkov photons striking that mirror arrive on the HPD plane, the alignment of the mirrors must be closely monitored, and the current state of the mirrors communicated to the reconstruction software. Within both RICH1 and RICH2 the mirror alignment is monitored using a Laser Alignment Monitoring System, or LAMS [81]. Schematics of the LAMS setups for RICH1 and RICH2 are shown in Fig. 2.12. The LAMS takes light from a laser source and passes it through a set of optical fibres. Each fibre ends inside the RICH, where the light is passed through a focalising unit. This unit focuses the beam and then splits it into two, with one beam aimed directly at a CCD camera mounted inside the RICH, and the other beam reflecting off a mirror segment, then arriving at the CCD camera. This creates two light spots on the CCD camera. Any change in the relative position of the two spots reveals a change in the orientation of the mirror segment.

### 2.2.2.5 RICH Electronics

The task of the RICH electronics [82] is to read out the data from the HPDs. The RICH electronics are divided into the Level 0 (L0) electronics and the Level 1 (L1) electronics. The L0 electronics are located on the detector and hence have to be radiation-hard. As well as the HPD pixel chip itself<sup>1</sup>, the L0 electronics comprise the L0 board, which interfaces the HPD with the LHCb online system (see Sec. 2.2.8), and the distribution system for the high and low voltages for the HPDs. The L1 electronics are located in the counting house, behind the protective concrete shield, and so need not be radiation-hard. They compress the data received from the L0 electronics (through zero-suppression for example), and also interface the custom format of this data with the commercial ethernet protocol used by the LHCb DAQ network (see Sec. 2.2.8). The data manipulation is carried out by a series of Field-Programmable Gate Arrays (FPGAs), which are controlled by the LHCb Readout Supervisor.

### 2.2.2.6 Particle Identification Algorithm and Performance

Using information from the LHCb tracking system, the RICH reconstruction software estimates the location on the HPD plane of the centre of the Cherenkov ring created by the photons emitted from each reconstructed track [70]. In doing this there are several sources of uncertainty that need to be taken into account. The most important sources are emission point error, chromatic dispersion, pixel resolution and track reconstruction.

- Emission point error: in the reconstruction it is assumed that all Cherenkov photons were emitted when the particle was at the midpoint of the radiator. In reality the photons are emitted randomly along the trajectory of the particle as it traverses the radiator, hence some uncertainty is introduced.
- Chromatic dispersion: the refractive index of each radiator is dependent on the wavelength of the emitted photon. As emitted photons will have different wavelengths following the Cherenkov emission spectrum [76], they experience a different refractive index, resulting in a spread in the Cherenkov angle (see Eqn. 2.1).
- Pixel resolution: The pixels on the HPD silicon sensor, which detect the photoelectrons produced by the Cherenkov photons, have a finite size, which introduces uncertainty to the location on the quartz window that was struck by the Cherenkov photon.

---

<sup>1</sup>The HPD pixel chip will be described in detail in Chapter 3.

- Track reconstruction : The finite resolution of the tracking system introduces uncertainty on the trajectory of the charged particle, and thus on the location of the centre of the Cherenkov ring.

The expected relative size of each of the above effects for each radiator is given in Table 2.1.

Detector	RICH1		RICH2
Radiator	Aerogel	C <sub>4</sub> F <sub>10</sub>	CF <sub>4</sub>
$\sigma_{\theta}^{\text{emis}}$ (mrad)	0.60	0.74	0.31
$\sigma_{\theta}^{\text{chrom}}$ (mrad)	1.61	0.81	0.42
$\sigma_{\theta}^{\text{pixel}}$ (mrad)	0.78	0.83	0.18
$\sigma_{\theta}^{\text{track}}$ (mrad)	0.26	0.42	0.20
$\sigma_{\theta}^{\text{total}}$ (mrad)	2.00	1.45	0.58
$N_{\text{pe}}$	6.5	30	22

**Table 2.1:** Contributions to the overall resolution on the Cherenkov angle in each of the three RICH radiators [74], and the expected number of detected photoelectrons in each radiator [70].

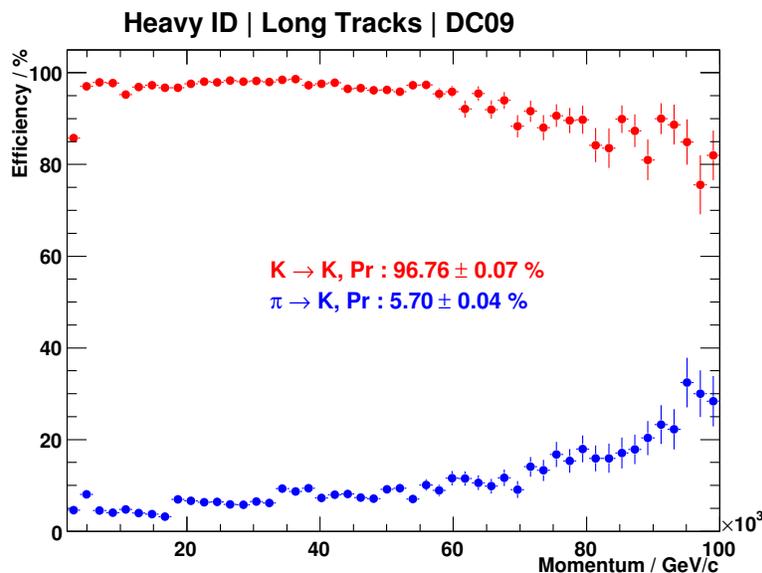
For each track the software then cycles through five different particle hypotheses ( $e$ ,  $\mu$ ,  $\pi$ ,  $K$  and  $p$ ). Each hypothesis is associated with an expected pattern of hits for that track. By comparing this pattern with the observed pattern of hits for each track, a likelihood can be constructed. Finally, the set of hypotheses is chosen that maximises the *total* likelihood when all tracks are considered. Each hypothesis also has a confidence level associated with it, calculated by considering the change in the total likelihood when that hypothesis is changed. The procedure described above is known as *global pattern-recognition*, because hits from all reconstructed tracks and all three radiators are considered simultaneously.

An estimation of the performance of the RICH hadron PID obtained using the above algorithm and the current known performance of the RICH hardware has been made. Figure 2.13 shows the expected efficiency and misidentification rate for kaons.

During data taking, the actual PID performance of the RICHes will be monitored by studying specially-obtained samples of certain decays [56]. These decays are chosen such that the type of particle associated to a track can be determined purely from the decay kinematics, without the use of PID information.

To measure the  $K - \pi$  separation power of the RICH, the decay  $D^{*+} \rightarrow D^0(K\pi)\pi^+$  (and its charge conjugate) will be used. Here the tightly constrained  $D^* - D^0$  mass difference is very effective in reducing the background rate while maintaining high signal efficiency.

To measure the  $p - \pi$  separation power of the RICH, the decay  $\Lambda \rightarrow p\pi^-$  (and its charge conjugate) will be used. The large lifetime of the  $\Lambda$  will be useful in distinguishing



**Figure 2.13:** Efficiency and mis-identification rates, as a function of momentum, for identifying a kaon as “heavy”, i.e. as a kaon or a proton.

it from most other particles. The  $K_s$  has a similar lifetime to the  $\Lambda$ , but they can be distinguished from each other using their distinctive distributions in the  $(p_T, \alpha)$  plane, where  $\alpha$  is the momentum asymmetry of the two tracks. A plot of such distributions is known as an Armenteros-Podolansky plot [83].

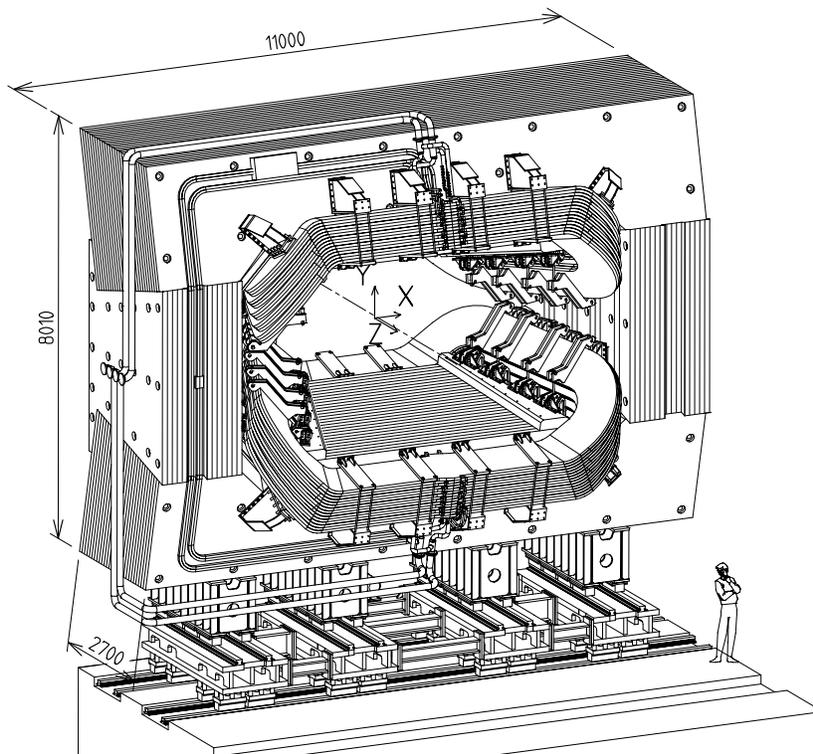
In both of the above channels, tight selection algorithms have been developed that are expected to give sample of the control channels with greater than 90% purity. Both channels will have dedicated streams in the LHCb Trigger framework (see Sec. 2.2.7).

### 2.2.3 Magnet

The LHCb detector possesses a powerful dipole magnet [70, 84] that is used to allow measurements of the momenta of charged particles to be made. A perspective view of the magnet is shown in Fig. 2.14. The magnet is a warm (non-superconducting) dipole. Its yoke is constructed from plates of laminated low carbon steel, with a total weight of 1500 tons, while its two coils are made from aluminium and have a total weight of 54 tons.

The magnet at full power produces an integrated field of 4 Tm. This large value is necessary to achieve the required precision on momentum measurements. The magnet polarity will be reversed on a regular basis to monitor differences in charged particle tracking, thus reducing systematic effects in  $\mathcal{CP}$  asymmetry measurements. The magnet is designed so as to maximise the field experienced by particles passing through the LHCb tracking system.

Accurate momentum measurements require an excellent knowledge of the magnitude



**Figure 2.14:** Perspective view of the LHCb dipole magnet, where the LHCb interaction point lies behind it. Distances shown are in mm. Reproduced from [70].

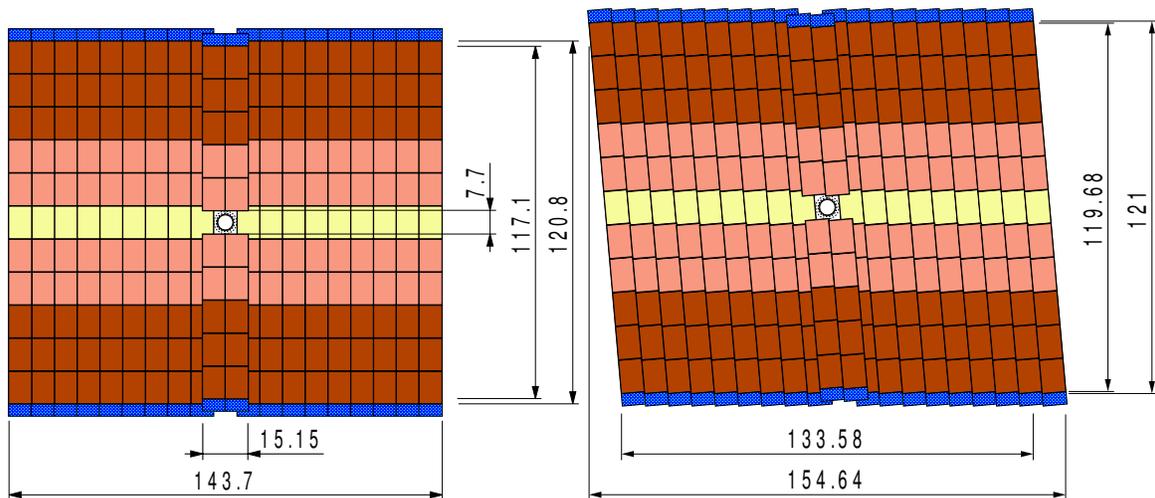
and direction of the field throughout the tracking volume. To acquire this knowledge, the field map of the magnet has been accurately measured by a survey campaign that used Hall probes to measure field at different space points. The survey encompassed both the tracking volume and the volume of the RICH detectors [85].

## 2.2.4 Tracking System

In order to measure the trajectories and momenta of charged particles, LHCb uses several tracking systems, located at different points along the beam line. Between RICH1 and the magnet there are two silicon tracker stations, collectively known as the Tracker Turicensis (TT). Between the magnet and RICH2 there are three tracker stations, known as T1, T2 and T3 respectively. The different components of the tracking systems will now be briefly described.

### 2.2.4.1 Tracker Turicensis

The TT system [70, 72], along with the VELO, allows the trajectories of charged particles between the interaction point and the magnet to be measured. In particular these tracks are used to identify displaced vertices for the purposes of the High Level Trigger (see Sec. 2.2.7.2).



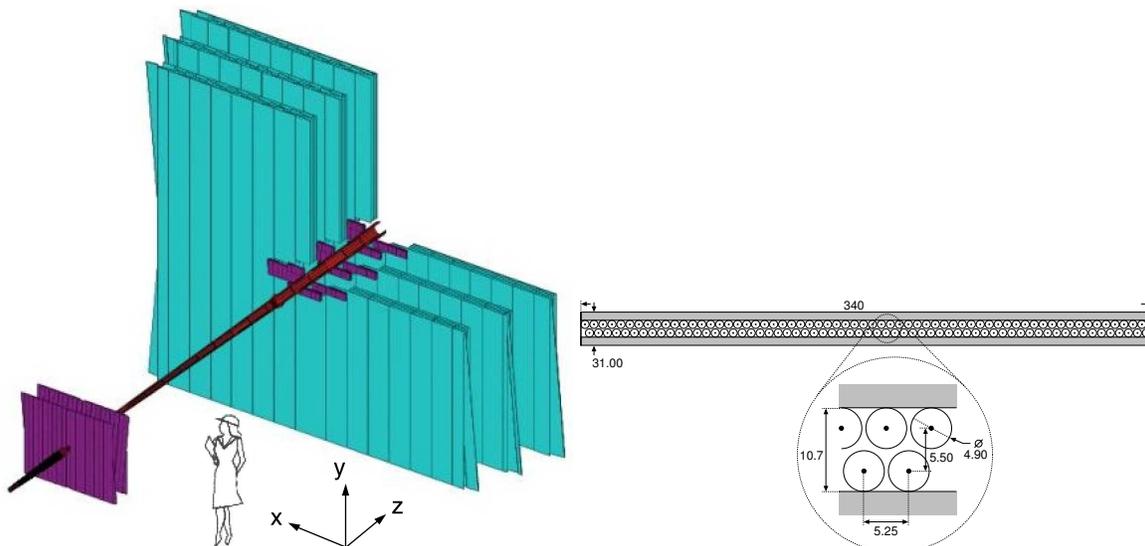
**Figure 2.15:** Layout of Tracker Turicensis layers, with a vertical  $x$  layer on the left and the rotated  $u$  layer on the right. Dimensions are given in cm. Both figures reproduced from [72].

They also provide associated tracks for Cherenkov rings in RICH1. The TT is comprised of four layers of silicon microstrip sensors, with pitches of around  $200 \mu\text{m}$ . The spatial resolution of the microstrips is around  $50 \mu\text{m}$ . The layout of the layers is shown in Fig. 2.15. The total active area of the TT is around  $8.4 \text{ m}^2$ , with a total of 143 360 readout strips. The occupancy per channel in the innermost region of the TT is at the level of a few %. The TT layers are grouped into two pairs, referred to as  $(x, u)$  and  $(v, x)$ . The strips in the  $x$  layers are aligned vertically, while those in the  $u$  ( $v$ ) layer are rotated at an angle of  $-5^\circ$  ( $+5^\circ$ ) to the vertical. This stereo angle geometry allows the transverse momentum ( $p_T$ ) of charged particles to be measured, and also reduces ambiguities between measured hits.

#### 2.2.4.2 Tracking Stations

Particle trajectories measured in the tracking stations allow tracks created in the VELO to be matched with objects identified in the calorimeters and the muon system. Tracks from the tracking stations are also needed to provide associated tracks for Cherenkov rings in RICH2. The outline of the tracking stations layout is shown on the left of Fig. 2.16. Each of the three tracking stations (T1,T2,T3) has an inner region composed of silicon (known as the Inner Tracker, or IT [70, 86]), and an outer region composed of straw tubes (known as the Outer Tracker or OT [70, 87]). The more expensive silicon technology is used near the beam pipe as the charged particle density there is much higher than in the outer regions of the detector, where the cheaper straw tube technology is used.

The IT part of each tracking station consists of four layers of silicon microstrip detectors,



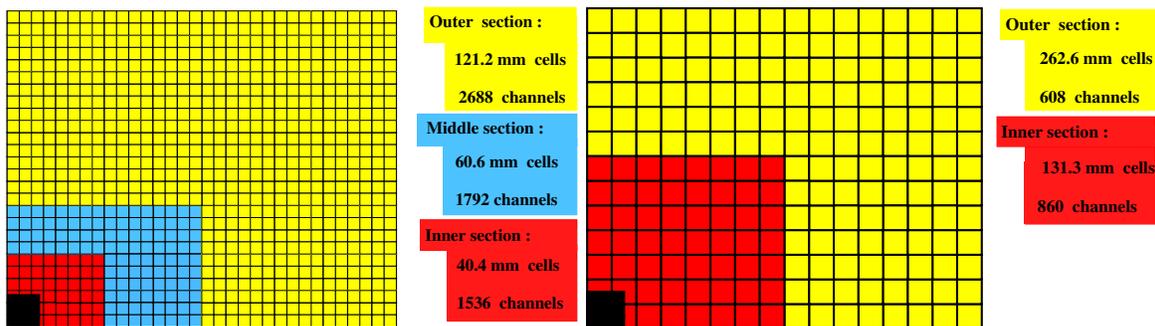
**Figure 2.16:** Left: Outline of the tracking stations layout. Along increasing  $z$  is the TT, followed by T1, T2 and T3. The TT and IT are shown in purple, and the OT is shown in aquamarine. Right: Cross-section of an OT module, where dimensions are given in mm. Both figures reproduced from [70].

which follow the same  $(x, u, v, x)$  geometry as the TT layers described above. As for the TT, the spatial resolution of the IT strips is around  $50 \mu\text{m}$ . The total active area of the IT is around  $4.0 \text{ m}^2$ , with a total of 129 024 readout strips. The occupancy per channel is  $\simeq 1\%$ .

The OT is an array of modules, each of which contains two staggered layers of straw tubes, with each layer having 64 tubes. The cross-section of an OT module is shown on the right of Fig. 2.16. The gas in the drift tubes is a mixture of Ar (70%) and  $\text{CO}_2$  (30%). This mixture is chosen to allow a fast drift time (below 50 ns) and good spatial resolution ( $\simeq 200 \mu\text{m}$ ). As with the IT, each tracking station contains 4 layers of OT modules, arranged in the  $(x, u, v, x)$  geometry. In total there are 55 000 single straw tube readout channels. The occupancy per channel in the most active region of the OT (where it meets the IT) is less than 10%.

## 2.2.5 Calorimeters

LHCb uses its calorimeter system [70, 88] to measure the hit position and energies of electrons, hadrons and photons. This information is used in the Level 0 hardware trigger (see Sec. 2.2.7 for a full description of the LHCb trigger), and also for particle identification. The calorimeter system consists of the Scintillator Pad Detector (SPD), Pre-Shower (PS), Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL). Each of these systems will now be described in turn.



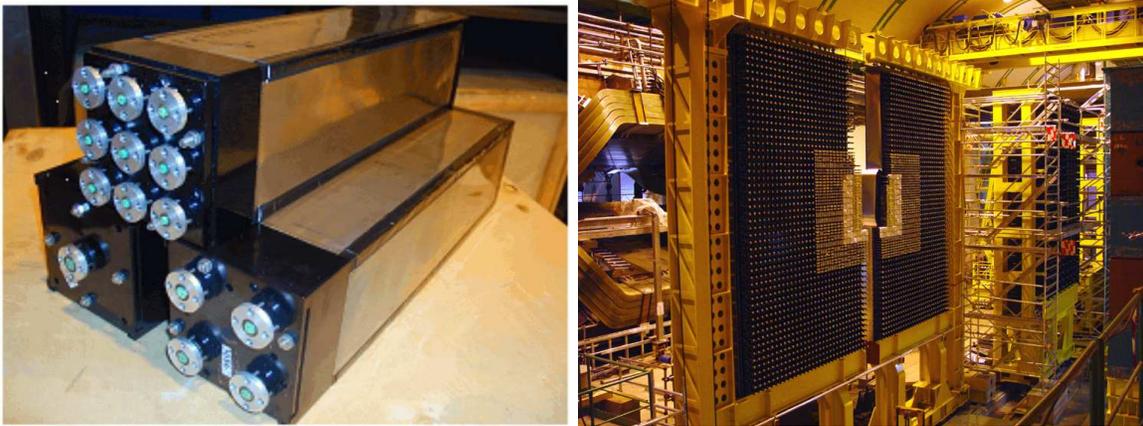
**Figure 2.17:** Segmentation scheme of the SPD/PS and the ECAL (left), and of the HCAL (right). The cell dimensions are also given. Reproduced from [70].

### 2.2.5.1 Scintillator Pad Detector and PreShower

The SPD/PS system consists of two planes of scintillator pads, with a 15 mm-thick lead converter plate sandwiched between them. The segmentation scheme of the SPD/PS pads (which is also the segmentation scheme for the ECAL) is shown on the left of Fig. 2.17. The SPD detects charged particles in order to reduce  $\pi^0$  contamination in the Level 0 electron trigger. The lead plate induces electromagnetic showers, causing signals in the PS that allow contamination from charged pions to be suppressed. The scintillation light produced by the SPD and PS is fed through Wavelength-Shifting (WLS) optical fibres to Multi-anode PhotoMultiplier Tubes (MaPMTs) that sit outside the detector. Each fibre is pointed at a particular pixel of a MaPMT. The MaPMTs amplify the scintillation light so that the signal becomes strong enough to be read out. All parts of the calorimeter system adopt variable lateral segmentation, because the occupancy varies hugely (by two orders of magnitude) between the inner and outer regions. The units of segmentation are known as cells, with each cell being read out by one WLS fibre.

### 2.2.5.2 Electromagnetic Calorimeter

The purpose of the ECAL is to select high-energy photons and electrons for the Level 0 trigger. It also provides information for particle identification of photons and electrons. It consists of alternating layers of lead absorber and scintillating material, divided into modules. Each module consists of 66 layers of 2 mm-thick lead, and 66 layers of 4 mm-thick scintillator tiles. The total radiation length provided by the ECAL is  $25X_0$ , where  $X_0$  is the characteristic radiation length. The readout system for the ECAL is similar to that for the SPD/PS, with the exception that the WLS fibres are organised into bundles, with the light from each bundle being readout by a single phototube. The number of cells per module decreases with distance from the beam pipe, with modules in the inner region having nine cells, those in the middle region having four, and those in the outer region having only one. All



**Figure 2.18:** Photographs of modules for the ECAL (left), and of the completed ECAL detector (right). Both figures reproduced from [70].

three types of module can be seen on the left of Fig. 2.18. The expected energy resolution of the ECAL is given by

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%, \quad (2.3)$$

where  $E$  is measured in GeV.

### 2.2.5.3 Hadronic Calorimeter

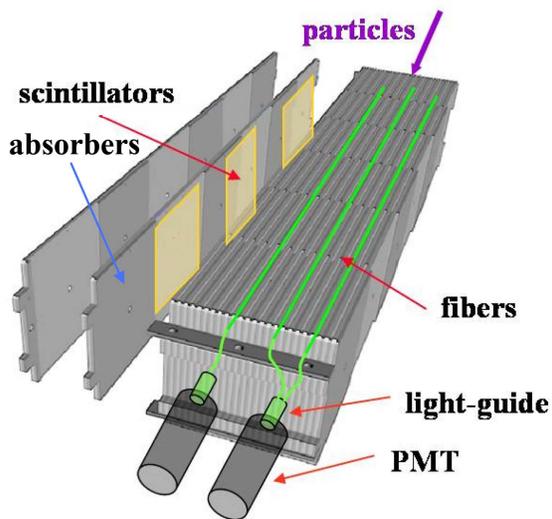
The purpose of the HCAL is to select hadrons with high transverse energy<sup>2</sup> for the Level 0 trigger. It also provides information for particle identification of hadrons. Similarly to the ECAL, the HCAL consists of alternating layers of iron absorber and scintillating material. However in the HCAL the scintillator tiles are oriented parallel to the beam axis. The scintillator-absorber layout is shown on the left of Fig. 2.19, and the segmentation of the HCAL is shown on the right of Fig. 2.17. The scintillator tiles are 4 mm thick, while the iron absorber plates are 12 mm thick. The scintillator-absorber slices are grouped into cells, with each cell read out by a group of WLS fibres that lead to a single PMT. The total radiation length provided by the HCAL is  $5.6\lambda_I$ , where  $\lambda_I$  is the characteristic interaction length for hadrons in steel. This is fewer interaction lengths than the number of radiation lengths in the ECAL because (unlike for the Level 0 photon trigger) the energy resolution necessary for the Level 0 hadronic trigger does not require the full shower to be contained.

The expected energy resolution of the HCAL is given by

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\%, \quad (2.4)$$

where  $E$  is measured in GeV.

<sup>2</sup>The transverse energy  $E_T$  is defined as  $E_T \equiv E \sin \theta$ , where  $\theta$  is the opening angle of the particle's trajectory.



**Figure 2.19:** Schematic diagram of scintillator-absorber layout for the HCAL (left), and photograph of the completed HCAL detector (right). Both figures reproduced from [70].

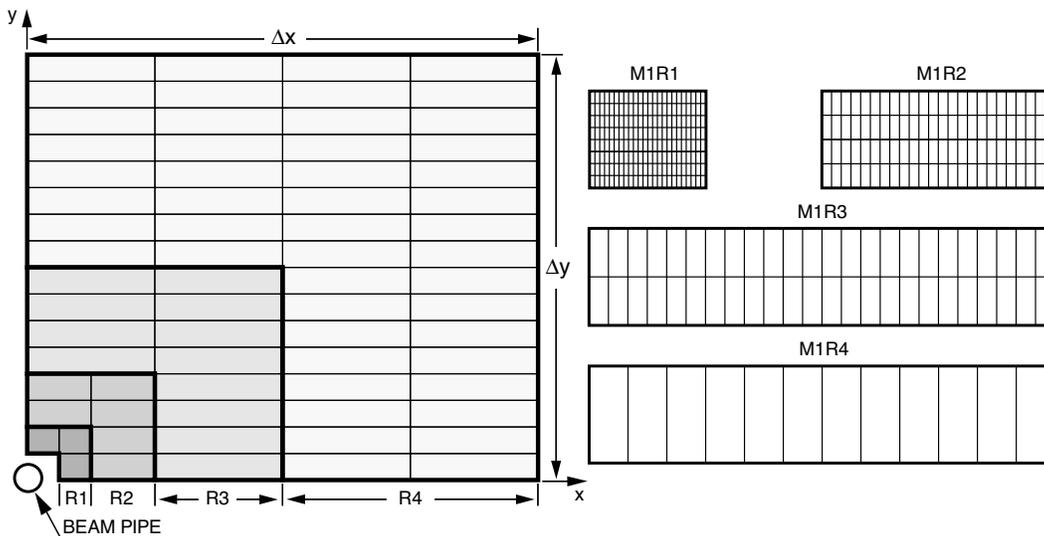
## 2.2.6 Muon System

### 2.2.6.1 Overview

LHCb uses a muon system [70, 89] to provide PID and momentum information on muons for the High Level Trigger and offline analyses, and also to provide transverse momentum ( $p_T$ ) information for the Level 0 hardware muon trigger. The LHCb muon system comprises five muon stations, referred to as M1, M2, M3, M4 and M5 respectively. M1 is located between RICH2 and the SPD/PS, and is only used to provide  $p_T$  information for the Level 0 trigger. The other four stations are located beyond the HCAL. Between each of M2–M5 an 80 cm thick iron filter is used in order to reduce contamination from electrons or hadrons. As a side-effect, low energy muons are also filtered out (a muon with an energy of less than 6 GeV will not reach M5).

Each station is divided into four regions, labelled R1–R4 moving away from the beam pipe. The chamber layout of a muon station is shown on the left of Fig. 2.20. Regions further from the beam pipe are larger with coarser segmentation, so that average particle occupancies across regions are similar.

Two different technologies have been adopted to perform the muon detection. Most of the regions use Multi-Wire Proportional Chambers (MWPCs). Region R1 of M1 uses Gas Electron Multipliers (GEMs). This is because the location of M1, in front of the calorimeter system, means that the particle flux in the R1 region is so high that an MWPC would suffer



**Figure 2.20:** Layout of chambers in a muon station quadrant (left), and segmentation of each region of the quadrant (right). Reproduced from [70].

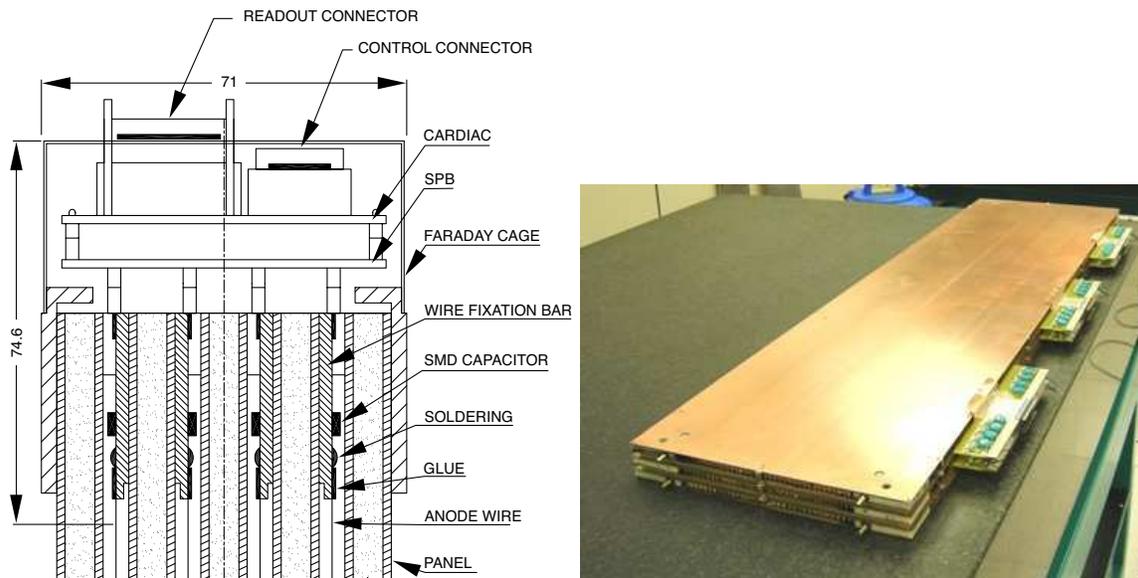
significantly degraded performance over the lifetime of the experiment, due to radiation damage. Elsewhere the flux is low enough that the simpler technology of MWPCs can be used. Both of these technologies and their implementation within LHCb will now be described.

### 2.2.6.2 Multi-Wire Proportional Chambers

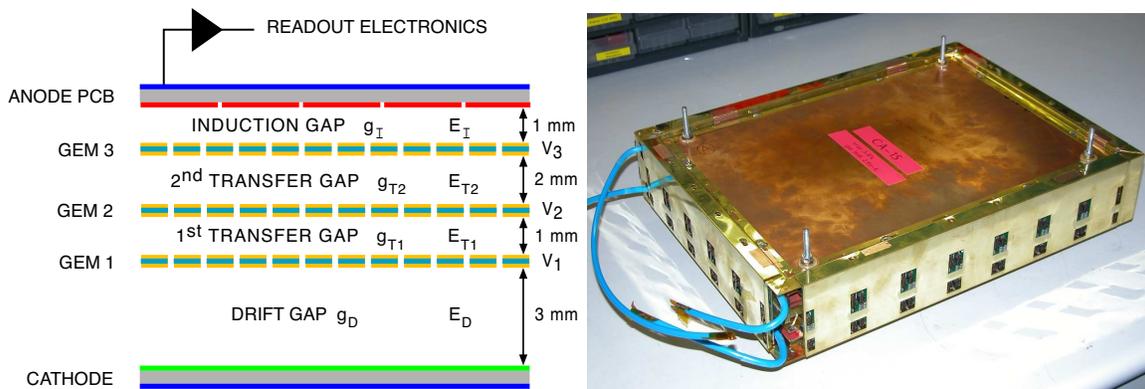
In total the LHCb muon system uses 1368 MWPCs. Within each MWPC layer, gold-plated tungsten wires are held 2.0 mm from each other. The wires sit in the centre of two cathodes (held at ground) that are 5 mm apart. The gap between the cathodes is filled by a gas that is a mixture of  $\text{CO}_2$  (55%), Ar (40%) and  $\text{CF}_4$  (5%). The signal from each pair of adjacent gas gaps is ORed to achieve an efficiency of greater than 95% in a 20 ns window. This efficiency requires a gain of  $\simeq 10^5$ , which is achieved by holding the wires at a voltage of around 2600 V. The MWPCs in M2–M5 have four layers, while those in M1 have two layers. A diagram of a four-layer MWPC and a photograph of a two-layer MWPC are shown in Fig. 2.21.

### 2.2.6.3 Gas Electron Multipliers

Region R1 of M1 is subdivided into twelve chambers, with each chamber containing two triple-GEM detectors. The signal from each pair of GEMs is ORed, achieving an efficiency better than 96% within a 20 ns window. A schematic diagram and a photograph of a triple-GEM detector are shown in Fig. 2.22. Each individual GEM foil is made 50  $\mu\text{m}$ -thick Kapton, with a 5  $\mu\text{m}$ -thick copper layer deposited on each side of the foil. The foil is permeated with bi-conical holes of 50  $\mu\text{m}$  (70  $\mu\text{m}$ ) internal (external) radius. The holes lie 140  $\mu\text{m}$  apart



**Figure 2.21:** Diagram of a four-layer MWPC (left), and photograph of a two-layer MWPC (right). Both figures reproduced from [70].



**Figure 2.22:** Schematic diagram (left) and photograph (right) of a triple-GEM detector. Both figures reproduced from [70].

from each other.

The foils are used in a triple-layer structure, in order to increase the overall gain. The magnitudes of the drift, transfer and induction fields (see Fig. 2.22) are  $E_D = 3.5$  kV/cm,  $E_T = 3.5$  kV/cm and  $E_I = 5.0$  kV/cm respectively. The gas mixture present between the foils was chosen as  $\text{CO}_2$  (15%), Ar (45%) and  $\text{CF}_4$  (40%), as this combination was found to offer the best time resolution (better than 3 ns).

### 2.2.7 Trigger

Due to the very high rate of bunch crossings at the LHC, and the large amount of storage needed for each event, it is not possible to record and store all the data from every single

collision that takes place inside the detector. However, only a small proportion of collisions are of interest for the physics programme of LHCb. Hence a trigger system [70, 90] is used that aims to select only these interesting events, at a low enough rate to allow the data from them to be stored. Data from events judged by the trigger to be uninteresting are not stored, and the event is discarded.

The trigger process is split into two distinct levels, known as the Level 0 Trigger (L0) and the High Level Trigger (HLT). The overall effect of these two levels is to reduce the event rate from 10 MHz (the rate of visible interactions for LHCb at nominal luminosity) to 2 kHz, which is a low enough rate such that all data can be read out and stored. The following sections describe how the L0 and the HLT achieve the necessary reduction in event rates while having high efficiency on interesting events.

### 2.2.7.1 Level 0 Trigger

The Level 0 Trigger reads out information from the VELO pile up system, calorimeters and muon system only. These subdetectors are read out at 40 MHz<sup>3</sup>. The rate of events passing L0 is 1 MHz, which is the speed at which information from the whole LHCb detector can be read out. The operation of L0 is purely hardware-based, due to the speed with which the decision to retain the event or not must be made (the latency time for the L0 decision is 1  $\mu$ s). The L0 uses a combination of the following information when making its decision:

- **High Energy Particle:** The products of  $B$  decays tend to have high transverse energies ( $E_T$ ) and momenta ( $p_T$ ), due to the high mass of  $B$  hadrons. The L0 uses information from the calorimeters and the muon system to identify particles ( $e$ ,  $\mu$ ,  $\pi$ ,  $K$ ,  $p$  or  $\gamma$ ) with high  $E_T$  or  $p_T$ .
- **Track Multiplicity:** Correct reconstruction of  $B$  decay products is very difficult in events with a high number of charged tracks. The L0 uses the number of fired cells in the SPD to estimate the number of charged tracks present in the forward direction. The VELO pile-up system is also used, to estimate the number of charged tracks in the backward direction.
- **Multiple Interactions:** The VELO pile-up system also provides an estimate of how many inelastic  $p - p$  interactions occurred in a given bunch crossing.

The above information is processed by the L0 Decision Unit, which makes the decision to keep or discard the events as follows: unless at least one particle having an  $E_T$  or  $p_T$  above

---

<sup>3</sup>Although the visible interactions only occur at 10 MHz, whether the event has a visible interaction or not is of course not known until some information from the detector is read out.

a certain threshold is found, the event is discarded. Even if such a particle is found in the event, the event may be discarded if the number of charged tracks or number of interactions is judged to be too high. The  $E_T/p_T$  thresholds for each particle type, and the thresholds for rejecting “noisy” events with too many charged tracks or interactions, can be varied as desired. For example the L0 traditionally rejected events with two or more  $p-p$  interactions, but this condition has recently been relaxed. The decision from the L0 is transmitted, via the readout supervisor, to the front end electronics of all subdetectors. If the event has passed L0, then all subdetectors are then read out. This data is sent to the LHCb processing farm, where the event is scrutinised by the HLT algorithms.

### 2.2.7.2 High Level Trigger

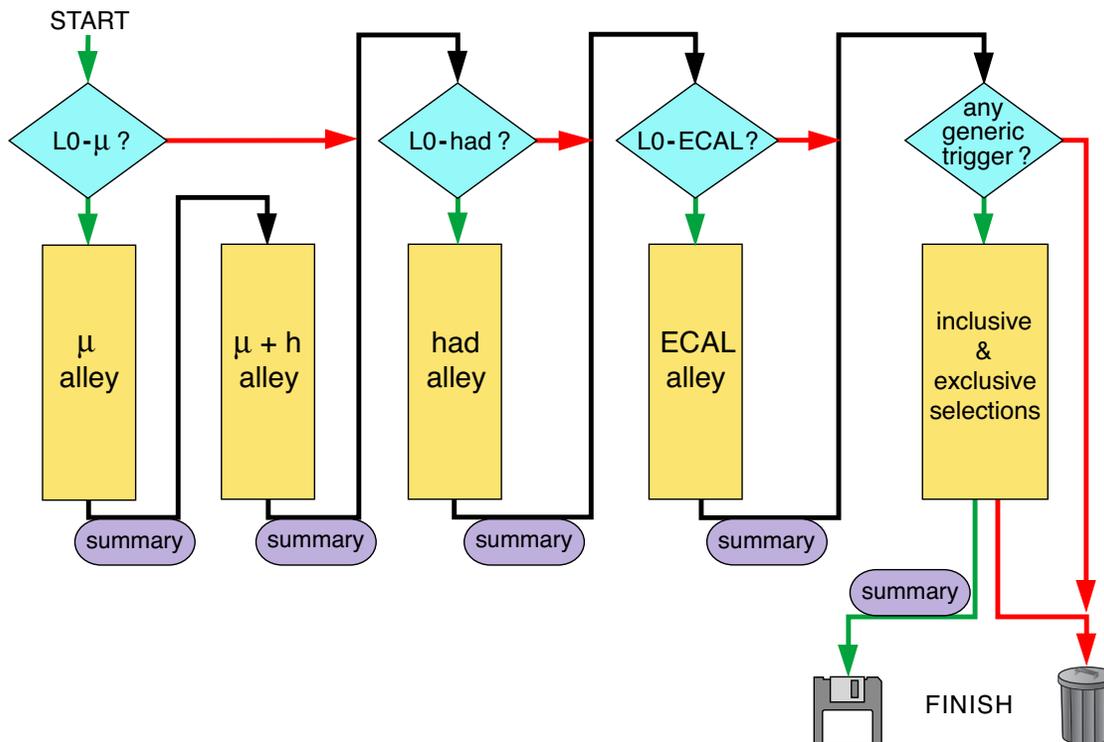
The HLT is a software-based system that runs on the LHCb Event Filter Farm (EFF), a linked group of up to 2000 computing nodes. An overview of the structure of the HLT is shown in Fig. 2.23. The algorithms for the HLT are implemented in C++. The input to the HLT is the data read out from the detector for events that have passed L0. This data arrives at a rate of 1 MHz and includes information from all LHCb subdetectors. The HLT reduces the rate to around 2 kHz, aiming to retain only events that are interesting enough to store permanently. The HLT is subdivided into two stages, known as HLT1 and HLT2. HLT1 is run on each event first, and if the event passes HLT1 it is then passed to HLT2.

The HLT1 stage is carried out as a series of “alleys”. Several different alleys exist, with each one being defined according to the type of particle(s) that were identified at the L0 stage. Only the relevant alley(s) is then run. The alley attempts to find a track in the VELO and T-stations, which corresponds to the L0 object. For example if the L0 object was a muon, the muon alley of HLT1 would attempt to find a track that matches the cluster in the muon system that caused the event to pass L0. In general, for an event to pass HLT1, a track must be found that matches the L0 object, and that has large  $p_T$  and a large impact parameter. The HLT1 reduces the event rate from 1 MHz to around 30 kHz. Events passing HLT1 undergo further analysis in the HLT2 stage.

The HLT2 stage is composed of a number of selection algorithms, each of which is designed to select a particular class of decay. This could be an inclusive decay (for example a  $B$  decaying to two charged tracks, or to a  $J/\psi$  meson plus something else), or an exclusive decay (for example  $B_s \rightarrow \phi\gamma$ ). The event rate coming out of HLT1 is designed to be low enough that a full track reconstruction<sup>4</sup> can be carried out by HLT2, allowing better discrimination between interesting and uninteresting events. For example, the HLT2 selections can cut on variables such as impact parameter significance (i.e. the impact parameter divided

---

<sup>4</sup>However the HLT2 reconstruction is not as accurate as the offline reconstruction, due to time constraints.

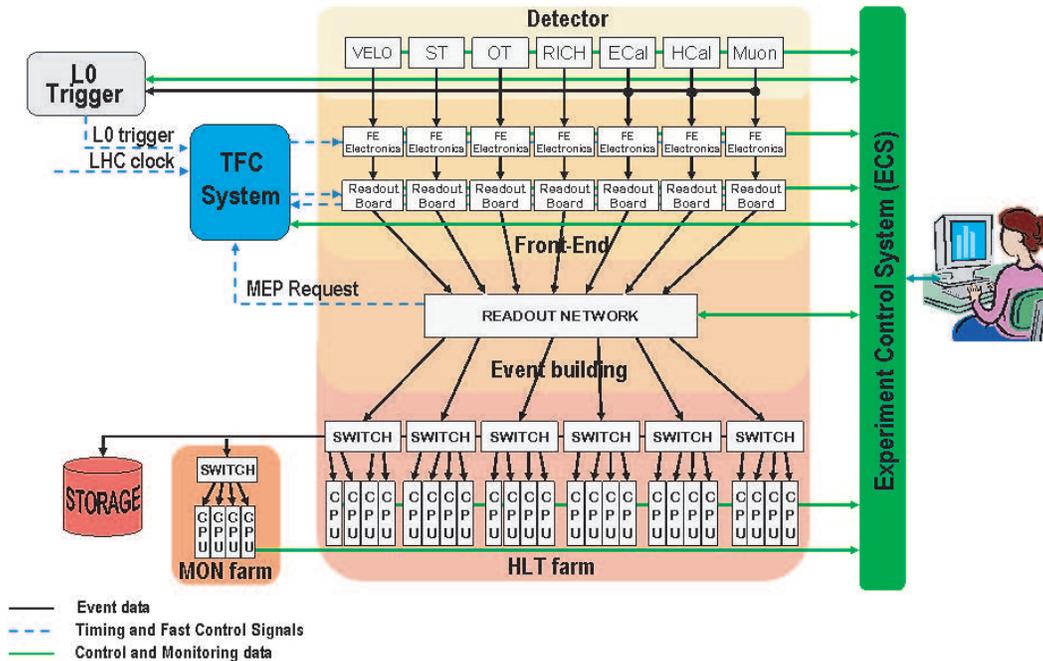


**Figure 2.23:** Overview of the structure of the LHCb High Level Trigger system. Reproduced from [70].

by its error), whereas the HLT1 cuts only on the raw impact parameter. The final decision from the HLT2 stage is taken as an OR of all the HLT2 selections, and the output rate from HLT2 will be around 2 kHz. The issue of which HLT2 inclusive and exclusive selections will be used, and the division of the HLT2 bandwidth between selections, will evolve with the experience of running the detector and the physics priorities of the experiment.

## 2.2.8 Online System

In order to handle the data being read out from the detector and pass it safely to permanent storage, and also to monitor the status and performance of each subdetector, LHCb uses a dedicated online system [70, 91]. The overall architecture of the online system, including how it is linked to the detector readout electronics and the storage of data, is shown in Fig. 2.24. This system consists of three parts: Data Acquisition (DAQ), Timing and Fast Control (TFC) and Experiment Control System (ECS). Each of these subsystems will be briefly described in the following sections.



**Figure 2.24:** Overview of the architecture of the LHCb Online system. Reproduced from [70].

### 2.2.8.1 Data Acquisition

The aim of the DAQ system is to ensure safe passage of the data from the front end (FE) electronics of the detector to the permanent storage facilities. This includes the passage of data to the trigger system, some of which will subsequently fail the trigger and be discarded, and also the passage of data from events that have passed the trigger to permanent storage. The main requirements for the DAQ system are that it is fast enough to cope with the high rates of data being produced by the detector, while being flexible enough to handle different amounts of data produced by different events. Note that the average LHCb event requires around 150 kB of space. The first stage in the DAQ system is to collect the data from the FE electronics. In most subdetectors the readout motherboard used to digitise and pre-process the data before sending it to the EFF is the TELL1 board [92], while the RICH uses the UKL1 board [74]. However the two different boards have the same behaviour from a data readout point of view. These readout boards are controlled by a readout supervisor [93], which sends the boards the trigger signal and the IP address of the EFF node that is to receive the data. Once the data arrives at this node, the HLT algorithms are run, and if the event passes the HLT it is sent to permanent storage.

### 2.2.8.2 Timing and Fast Control

The purpose of the TFC system is to control the DAQ system by providing the readout boards and the FE electronics with a clock signal that is synchronised with the LHC clock (i.e. synchronised with the timing of the bunch crossings). It also provides the readout boards with the L0 trigger decision. The synchronisation of these signals, and their transmission to the readout boards and the FE electronics, are again controlled by the readout supervisor.

### 2.2.8.3 Experiment Control System

The ECS provides control of the entire detector on a far slower timescale than the TFC system. The ECS controls and monitors not only the status of the subdetectors (e.g. status of voltage supplies, gas pressures, temperatures etc.), but also the status of the trigger, DAQ and TFC systems. For the readout of subdetector status, specially-developed radiation-hard serial buses are used. Away from the high radiation zone (e.g. in the electronics barracks and on the surface) commercial ethernet is used. The ECS software is based upon PVSS II [94], which is a commercially available Supervisory Control and Data Acquisition (SCADA) system. The software has to integrate the diverse monitoring hardware across different subdetectors into one system. It is also capable of issuing warnings in case problems are identified in a certain subdetector, and in emergency cases it can take action to prevent the problem causing damage to part of the detector.

## 2.3 Summary

This chapter has provided background information on the overall LHC accelerator complex, and on the LHCb detector and its different components. This information will enable the work of the forthcoming chapters to be understood in the wider context of the whole detector and its capabilities. The RICH detectors were described in detail as the following chapter will describe the concept and operation of the Hybrid Pixel Detectors (HPDs) used in the RICHes, and give results from their characterisation program.

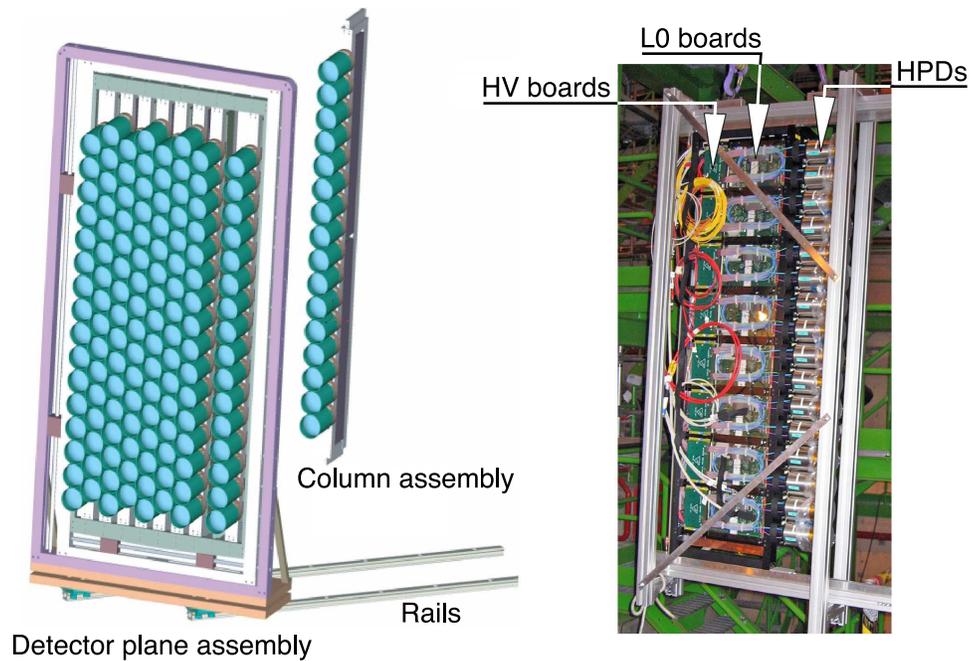
# Chapter 3

## Characterisation of Hybrid Photon Detectors for the RICH

This chapter provides a detailed description of the Hybrid Photon Detectors (HPDs) that are used to detect Cherenkov photons in the LHCb Ring Imaging Cherenkov counters (RICHes), and gives the results from their characterisation. Section 3.1 describes the design and operational principles of the HPDs used in the RICHes. Section 3.2 describes the programme of general quality assurance tests that were carried out on all HPDs manufactured for LHCb. Section 3.3 briefly describes the measurements of the quantum efficiency of the HPD photocathode. Section 3.4 describes in detail the implementation of and results from a specialised test that was carried out on a subset of HPDs. This specialised test, also reported on in [95], measured the photoelectron detection efficiency of the HPD anode. The measurements of quantum efficiency and photoelectron detection efficiency were not carried out on all HPDs due to their time-consuming nature. Section 3.5 summarises and discusses the performance of the HPDs in the general tests.

### 3.1 Hybrid Photon Detectors in LHCb

The two RICH detectors of LHCb, and their importance for the physics programme of LHCb, have been described in Sec. 2.2.2. Both RICH detectors make use of arrays of pixel Hybrid Photon Detectors (HPDs) [79] to detect the Cherenkov photons that are emitted by charged particles as they traverse the RICH. The RICHes use 484 HPDs between them, 196 being in RICH1 and 288 being in RICH2. The HPDs are mounted onto columns, which are then installed side-by-side within the RICHes to form a rectangular array of HPDs. Figure 3.1 shows the column layout for RICH2 and a photograph of a single RICH2 column populated with HPDs. In RICH1 each photodetector plane consists of seven columns, with each column



**Figure 3.1:** Schematic diagram of the column mounting scheme for RICH2 (left), and photograph of a single RICH2 column populated with HPDs (right). Both figures reproduced from [70].

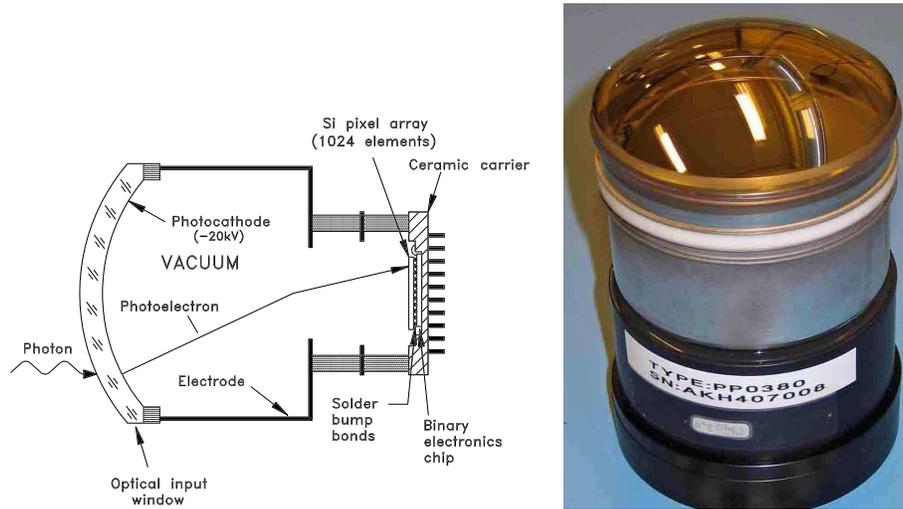
holding 14 HPDs. In RICH2 each photodetector plane consists of nine columns, with each column holding 16 HPDs.

### 3.1.1 Requirements

As the chosen photodetector technology for the RICHes, the HPDs must meet stringent performance requirements if they are to allow the RICHes to provide sufficient Particle Identification (PID) capability over the lifetime of the LHC. These requirements include:

- Ability to detect single photons with wavelengths between 200 and 600 nm
- Time precision of less than 25 ns (the bunch crossing time of the LHC)
- Spatial precision of 2.5 mm x 2.5 mm for Cherenkov photons
- Readout rate of 1 MHz (the rate after the level 0 trigger)
- Large area coverage (around 3.5 m<sup>2</sup>) with high active-to-total area ratio (at least 64%)
- Sufficiently low noise rate
- Ability to operate in a magnetic field of 5 mT (in RICH1).

The next section describes how the HPDs fulfil these requirements.



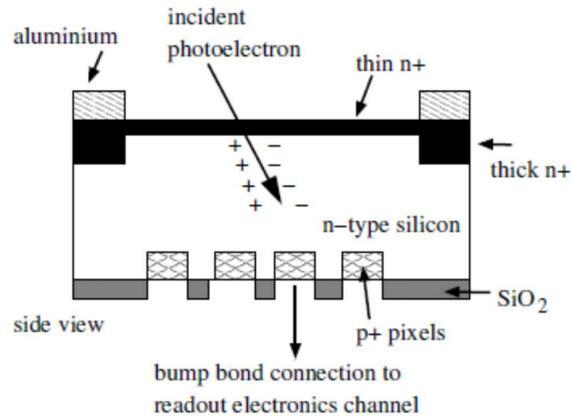
**Figure 3.2:** Schematic diagram (left) and photograph (right) of the pixel hybrid photon detector. Both figures reproduced from [70].

### 3.1.2 Design and Operation

The pixel HPD combines the advantages of vacuum and silicon technology by encapsulating a pixellated silicon sensor inside a vacuum-sealed tube with a quartz window where Cherenkov photons arrive. Figure 3.2 shows a schematic diagram of an HPD. Its basic principle of operation is as follows: an incoming Cherenkov photon reaches the quartz window of the HPD and strikes a layer of photocathode material, liberating a photoelectron. The photoelectron is accelerated toward the silicon sensor by a strong applied electric field. On striking the sensor, the photoelectron creates a large number of electron-hole pairs. The electrons form the signal from which the arrival location of the Cherenkov photon is inferred. The nature of each component of the HPD will now be described in detail.

The HPD quartz window is 7 mm thick and has a diameter of 83 mm, with the active diameter being 72 mm. A thin coating of S20-type multialkali photocathode material is deposited on the vacuum side of the window. To create a high-quality vacuum inside the HPD, the carrier-tube body system is subjected to a bake-out process whereby it is placed in a vacuum and heated at around 350 °C for 5 hours. Gas molecules inside the tube body are boiled off to create a hard vacuum. After the bake-out the quartz window is placed onto the tube body and secured in place with an indium seal.

The photoelectrons that are liberated by Cherenkov photons are accelerated and focused onto a pixellated silicon sensor by a 20 kV electric field. There are four electrodes involved in accelerating and focusing the photoelectron from the photocathode to the silicon sensor. This is known as a tetrode structure. The anode (the silicon sensor) is held at ground, and the photocathode is held at  $-20.0$  kV. The two further intermediate electrodes are situated near



**Figure 3.3:** Cross-section diagram of the HPD silicon sensor. Reproduced from [96].

the anode. They are known as the zoom and focus electrodes, and are held at  $-16.4$  kV and  $-19.7$  kV respectively. The purpose of these extra electrodes is to cross-focus the photoelectrons so that they all strike the silicon sensor, which is smaller than the photocathode. This cross-focusing results in the image from the photocathode being demagnified by a factor of about 5.

The sensor cross-section is shown in Fig. 3.3. It is of the p-on-n type, i.e. it consists of p+ type pixel segmented implants in a near intrinsic n-type substrate, with n+ implants on the backplane layer. This n+ layer forms an ohmic contact with the bulk (lightly doped) n substrate and is kept as thin as possible in order to minimise the energy lost by the photoelectron before it reaches the sensor bulk. A thicker layer of n+ implants around the edge of the sensor backplane, with an aluminium contact frame placed on top, forms the guard ring. A layer of  $\text{SiO}_2$  on the front side of the sensor insulates the sensor bulk from the readout electronics, and the bump bonds from each other.

When a 20 kV photoelectron strikes the silicon sensor, around 5,000 electron-hole pairs are created (the ionisation energy of silicon is 3.6 eV). Beyond the obvious advantage of having a high signal, the large number of electron-hole pairs created causes the downward fluctuations from the most probable value (of the Landau distribution that describes the signal size) to be small, around 1.4%. This means that the signal size stays well above the threshold value, which is around  $1100 e^-$ . The pixellated silicon sensor is bump-bonded pixel-by-pixel to a binary (pixel) readout chip, creating what is referred to as the flip-chip assembly. The use of bump-bonding results in a low capacitance load to the front-end electronics, allowing the sensor readout to be performed at 40 MHz with low noise. It is this low noise rate that makes the signal-to-noise ratio large enough that a binary readout can be used. The binary readout has the advantage that the chip consumes relatively little power (about  $50 \mu\text{W}$  per pixel). The readout chip was a joint development [97] with the ALICE experiment, with

different final versions used by the two experiments. Fabrication of the pixel chip was done using  $0.25 \mu\text{m}$  CMOS technology, with the layout adapted to achieve the necessary radiation tolerance.

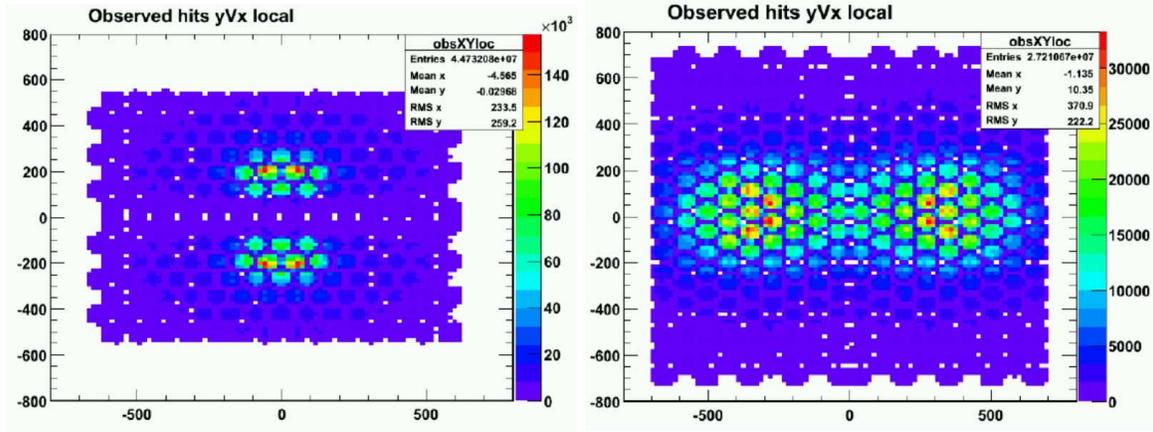
The flip-chip assembly is wire-bonded to a Pin Grid Array (PGA) ceramic carrier, to create the HPD anode. The silicon sensor has 8192 pixels, of dimensions  $62.5 \mu\text{m} \times 500 \mu\text{m}$ . These are known as ALICE pixels. During LHCb running the ALICE pixels are ORed in groups of eight to make 1024 superpixels (or LHCb pixels), in a  $32 \times 32$  square grid. Each superpixel then has dimensions of  $500 \mu\text{m} \times 500 \mu\text{m}$ . Due to the demagnification each pixel represents an area of  $2.5 \text{ mm} \times 2.5 \text{ mm}$  on the photocathode.

The HPD manufacturing process, from fabrication of the silicon wafers through to final tube assembly, involved a number of different international companies, with the final stages carried out by Photonis-DEP<sup>1</sup>. With the requirements as given in Sec. 3.1.1 in mind, a series of specifications [98] for each aspect of the HPDs' performance were drawn up and agreed with DEP. The key specifications are given in Table 3.1. The meaning of each property and its relevance to the performance of the HPD will be discussed in the following sections, as the relevant tests to measure each property are described.

Property	Specification
Working Pixels	95% minimum
Photoelectron Detection Threshold	$1500e^-$ maximum
Photoelectron Detection Noise	$145e^-$ typical
Photoelectron Detection Efficiency	85% typical
Leakage Current at 80 V Reverse Bias	$1 \mu\text{A}$ typical
Quantum Efficiency at 270 nm	20.0% minimum
Quantum Efficiency at 400 nm	15.7% minimum
Quantum Efficiency at 520 nm	8.3% minimum
Dark-Count Rate	$5 \text{ kHz/cm}^2$ typical
Ion Feedback Probability	1% maximum
Lifetime	10 years minimum
Radiation Hardness	30 krad minimum

**Table 3.1:** Selected contract specifications for the manufactured HPDs.

<sup>1</sup>Photonis Netherlands B.V., Dwazziewegen 2, P.O. Box 60, NL-9300 AB Roden, Netherlands. Formerly Delft Electronic Products.



**Figure 3.4:** Expected distributions of Cherenkov photons in RICH1 (left) and RICH2 (right) for collisions that contain a  $b$  quark. The scale and units are arbitrary. Both figures reproduced from [99].

## 3.2 Photon Detector Test Facilities

Two dedicated Photon Detector Test Facilities (PDTFs) were constructed, one at the University of Glasgow and one at the University of Edinburgh. To allow for the possible failure of HPDs during the lifetime of the detector, a total of 550 HPDs were manufactured for LHCb. Quality Assurance (QA) tests were carried out on all of these 550 HPDs between October 2005 and July 2007. The purpose of the QA tests is to ensure that all HPDs going into the RICHes meet the specifications, and also to provide values for certain parameters (for example dark-count rate, ion feedback rate, quantum efficiency and photoelectron detection efficiency) that must be known by the RICH reconstruction software (see Sec. 2.2.2.6) in order to optimise the PID performance of the RICHes. Furthermore, data from the PDTFs is used to decide the placement of HPDs within the RICHes, as the best-performing HPDs should be installed in the highest-occupancy areas. These tend to be in central area of the RICHes, with the occupancy being at its highest in the centre of RICH1. The expected distributions of Cherenkov photons in RICH1 and RICH2 for collisions that contain a  $b$  quark are shown in Fig. 3.4.

Section 3.2.1 describes the setup of the PDTF stations. Sections 3.2.2 and 3.2.3 describe the general test programme that all HPDs were subjected to.

### 3.2.1 PDTF Setup

Each PDTF consists of two test stations. The four test stations are identical to each other, except for additional equipment added to two of the stations to allow specialised tests to be carried out: one station in Edinburgh was used to measure the quantum efficiency of the

photocathode (see Sec. 3.3), and one station in Glasgow was used to measure the photoelectron detection efficiency of the HPD anode (see Sec. 3.4). One station at each PDTF site is used as the default test station where the majority of HPDs are tested in the first instance. The station at each PDTF with the specialised equipment added is used to carry out retests on HPDs that showed behaviour in their first test that requires further investigation, and also of course to carry out the specialised tests. Should the default station develop a fault, the second station is also used as backup for the main test programme while the fault on the default station is repaired.

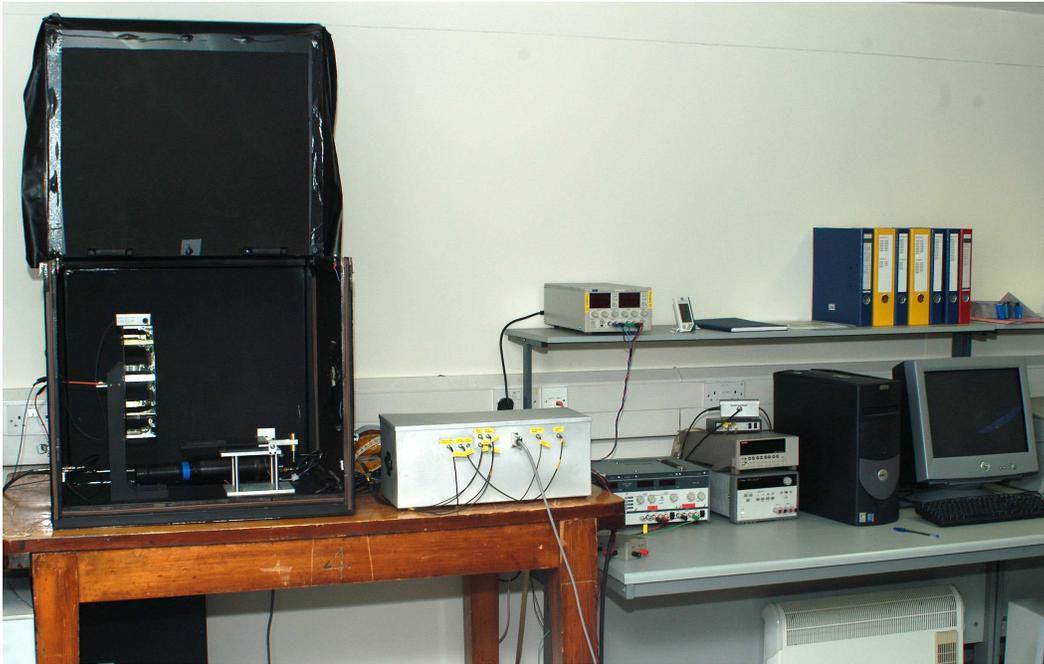
Batches of HPDs arrive on a monthly basis at each PDTF. They are then tested at a rate of about one HPD per site per day, and all HPDs that are judged suitable for use in the RICHes are then shipped to CERN in batches. When not under test, the HPDs are stored in fridges containing a nitrogen atmosphere. Normal air is not a suitable atmosphere for storage of HPDs as it contains a small amount (around 5 parts per million) of helium. Helium atoms are small enough to diffuse through the quartz window, increasing the concentration of gas molecules inside the HPD. This increases the ion feedback rate (see Sec. 3.2.3.2), degrading the performance of the HPD.

Figure 3.5 shows a photograph of the default test station at Glasgow. A schematic of the setup is shown in Fig. 3.6. At the most basic level, a test station consists of:

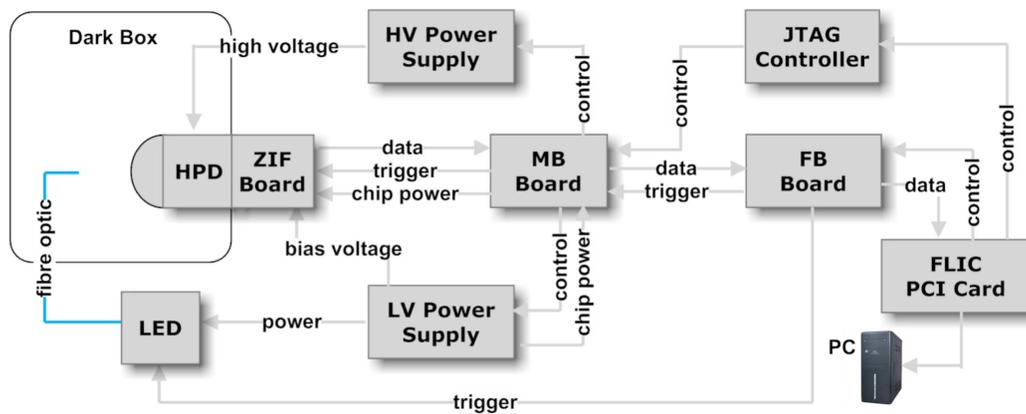
- A light-tight box within which the HPD is mounted, and with an optical fibre feed-through carrying light from an LED.
- A set of electronics devices through which the HPD can be powered, controlled, configured and read out.
- A PC that is used to acquire and store data from the tests.

A photomultiplier tube is placed inside the light-tight box, and is read out using an oscilloscope placed outside the dark box. The PMT is used to check that the light level inside the box once it is closed and sealed is not too high, as exposure to high light levels when the high voltage is on can damage the HPD. The light-tight box also contains a “pinhole mask”, which is a small rectangular slice of plastic with a grid of holes drilled through it. The pinhole mask is mounted onto a small table with a hinge, and sits between the end of the optical fibre and the HPD. It can be raised and lowered via the activation of an electromagnet that is attached to the hinge. The pinhole mask is used in one part of the software test programme (see Sec. 3.2.3.2).

An Agilent E6347A DC power supply provides the 80 V reverse bias voltage to the silicon sensor, while a Keithley 6485 picoammeter is used to measure the leakage current from



**Figure 3.5:** Photograph of the default HPD test station at Glasgow. From left to right can be seen: light-tight dark box (with LED fibre mount, pinhole mask and HPD mount point), high voltage supply box, LED control electronics, pixel chip readout electronics and DAQ PC.



**Figure 3.6:** Schematic diagram of the standard PDTF setup.

the sensor. The voltage that produces the accelerating electric field for the photoelectrons (referred to as the high voltage, or HV) is provided by a Matsusada high voltage power supply, which is internally limited to supply no more than the required 20 kV. A custom voltage divider is used that receives the voltage from the supply, divides it and delivers it at the correct level to the various electrodes along the HPD body (see Sec. 3.1.2). Each clip used to secure the lid of the dark box is interlocked to the HV supply so that the HV drops to 0 V should the dark box be opened while the HV is running.

An LED emitting blue light (470 nm) is used to illuminate the quartz window of the HPD. The light from the LED is delivered to the dark box via a fibre optic cable. The distance from the end of the fibre optic cable inside the dark box to the quartz window is sufficiently large to allow approximately uniform illumination of the window. The intensity of the LED is controlled via a voltage delivered from a Thurlby Thandar PL330DP programmable DC power supply, and is tuned so that the mean number of hits on the pixel chip per LED burst is around 2 to 3. The LED is contained inside a box, external to the HPD dark box, that contains the necessary electronics to allow the frequency and duration of the LED trigger to be tuned. These electronics also allow the tuning of the delay of the trigger with respect to the readout trigger of the HPD pixel chip.

The HPD is mounted inside the dark box such that the Pin Grid Array (PGA) ceramic carrier of the anode sits roughly in the plane of the near side of the dark box, with the quartz window pointing toward the end of the cable carrying the LED light. The signals passing through the PGA are read out using a Zero Insertion Force (ZIF) board, which is mounted directly onto the PGA. To allow the ZIF board to move independently of the rest of the electronics, two double-layer twisted pair cables are used to carry the signal from the ZIF board to the so-called MB and FB printed circuit boards<sup>2</sup>, which form the next layer of electronics [100].

The twisted pair cables are connected via gender adaptors to the MB board. This board supplies power to the pixel chip, and in addition is fitted with a set of ADCs that monitor the voltage and current levels of the various power supplies. Another task performed by the MB board is the adjustment of voltage and current levels of the incoming and outgoing signals to correspond to the logic system used at each stage. The signals produced by the HPD pixel chip use Gunning Transceiver Logic (GTL). These signals are translated to Transistor-Transistor Logic (TTL) for processing on the MB board, and finally to Low-Voltage Differential Signalling (LVDS) logic for use by the FB board. This translation process happens in reverse in the case of signals being sent to the pixel chip. The JTAG (Joint Test Action

---

<sup>2</sup>The MB and the FB board are named after their respective developers at CERN, Mike Burns and François Bal.

Group) protocol is used to configure the MB board and the pixel chip. This interface provides TTL signals that control and monitor the voltages on the MB board and the pixel chip. Connection between the MB and FB boards is provided by two twisted pair cables, one taking signals in each direction.

The FB board generates the clock signal and triggers needed by the pixel chip. These signals are generated by an FPGA (Field-Programmable Gate Array) on the FB board, and are sent to the pixel chip via the MB board. The other main task of the FB board is to format and serialise the data from the pixel chip (received via the MB board) and then send it to a mezzanine receiver card inside the DAQ PC. Transmission to the PC occurs at a rate of 40 MHz along a shielded twisted pair cable, using the S-link protocol [101]. The receiver card de-serialises the data and transmits it to the FLIC (Flexible Input/Output Card) PCI card on which it is mounted. The data is then written to memory by the FLIC card. The FLIC card also generates the readout and reset commands that are passed to the FB board and then on to the pixel chip.

The equipment described above is used to subject each HPD to a rigorous programme of quality assurance tests. This programme is described in detail in the following sections.

### **3.2.2 Visual Inspection and Mechanical Tests**

The first step in the test programme is to remove the HPD from its protective packaging and inspect it for visible defects. During inspection, unpowdered latex gloves are worn to keep the HPD as clean as possible. Also a grounding strap is worn to prevent any electrostatic discharges.

Firstly the pin grid array is inspected to check for bent or otherwise damaged pins. Pins that are only slightly bent can be corrected with a simple mechanical straightening tool, while more severe damage requires expert intervention and may cause the HPD to be unusable. The HV cables are also inspected to check that their insulation is intact. Next, a dry brush is used to clear any dust present within the PGA. Then the PGA is cleaned using de-ionised water and the brush, and dried using a dry air spray. The HV connections inside the HPD are then checked using a multimeter. Each electrode should be electrically connected to one specific part of the HPD, and isolated from those parts connected to the other electrodes. These connections are checked simply by measuring the resistance between the relevant parts of the HPD. If any of the HV connections are not as expected then the HPD needs expert attention and may be unusable.

The final area of the HPD to be inspected is the quartz window. The window is first checked for scratches or other damage, then any dust present is removed with a dry air spray. Next the window is wiped with an ethanol-soaked lens tissue to remove any dirt or residue.

This last step is particularly important as any unwanted material on the quartz window is likely to cause microdischarges in that area of the window once the HV is activated, which means the estimate of the dark-count rate will be affected. After the window is clean, a strip of teflon tape is applied to the indium seal that attaches the window to the rest of the HPD. The teflon insulates the indium seal to prevent microdischarges from occurring once the HV is activated. This completes the visual inspection.

After the visual inspection the mechanical dimensions of the HPD are checked using a custom-made jig. The HPD is mounted into the jig, which is then closed around the HPD. If the HPD has been manufactured correctly, the distance between the rim of the HPD quartz window and the jaws of the jig will then have a uniform value (of about 0.5 mm) all the way round the HPD. If it does not (e.g. if the HPD is leaning to one side) then this may cause problems when the HPD is mounted in the RICH, as the MuMetal shield of the leaning HPD may touch that of a neighbouring HPD.

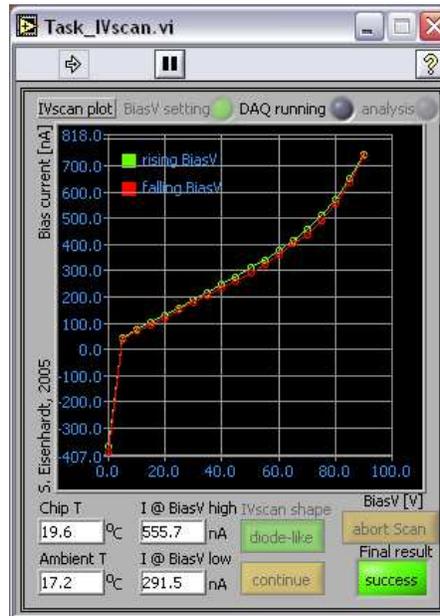
### 3.2.3 Software Tests

Assuming no problems were found during the visual inspection and mechanical tests, the HPD is then mounted inside the dark box, and the HV cables are connected to the correct terminals inside the HV distribution box. The ZIF board is then attached to the pin grid array, and the dark box is closed and sealed. Before the electronics are turned on, the signal from the photomultiplier tube inside the dark box is checked on the oscilloscope. If the rate of single photons is above 10 kHz this probably means there is a light leak inside the box, and the dark box is opened up and the source of the light leak searched for.

After the dark box is sealed and the light level has been checked with the PMT, the HPD can be powered on, and the software-controlled tests can begin. The software used to control the various electronics devices detailed in Sec. 3.2.1, and hence control the HPD, is Labview [102] version 7.1, running on Windows XP. A Labview program has been written that automates much of the test programme, only requiring input from the tester at certain points (for example to confirm at the end of a task that the results are satisfactory and the next task can start). The program also has inbuilt safety checks to protect both the tester and the HPD: for example if the leakage current from the pixel chip exceeds  $10 \mu\text{A}$  at any time, the program instructs the Agilent power supply to drop the bias voltage to the chip to 0 V. Also automated is the writing of results from the tests onto a dedicated hard disk on the PC.

When started, the program displays the main Virtual Instrument (VI) that can be used to run some sequence of sub-VIs that perform individual tasks as described below. A screenshot of the main VI is shown in Fig. 3.7. There are also sub-VIs that display the current levels of important parameters such as the bias voltage, the high voltage and the chip temperature.





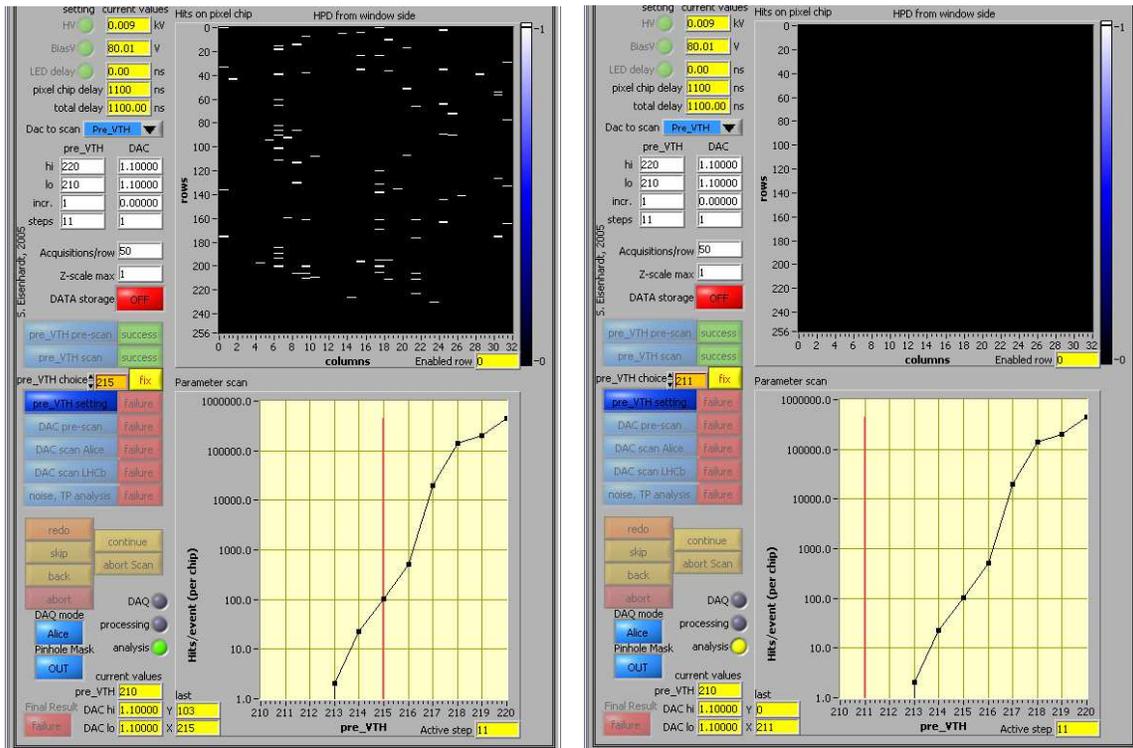
**Figure 3.8:** Screenshot from PDTF showing current-voltage behaviour of HPD pixel chip.

### 3.2.3.1 Pixel Chip Tests

The first task of the software is to power on the pixel chip and ensure that signals can be sent to and received from the chip. Then the responses of the various Digital to Analogue Converters (DACs) within the pixel chip are tested to ensure that they are linear. At this stage the tester also inspects the voltage and currents being read out from the chip (as displayed in the main VI), and compares them to reference values that vary from test station to test station. If the voltage or currents are not as expected this indicates either an internal problem of the pixel chip electronics, or (more likely) a poor connection, either between the chip and the readout boards, or between the readout boards and the power supplies.

The next test is an IV scan, measuring the current-voltage characteristic of the sensor. A screenshot of a typical IV scan is shown in Fig. 3.8. The bias voltage is ramped from 0 V to  $-90$  V and back to 0 V, in steps of 5 V, and the leakage current at each voltage step is recorded. Due to the p-n nature of the sensor, the leakage current dependence on voltage should be diode-like, rather than resistive (linear).

When installed in the RICHes, each HPD is placed in series with a  $2\text{ M}\Omega$  resistor, in order to protect the power supplies in case of a short within the HPD. Hence the leakage current at  $-80$  V acts to diminish the voltage actually depleting the sensor, so if it is too high the sensor may not deplete fully. The actual depletion voltage also affects the drift time of the electrons in the sensor, which in turn affects the exact timing settings required when

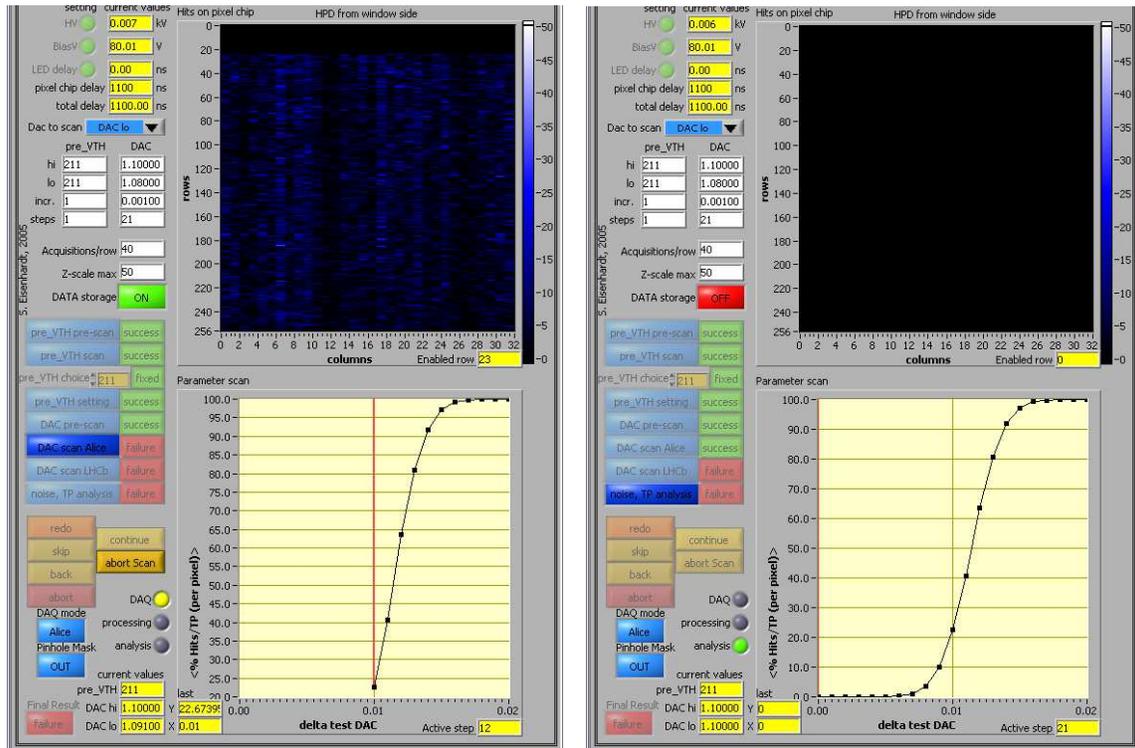


**Figure 3.9:** Screenshots from PETF showing response of the HPD pixel chip with different thresholds. Left, the response at a lower than desired threshold, where some hits can still be seen. Right, the response at the chosen threshold, where no hits are recorded. This particular HPD does not have any noisy pixels.

reading out the HPD. Finally, high leakage currents cause more heat to be dissipated in the chip, increasing the strain on the RICH cooling systems. After the IV scan the bias voltage is increased to  $-80$  V and held there for the remaining tests.

The next test is known as the threshold scan. It calibrates the voltage discriminator threshold within the chip. This is the voltage above which a signal (i.e. the voltage pulse formed from the charge collected in a given pixel) is considered as constituting a “hit” on that pixel. This threshold is controlled via the DAC supply line known as “pre\_VTH”. The higher the value of pre\_VTH, the lower the threshold. The threshold scan consists of several stages. In each stage the pixels are read out in ALICE mode, to allow individual ALICE pixels to be identified as dead or noisy. Noisy pixels are not masked online, rather they are masked during the processing of the test data after the test programme is complete.

The first stage of the threshold scan involves changing the value of pre\_VTH and observing the amount of electronic noise seen across the chip (no test pulse is injected at this stage). Screenshots of this first stage of the threshold scan are shown in Fig. 3.9. At the end of the scan pre\_VTH is fixed by the tester to be one DAC setting lower than the highest value



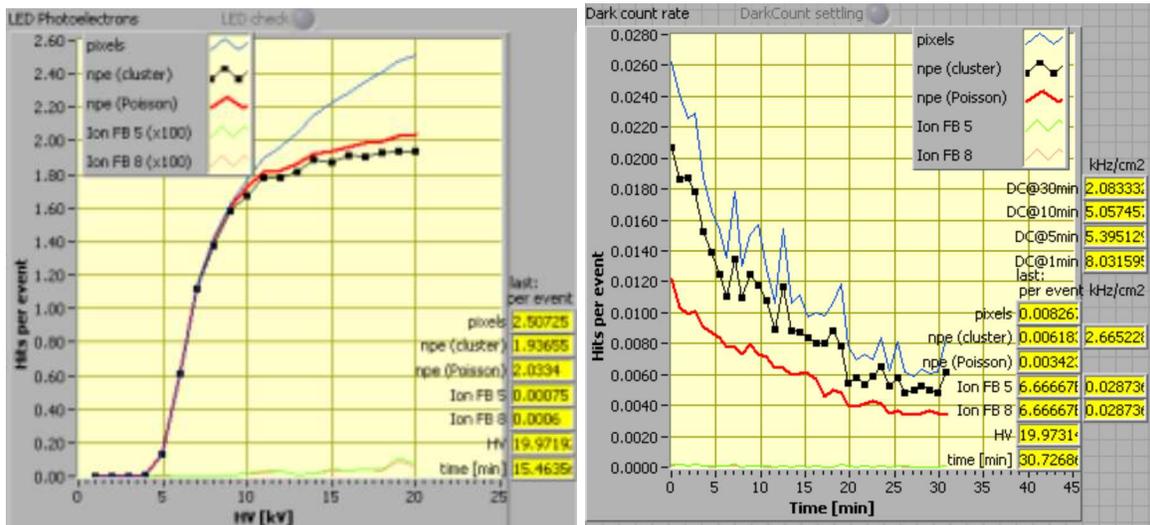
**Figure 3.10:** Screenshots from PETF showing response of an HPD pixel chip to a test pulse. The conversion factor between the DAC value and the number of electrons injected is 100,000, e.g. 0.01 corresponds to 1,000 electrons. Left, the response to a test pulse of 1,000 electrons. Right, the full S-curve after all stages of the scan are complete.

that gives no noise hits across the whole chip. This choice minimises the electronic noise while maximising the chance for genuine photoelectron hits to be detected. The program allows the tester to scroll back through pre\_VTH values to inspect each hitmap, so that any noisy pixels (that consistently give hits at lower pre\_VTH values than all other pixels) can be discarded when choosing the optimal pre\_VTH value.

Once the pre\_VTH value has been chosen, it is held at this value and a second type of scan is performed. Screenshots of this second stage of the threshold scan are shown in Fig. 3.10. At each step of this scan, a test pulse of a given charge is injected into the chip. The test pulse is injected row-by-row to avoid overloading the chip. Forty triggers are sent to each pixel and the number of responses from each pixel is recorded. As the test pulse decreases, the pixel response forms an S-curve. The derivative of this S-curve is a Gaussian. For an individual pixel<sup>3</sup>, the mean of this Gaussian gives the threshold value for that pixel, and the width gives the amount of electronic noise affecting that pixel.

The software can also identify any pixels that may be dead/poor (i.e. no response or

<sup>3</sup>The software displays only one S-curve, which is an average after taking all pixels into account.



**Figure 3.11:** Screenshots from PDF showing HPD hit rate during high voltage ramp-up (left), and dark-count rate following high voltage ramp-up (right). Over 30 minutes the dark-count rate drops from 9.2 kHz/cm<sup>2</sup> to 2.6 kHz/cm<sup>2</sup>.

very low response regardless of the size of the test pulse) or always-on/noisy (i.e. always or usually respond regardless of the size of the test pulse) using their S-curves. Such pixels can be taken into account offline when analysing the test data<sup>4</sup>.

Comparing the means for different pixels allows the spread of thresholds across pixels to be measured. The threshold spread and the individual pixel noise both contribute to the noise observed when reading out the pixel chip. Each pixel has a 3-bit threshold adjust that allows fine-tuning of thresholds on a per-pixel basis. However in practice it is found that the spread of thresholds across pixels is narrow (typical width is 90 electrons), hence this functionality is not used and the same threshold (dictated by the chosen value of pre\_VTH) is used across the whole chip.

### 3.2.3.2 Tests Under High Voltage

Once the discriminator threshold for the chip has been chosen, the high voltage can now be ramped up in the knowledge that any hits on the pixel chip should come from genuine photoelectrons and not electronic noise. The response of a typical HPD during the HV ramp-up is shown on the left of Fig. 3.11. The high voltage is ramped up slowly, at a rate of 50 V/s. After each kV increase, the photocathode is illuminated with light from the calibration LED. Typically no hits are recorded on the pixel chip until the HV reaches 5 kV. This is expected as the number of electron-hole pairs produced by a 5 kV photoelectron striking the sensor is around 1300, so below 5 kV the signal is insufficient to pass the typical discriminator

<sup>4</sup>In the full RICH monitoring software such pixels can be dealt with online.

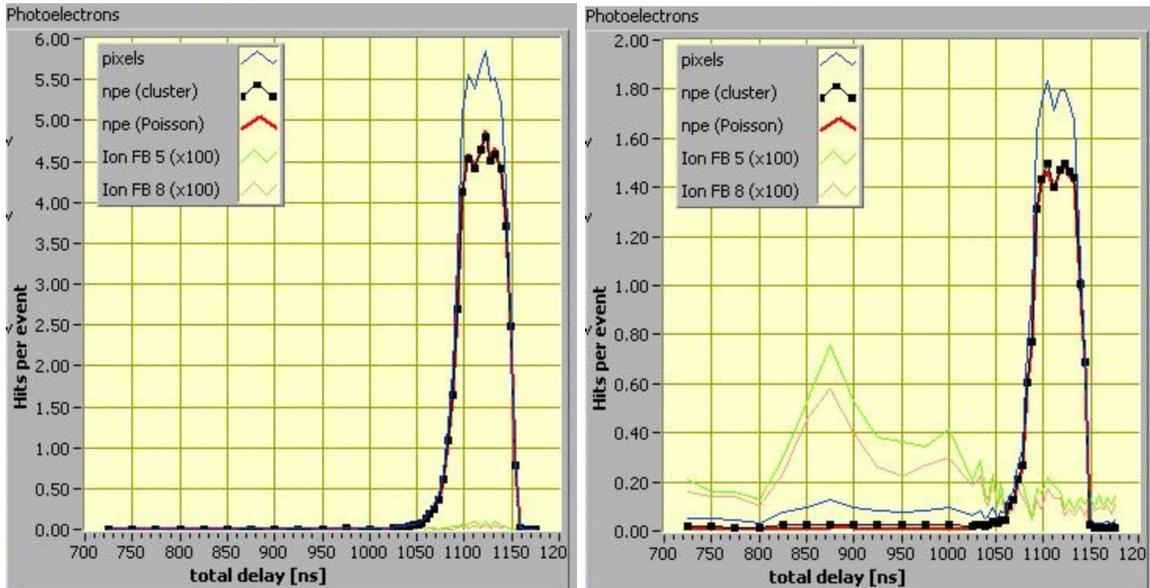
threshold (see Table 3.1). Between 5 kV and  $\simeq 10$  kV the hit rate rises sharply, then above 10 kV the rise is more shallow.

The hit rate is calculated using three methods, with all three of these rates being shown in the VIs. The first and simplest method just counts the number of pixels firing each time, and takes the mean. The second method uses a clustering algorithm to form clusters from adjacent hit pixels, and counts the number of clusters, again taking the mean. The third method calculates the average number of hits by counting the frequency with which no hits are recorded, and assuming the hit distribution is Poisson. In general the second and third methods give very similar results, while the first method gives somewhat higher rates, as often clusters are formed from more than one pixel (due to for example charge sharing between adjacent pixels). As stated previously, the intensity of the LED is tuned such that the average number of hits per LED burst at 20 kV is  $\simeq 2-3$ .

Once the HV reaches 20 kV, it is held there and the hit rate without any illumination from the LED (i.e. the dark-count rate) is measured for 30 minutes. The dark-count rate typically shows a settling effect over the 30 minutes as the HV stabilises and the rate of any microdischarges drops. The evolution of the dark-count rate after the HV ramp-up for a typical HPD is shown on the right of Fig. 3.11. Very high dark-count rates are not desirable during LHCb running as they create spurious hits on the photodetector plane, reducing the performance of the PID algorithm. Hence for optimal performance the dark-count rate should be as low as possible.

After the dark-count measurement a so-called strobescan is performed. Screenshots of strobescans for both a typical HPD and an HPD with large ion feedback probability (see below) are shown in Fig. 3.12. Each step in the strobescan involves illuminating the HPD with the LED with a different delay time between the triggering of the LED and the readout of the pixel chip. The readout trigger to the chip (generated on the FB board) can be moved in steps of 25 ns, while the LED trigger can be moved in steps of 2 ns. The strobescan serves two purposes: firstly it checks that the default delay (used in the rest of the tests that involve illumination of the HPD) is the delay that maximises the observed number of photoelectrons. Secondly, and more importantly for the characterisation of the HPD, it measures the ion feedback rate within the HPD.

Ion feedback (IFB) occurs when a photoelectron strikes a residual gas molecule inside the HPD, ejecting an electron from the molecule and creating a positive ion. Due to the applied field from the HV, the ion then moves toward the photocathode. When the ion strikes the photocathode, a large number ( $\simeq 30-40$ ) of photoelectrons are produced simultaneously due to the large amount of energy deposited by the ion. These photoelectrons are then accelerated toward the anode, where they cause large clusters of hits (typically more than 10 pixels). So



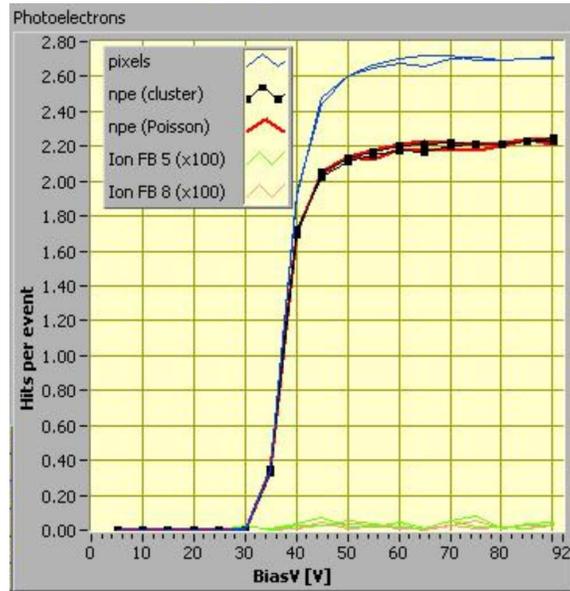
**Figure 3.12:** Screenshots of strobescans at PDTF. The delay shown refers to the trigger of the LED, so that a larger value corresponds to a shorter time between the LED firing and the chip being read out. The IFB rate is shown multiplied by a factor of 100 on the VI, to enable it to be seen next to the main peak from the direct photoelectrons. The main peak occurs at a timing of around 1100ns, while the hits from IFB occur at around 850 ns. Left: A typical HPD with negligible IFB rate (0.008%). Right: An atypical HPD with large IFB rate of 0.50%.

the rate of very large clusters is a measure of the quality of the vacuum inside the HPD. Due to the cross-focusing optics inside the HPD, most clusters from IFB photoelectrons will strike near the centre of the anode.

If the density of residual gas molecules inside the HPD is too high, then a point is reached where the probability of at least one of the  $\approx 30$ – $40$  photoelectrons produced by the ion striking another molecule on its way to the anode reaches 100%, so that an avalanche effect occurs. The photoelectrons from IFB will then dominate over the signal photoelectrons, and the HPD will no longer be usable.

For the purposes of the strobescan, a cluster is defined as being the result of ion feedback if it is larger than five pixels in size. If the ion feedback rate is non-negligible, a peak in the rate of large clusters can be seen when the delay is about 250 ns longer than the optimal delay for detection of standard photoelectrons. This corresponds to the extra time taken for the ion to reach the photocathode, and for the resulting photoelectrons to reach the anode.

Following the strobescan, a scan in bias voltage is performed. A screenshot of a typical bias voltage scan is shown in Fig. 3.13. The bias voltage is changed in steps of 5V, dropping from 80 V to 0 V and then rising back up to 80 V. The HV is held at 20 kV and the HPD is



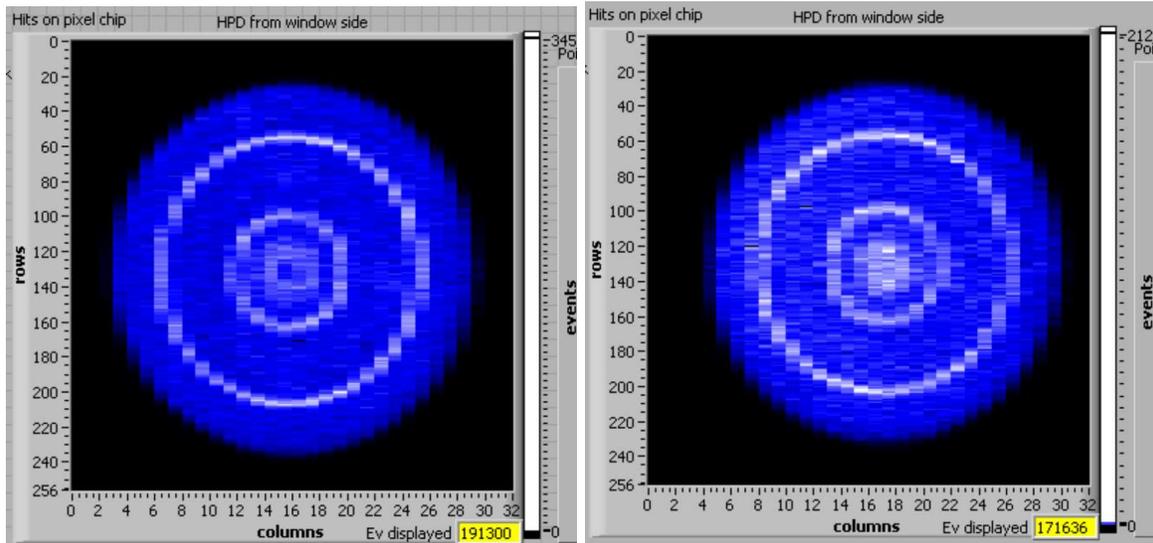
**Figure 3.13:** Screenshot of bias voltage scan at PDTF.

illuminated at each step. The main purpose of the bias scan is to confirm that the sensor responds as expected to different biases. Typically there is no activity below  $\simeq 35$  V, and above 35 V the response turns on very sharply, and plateaus above  $\simeq 50$  V where full depletion is reached. If the level of activity fails to plateau this may indicate that the LED delay is not optimal.

The final scan performed is a scan in HV. The HV is ramped down to 0 V and then back up to 20 kV in steps of 1 kV. The HPD is illuminated at each step. The number of photoelectrons observed as a function of the applied HV should follow a similar curve to that observed during the HV ramp-up.

Following the various scans, several stable runs of different types are performed. The first is the “long LED run”, where the HPD is illuminated over 200,000 LED bursts, with the HV and bias voltage at their nominal values. The long LED run allows the position and size of the photocathode image to be precisely determined, and also allows pixels that respond poorly (or not at all) to light to be identified. The long LED run is carried out twice, once with the pixels read out in ALICE mode and once with the pixels read out in LHCb mode. The ALICE run allows individual faulty pixels to be found, while the LHCb run checks for problems with the OR logic that makes the superpixels. Screenshots of long LED runs in ALICE mode for two different HPDs, illustrating the displacement of the photocathode image, are shown in Fig. 3.14.

It can be seen from Fig. 3.14 that the photocathode image does not fill the entire pixel chip. Rather it sits near the centre of the pixel chip, leaving a clear boundary around the image where no hits are visible. A very small number of hits do occur outside this main

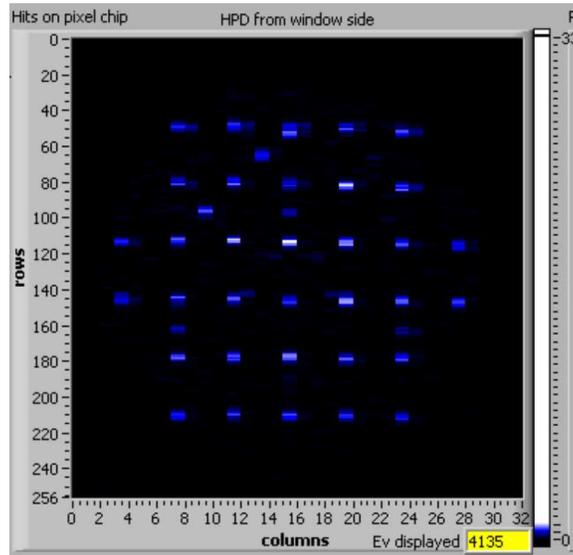


**Figure 3.14:** Screenshots of long LED runs in ALICE mode at PDTF. On the left, an HPD with little displacement of the photocathode image. On the right, an HPD with significant displacement (by 1.2 LHCb pixels) of the photocathode image in the positive  $x$  direction.

photocathode image, due not to electronic noise but rather to genuine photoelectrons that have arrived at the sensor within the normal photocathode image area, undergone backscattering and subsequently landed outside the normal image area (see Sec. 3.4.3 for a discussion of backscattering). Such hits however are not visible in Fig. 3.14, due to the high  $z$  scale required to allow the hit distribution in the photocathode image to be seen.

If the centre of the photocathode image is significantly displaced from the centre of the chip, due to for example irregular cross-focusing of the photoelectrons, some photoelectrons may miss the sensor and be lost. The possibility of distortion of the image due to the residual field from the LHCb magnet means that, even if the image is contained within the sensor at the PDTF, within the RICH part of the image may fall outside of the chip. Hence a clear boundary around the image as seen at the PDTF is desirable, to act as a safety margin.

In addition to locating the centre of the photocathode image, the long LED run also reveals the structure of hit frequency across the image. The hit distribution (see Fig. 3.14) is not uniform, but varies with radius. Three concentric rings of higher activity can be seen. These rings appear due to those incoming photons which strike the chromium coating on the underside edge of the quartz window. If the photon reflects back from the chromium coating it may then undergo total internal reflection between the outer and inner surfaces of the quartz window. Each time the photon strikes the inner surface, it may cause a photoelectron to be emitted. The three rings that are observed correspond to the three possible radii where the photon can strike the inner surface during this total internal reflection, before the angle of

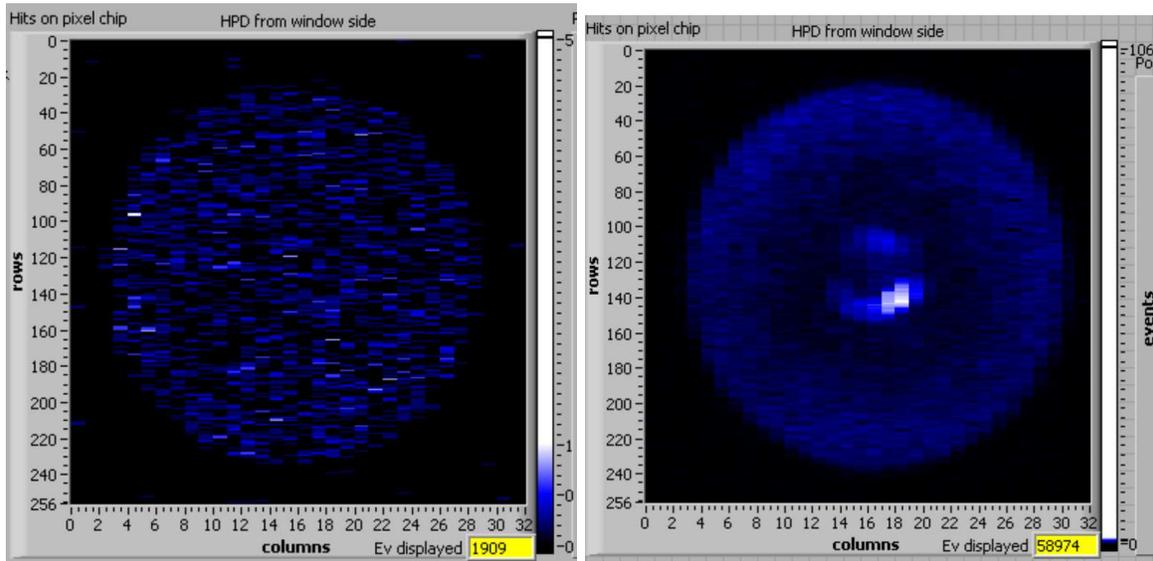


**Figure 3.15:** Screenshot of a distortion map run at PDTF. This HPD shows no significant rotation or distortion of the photocathode image.

incidence becomes too small for total internal reflection to continue (a detailed discussion of these effects, including a full simulation, can be found in chapter 5 of [99]). At the PDTFs these reflections cause concentric rings because the photons originate from a point source and arrive approximately perpendicular to the quartz window. During LHCb running the Cherenkov photons do not originate from a point source, and the reflections instead cause a small signal-correlated background to appear. This background is taken into consideration during the RICH reconstruction process.

After the long LED runs, the pinhole mask (see Sec. 3.2.1) is raised. This is done automatically by the software, which activates the electromagnet, pulling the mask up from a horizontal position to be vertical and between the end of the optical fibre and the HPD. The holes in the mask have a diameter of 1 mm and a pitch of 11 mm. They are arranged in a square grid, with two extra holes near the centre so that the absolute orientation is known. Data is taken over 200,000 LED bursts, then the pinhole mask is lowered again. A screenshot showing the resulting hit pattern for a typical HPD can be seen in Fig. 3.15. The pattern of hits from light passing through the pinhole mask allows the demagnification of the photocathode image to be measured. It also allows any rotation of the chip with respect to its expected orientation to be identified. For these reasons this test is known as the distortion map.

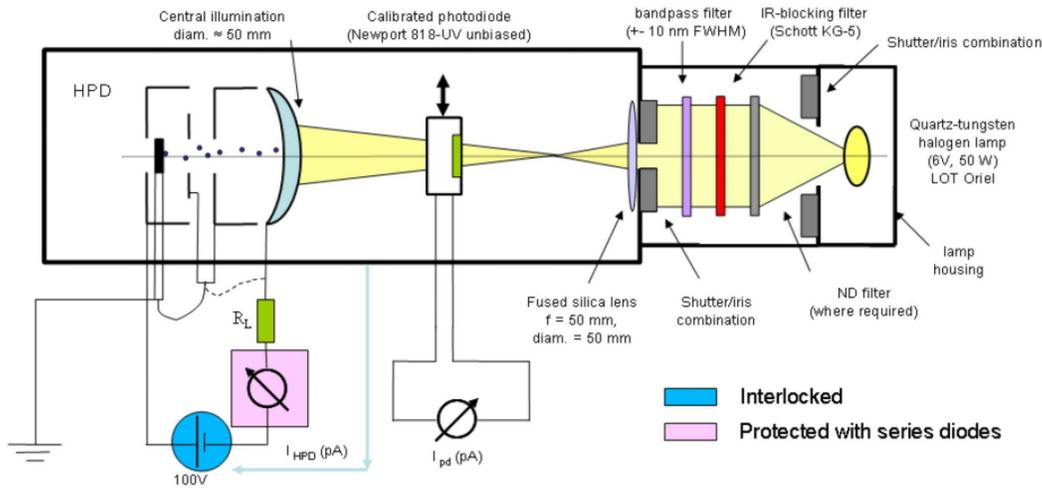
Finally, after the run with the pinhole mask, a dark-count run is made (again in both ALICE and LHCb mode). Data is accumulated from 5,000,000 triggers without illumination from the LED. The dark-count run is performed after all the other tests to give the dark-count



**Figure 3.16:** Screenshots of dark-count runs at PDTF. On the left, an HPD with low dark-count rate ( $0.16 \text{ kHz/cm}^2$ ) spread uniformly across the photocathode image. On the right, an HPD with significant ion feedback rate (1.3%), manifesting itself as increased activity in the centre of the photocathode image. The dark-count rate for this HPD was measured to be  $5.5 \text{ kHz/cm}^2$ .

rate time to stabilise, so that this run can give the best possible estimate of the dark-count rate in the HPD. The dark-count runs also give another estimate of the IFB rate, from counting the rate of large clusters of hits. Any source of microdischarges can also be identified from observation of an increased hit rate in a localised area of the photocathode image. Screenshots of dark-count runs in ALICE mode for two different HPDs, illustrating the effect of ion feedback, are shown in Fig. 3.16. It can be seen from Fig. 3.16 that the dark-count hits mostly fall within the area of the photocathode image. This shows that these hits are not from electronics noise, but rather from genuine photoelectrons thermally emitted from the photocathode.

The dark-count runs are the final stage of the software testing. After they are finished, the tester can instruct the software to either rerun one or more of the individual tests (if some issue was found in a certain test), or to begin the shutdown sequence, where the HV and bias voltage are ramped down and the pixel chip powered off. After the shutdown sequence is completed, the tester turns off all the electronics, dismounts the HPD and returns it to storage. The collated results for all production HPDs of the tests that have been described in this section are given in Sec. 3.5.



**Figure 3.17:** Schematic of the setup used to measure HPD quantum efficiency at PDTF. Refer to the text for details. Reproduced from [99].

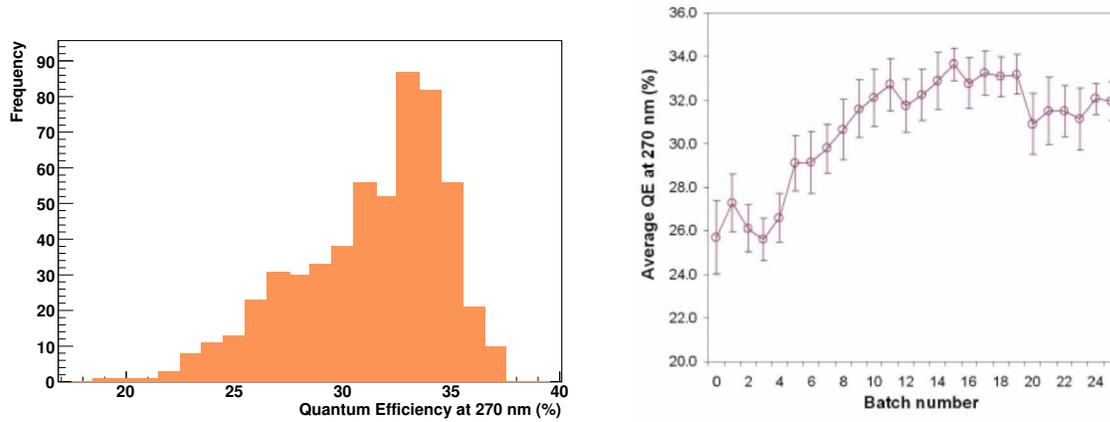
### 3.3 Quantum Efficiency Measurements

This section summarises measurements of the quantum efficiency of the production HPDs. These measurements were carried out at DEP and at the PDTF at the University of Edinburgh. A full and detailed account of the measurements can be found in chapter 4 of [99].

The quantum efficiency (QE) of a photocathode is defined as the ratio of the number of photoelectrons emitted by the photocathode to the number of photons incident on the photocathode. For any photocathode material the QE is a strong function of the wavelength of the incoming photon. The S20-type multialkali material chosen as the photocathode for the LHCb HPDs has a QE spectrum that peaks in the relevant region for Cherenkov photons produced in the LHCb RICHes, which is 200 nm–600 nm.

All HPDs had their QE measured by Photonis-DEP just after production, with around 10% of HPDs having the measurement repeated at the PDTF. The PDTF measurements serve to validate the DEP results, and also allow more detailed studies to be carried out. Note that the quantity that was measured is not the bare QE of the photocathode, rather it is the probability of photoelectron emission given an incident photon at the quartz window. This quantity will be sensitive to such effects as reflection at the outside of the window, and scattering within the photocathode. However this is not a problem, as it is this “overall” QE that is the relevant quantity for characterising the HPD performance.

The method used, which was similar at DEP and at the PDTFs, will now be briefly described. A schematic of the setup used to measure QE at PDTF is shown in Fig. 3.17. The HPD is mounted in the dark box as normal, but illumination comes from a quartz-tungsten-



**Figure 3.18:** Left: Distribution of quantum efficiency at 270 nm for all HPDs, as measured by DEP. Right: Evolution of QE at 270 nm with batch number, again as measured by DEP.

Property	Mean Value	RMS
Quantum Efficiency at 270 nm	30.8%	3.4%
Quantum Efficiency at 400 nm	24.2%	2.4%
Quantum Efficiency at 520 nm	13.4%	1.4%
Quantum Efficiency at 800 nm	0.5%	0.8%

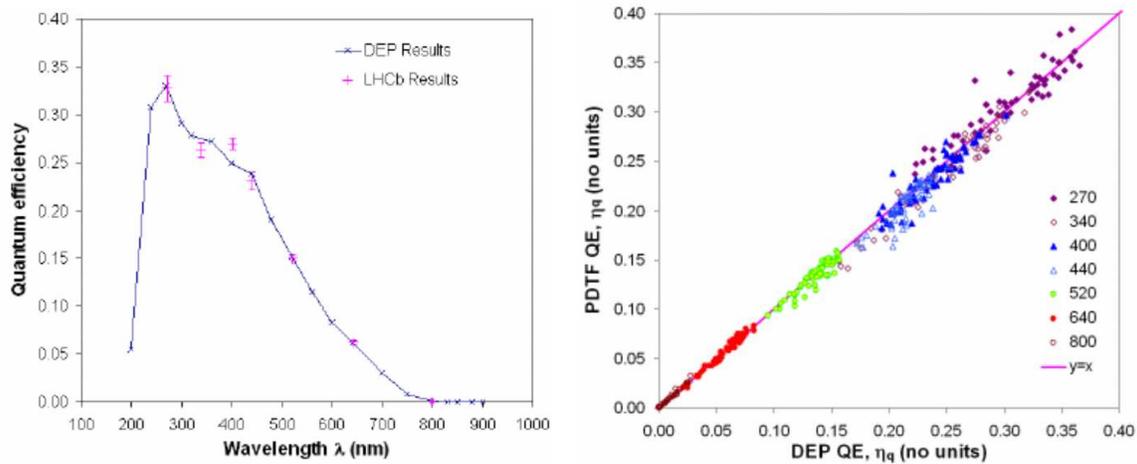
**Table 3.2:** Quantum efficiency performance of the manufactured HPDs, as measured by DEP. Values given are the raw mean and RMS of the distribution.

halogen bulb rather than an LED. A series of different wavelength filters are used to allow the QE to be measured as a function of wavelength. The light output from the lamp is measured by placing a calibrated photodiode (PD) of known QE ( $QE_{PD}$ ) between the light source and the HPD, and measuring the resulting photocurrent,  $I_{PD}$ . The photodiode is then removed and the photocurrent  $I_{HPD}$  produced at the HPD photocathode (which is biased to  $-100V$ ) is measured. The QE of the HPD is then simply:

$$QE_{HPD} = QE_{PD} \frac{I_{HPD}}{I_{PD}}. \quad (3.1)$$

The distribution of QE at 270 nm for all HPDs (as measured by DEP) is shown on the left of Fig. 3.18. The average QE values (again as measured by DEP) at key wavelengths are given in Table 3.2.

The average QE at 270 nm was found to be 30.8%, which is a relative improvement of 32% over the typical value of 23.3% specified in the contract (the contractual value was based on measurements on pre-production HPDs). The reason for this improvement is refinement by DEP of their procedure for fabrication of the photocathode. This refinement is reflected in the improvement of the QE at 270 nm over the production process, which increased from



**Figure 3.19:** Comparison of quantum efficiency measurements at PDTF and DEP. Left: QE spectrum for a single HPD as measured by DEP and by LHCb at PDTF. Right: scatterplot of QE at different wavelengths as measured by DEP and by LHCb at PDTF, showing all 60 HPDs that had their QE measured at PDTF.

$\simeq 26\%$  for early batches of HPDs to  $\simeq 32\%$  for the last batches produced. The evolution of QE with batch is shown on the right of Fig 3.18. The improved photocathode fabrication also led to a decrease with batch number of the QE at 800 nm, from  $\simeq 1.4\%$  to  $\simeq 0.3\%$ . This is beneficial as lower QE at 800 nm reduces the thermal emission of photoelectrons and hence reduces the dark-count rate.

These improvements in QE will increase the number of detected Cherenkov photons in the RICH, and thus improve the PID performance. The increase in photon yields compared to the expectations from preseries HPDs is expected [99] to be 22% in RICH2 and 36% for the gas radiator of RICH1. The increase is larger for RICH1 as the HPDs with the highest QE were placed in the central region of RICH1, which is the highest occupancy area of the RICHes. A preliminary simulation [99] suggests that with the photon yields increased as above, the kaon identification efficiency can be kept roughly constant while achieving a relative 32% decrease in the pion misidentification rate.

A total of 60 HPDs had their QE spectra measured at PDTF. In general, excellent agreement is found between these spectra and the QE spectra as measured by DEP. This agreement, illustrated in Fig. 3.19, reinforces confidence in the DEP measurements.

In summary, the quantum efficiency of the HPD photocathode has been measured by both DEP and PDTF. These measurements are consistent with each other, and indicate a significant increase in the quantum efficiency at the relevant wavelengths, compared to expectations. This increase will have a positive impact on the PID performance of the RICHes.

## 3.4 Measurement of the Photoelectron Detection Efficiency of the HPD Anode

This section describes measurements carried out on two HPDs at the PDTF at the University of Glasgow. The purpose of these measurements is to find the single photoelectron detection efficiency,  $\eta$ , of the HPD anode. Knowledge of  $\eta$  is required for an accurate simulation of the RICH detectors. A direct measurement of  $\eta$  involves comparing the average number of photoelectrons seen by the readout chip to the number arriving at the backplane of the silicon sensor. Since the measurement at the backplane is the interesting and challenging part of the efficiency measurement, the whole measurement is known as the “backpulse” measurement.

Section 3.4.1 explains the importance of measuring the single photoelectron detection efficiency, and Section 3.4.2 describes the general principles of the method to directly measure it at the PDTF. Section 3.4.3 outlines the effects within the sensor that cause  $\eta$  to be less than unity. The experimental setup and data taking methodology used at the Glasgow PDTF to carry out the measurement of  $\eta$  are respectively outlined in Secs. 3.4.4 and 3.4.5. Section 3.4.6 contains the results from the  $\eta$  measurement, including a discussion of how the errors were calculated.

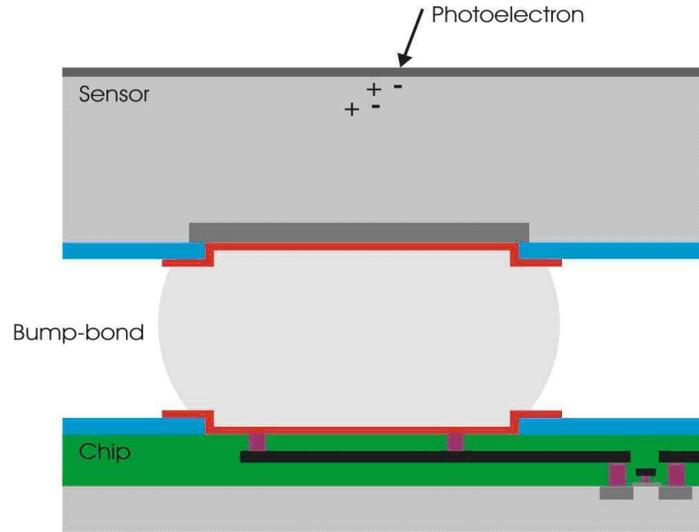
### 3.4.1 Why Measure the Photoelectron Detection Efficiency?

Given a saturated (i.e.  $\beta \approx 1$ ) charged particle traversing a RICH detector, the expected number of hits  $N_{pe}$  expected in the photodetector plane is given by [103]

$$N_{pe} = \left( \frac{\alpha}{\hbar c} \right) L \epsilon_A \eta \int \mathcal{QRT} \sin^2 \theta_c \, dE_\gamma. \quad (3.2)$$

Apart from the leading factor, which is made up of fundamental physical constants, all of the terms in Eqn. 3.2 are properties of the RICH system in question. The factors in front of the energy integral are the length  $L$  of the radiator, the fractional coverage  $\epsilon_A$  provided by the photodetectors, and the single photoelectron detection efficiency  $\eta$ . The following sections describe a direct measurement of  $\eta$  at the PDTF. It is also possible to make an indirect measurement of  $\eta$  in a testbeam setup, by measuring the average  $N_{pe}$  seen by the HPDs, and then using estimates for the other quantities in Eqn. 3.2 to deduce the value of  $\eta$ . The remaining terms are inside the integral as they are dependent on the energy of the emitted Cherenkov photon. They are the quantum efficiency  $\mathcal{Q}$  of the HPD entrance window, the reflectivity  $\mathcal{R}$  of the mirror system, the transmission  $\mathcal{T}$  of the quartz window that encloses the photodetector system, and the Cherenkov angle  $\theta_c$ .

The salient feature of Eqn. 3.2 for the current study is that  $N_{pe}$  is directly proportional to  $\eta$ , which demonstrates that accurate knowledge of  $\eta$  is required to correctly reproduce  $N_{pe}$



**Figure 3.20:** Diagram showing a Si pixel bump-bonded to a channel of the readout chip. At the top of the diagram is the sensor backplane, where the “analogue  $\langle npe \rangle$ ” is measured. At the bottom of the diagram is the readout chip input stage, where the “digital  $\langle npe \rangle$ ” is measured.

in simulations of the RICH detectors. An accurate simulation of the detector is needed in order to correctly predict the detector performance, for example the PID likelihood distributions. So measuring  $\eta$  will improve the detector simulation, allowing the actual detector performance to be confidently compared to the predictions from the simulation, so that any discrepancies can be investigated.

### 3.4.2 How to Measure the Photoelectron Detection Efficiency

To measure  $\eta$ , light is shone onto the quartz window using a pulsed LED, producing photoelectrons inside the HPD vacuum. The number of photoelectrons produced is counted at two different areas of the HPD anode. These areas can be seen in Fig. 3.20. Firstly, one counts the number of hits registered by the digital readout chip, as measured by the standard PETF Labview software. This number is the “digital  $\langle npe \rangle$ ”, where  $\langle npe \rangle$  means number of photoelectrons.

Secondly, one counts the number of photoelectrons arriving at the backplane of the silicon sensor. This is done by tapping in to the backplane directly and measuring the amount of charge deposited there. Over a large number of events a histogram of the charge can be built up that should show peaks at multiples of the charge deposited by one 20 keV photoelectron, which is around 5,000 electrons. The capacitive noise level associated with this measurement is very high, as the whole chip is being read out at once. This leads to a very large capacitance (proportional to the area being read out) of around 90 pF, compared to the

90 fF that is associated with an individual pixel. From the relative height of each peak, one can deduce the average number of photoelectrons arriving at the backplane per LED strobe. This number is the “analogue  $\langle npe \rangle$ ”. The ratio of digital to analogue  $\langle npe \rangle$  is the efficiency of the chip:

$$\eta = \frac{\langle npe \rangle_{\text{digital}}}{\langle npe \rangle_{\text{analogue}}}. \quad (3.3)$$

### 3.4.3 Factors Affecting Detection Efficiency

There are two main mechanisms by which the chip can fail to detect an incoming photoelectron, giving rise to an efficiency that is less than 100%. The first is called charge sharing. This occurs when a photoelectron hits the sensor near to the edge of a pixel, causing the electron-hole pairs to be detected by more than one pixel on the readout chip. It is possible for the electron-hole pairs to be shared between pixels in such a way that no individual pixel sees enough of a signal in the required time window to reach its threshold, so no pixel fires and the photoelectron is not detected by the chip.

The second, and predominant, mechanism is known as backscattering. When a photoelectron strikes a pixel on the sensor there is a chance that it will not deposit all of its energy there, but rather rebound from the sensor surface having deposited only part of its energy. This initial energy deposit may not be sufficient to reach the threshold and cause the pixel to fire. In this case several different scenarios are possible. If the backscattered photoelectron lands again outside of the detector chip, it will not be detected. If it lands again inside the detector chip, it may still be undetected, if its residual kinetic energy is below the discrimination threshold. The more likely outcome, given the typical threshold used in the HPD pixel chip, is that sufficient charge will be deposited at this stage to reach the threshold. However it is possible that by the time the photoelectron lands again and the charge from this second strike is collected, the readout window of the chip will have closed. In this case the photoelectron will be detected, but will be assigned to the wrong time window. This is known as a “timewalked” hit. Clearly the amount of timewalked hits depends strongly on the length of the readout window.

The probability that an electron striking a thick sensor will backscatter is a function only of the electron energy, angle of electron incidence and the atomic number of the sensor material [104]. Here, “thick” means sufficiently thick that the electron will not pass through the sensor. As the stopping range of a 20 keV electron in silicon is about  $5 \mu\text{m}$ , this condition is comfortably satisfied for the case of the HPD sensor. For HPD operation, the electron energy (20keV) and atomic number of the sensor (14 for Si) are known, and the angle of incidence is restricted as the electron must originate at the HPD quartz window and strike

the sensor. This means that the backscatter probability can be accurately estimated as 18%, with an error of less than 1% [104].

Note that both mechanisms described above depend on the length and position of the electronic “window” that defines when the chip is read out. This dependency will be explored in Sec. 3.4.6.3.

In addition to the above effects occurring within the HPD sensor, the observed photoelectron detection efficiency may also be affected by aspects of the measurement setup, such as the time distribution of the LED pulse. If this is too wide or contains a tail, this will increase the amount of hits seen in the wrong time window.

### 3.4.4 Backpulse Setup at PDTF

#### 3.4.4.1 Digital Measurement

The digital  $\langle n_{pe} \rangle$  is measured by running a strobescan (see Sec. 3.2.3.2). The relevant point on the strobescan for the  $\eta$  measurement is the point that gives the maximum number of clusters. This is the peak timing of the readout, which is what should be compared with the analogue  $\langle n_{pe} \rangle$ . Note that each point on the strobescan is measured by running 20,000 triggers.

The number of photoelectrons seen by the digital chip will of course depend on the length and position of the digital readout window. The digital readout window’s position with respect to the signal from the chip can be altered via the standard PDTF Labview software. However the length of the window is hardwired into one of the chips on the PDTF electronics boards as 50 ns, and cannot be easily changed. The bunch crossing frequency of the LHC demands that a 25 ns-long window be used during LHCb data taking. Given the constraints on the window length that have just been mentioned, one constructs a 25 ns window by shifting the start of the 50 ns window back by 25 ns with respect to the signal, which leaves only the second 25 ns of the window with signal in it. It is expected that the  $\eta$  value obtained using the 50 ns gate will be larger than with the 25 ns gate, as the longer readout gate will pick up more timewalked hits (see Sec. 3.4.3).

#### 3.4.4.2 Analogue Measurement

A schematic of the setup of the equipment used for the analogue measurement is shown in Fig. 3.21. The readout chain is operated as follows: the current pulse from the backplane is input to a low-noise charge preamplifier (model PR304 from Eurorad) that outputs a signal whose height is proportional to the amount of charge deposited at the backplane. To smooth this signal out it is passed through a commercial buffer amplifier (model EL2002CN from



allow the starting point of the analogue readout gate and its length to be changed. This is important as the correct readout of the analogue signal is very sensitive to the exact setup of the analogue readout gate (see Sec. 3.4.5.6).

### 3.4.5 Procedure for Data Taking and Fitting

#### 3.4.5.1 General Strategy

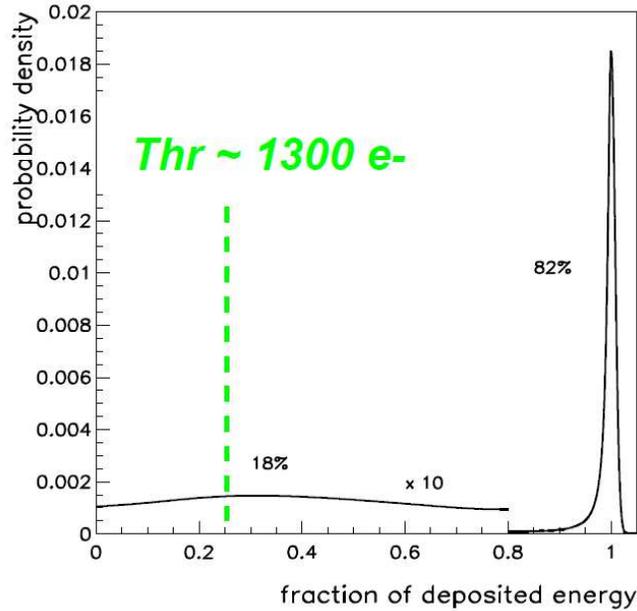
As the data taking procedure takes a finite amount of time, there will inevitably be changes in the environmental temperature. The size of the change in temperature depends on factors such as the time of day the measurement is carried out. Environmental temperature changes cause a change in the temperature of the LED, which in turn affects the average number of photons emitted by the LED each time it is triggered. To cope with this, one needs to monitor the changes in the LED output level over the time that the measurements are carried out. This is achieved by introducing a measurement procedure where each analogue measurement is sandwiched between two digital measurements. The average of the two digital  $\langle npe \rangle$  is then taken as an estimate for what the digital  $\langle npe \rangle$  would have been during the analogue measurement. From here on,  $\bar{\mu}_{\text{dig}}$  will be used to refer to the average digital  $\langle npe \rangle$  (the value used to calculate  $\eta$ ), while  $\mu_{\text{dig}}^1$  and  $\mu_{\text{dig}}^2$  will be used to refer to the digital  $\langle npe \rangle$  measured by the first and second strobescans, respectively.

Demonstrating that the  $\eta$  value obtained for a given HPD is independent of the LED intensity (i.e. independent of the digital  $\langle npe \rangle$ ) is vital in proving that a genuine efficiency is being measured. To check this, for each HPD studied measurements are performed at different levels of LED intensity. One can then check if the value of  $\eta$  changes with the LED intensity or not. However the LED intensity used should not be too high, to ensure that the probability of two or more photoelectrons striking the same pixel (thus distorting the measured digital  $\langle npe \rangle$ ) remains very low.

#### 3.4.5.2 Model for Charge Deposition Spectrum at Backplane

Before the analogue  $\langle npe \rangle$  can be measured, a model needs to be defined that describes the sensor response to the arrival of a number of 20 keV photoelectrons at the backplane. This model will be used to perform a fit to the observed charge deposition spectrum, and extract the analogue  $\langle npe \rangle$ . The starting point for this model is the Single Photoelectron Response (SPR) function. This describes the probability that a single photoelectron striking the sensor will deposit a certain fraction of its energy. The SPR used for the HPD sensor is shown in Fig. 3.22. This function was characterised previously by studying prototype HPDs that had a smaller sensor, and hence lower noise levels. In particular the spectra from these

### Single Photoelectron Response



**Figure 3.22:** The single photoelectron response (SPR) of the HPD sensor. There is an 82% probability for the photoelectron to not backscatter, and so deposit all (or almost all) of its energy into the sensor upon striking it for the first time. If the photoelectron does backscatter (18% probability), the fraction of its energy that is deposited in the initial strike follows an approximately flat distribution (shown multiplied by a factor of 10 to make it visible). The green vertical line indicates the typical amount of charge required to be deposited in one pixel to cause it to fire. Reproduced from [106].

prototype HPDs have low enough noise levels such that a fit can be performed to measure the backscatter probability. The value found for the backscatter probability was around 18%, which is in agreement with the accepted value described in the previous section.

The energy deposition spectrum for the arrival of two photoelectrons within the same readout window is just the convolution of the SPR with itself. This is because the arrival of photoelectrons is independent. Another consequence of the independence of the photoelectrons is that the probability that a given number will arrive in the same time window follows a Poisson distribution. This is a key factor that allows the backpulse measurement to be carried out, as the relative height of each photoelectron peak will follow a Poisson distribution, whose mean  $\mu$  is the analogue  $\langle npe \rangle$ .

So in the absence of noise in the readout chain, the charge spectrum would consist of a Poisson-weighted sum of the deposition spectrum for  $k$  electrons, where  $k = \{1, 2, \dots, \infty\}$  and the pedestal, which at this stage is just represented by a delta function at zero. Let this idealised charge spectrum be  $f(x)$ , where  $x$  represents the total charge deposited, in units of

the charge deposited by a single non-backscattered electron (which is  $\approx 5,000$  electrons). To account for the fact that there is electronic noise present both at the sensor and in the readout chain, one convolutes  $f(x)$  with some global noise function  $g(x)$  to obtain  $h(x)$ , say. This  $h(x)$  is then a realistic model for the charge spectrum. In practice,  $g(x)$  takes the form of a Gaussian. Applying a Fourier Transform to  $h(x)$  gives (using the convention of capitalising transformed functions):

$$H(k) = F(k) \cdot G(k). \quad (3.4)$$

Under the assumption that the charge spectrum follows a Poisson distribution, it has been shown that a Light Sum Rule exists for the Fourier Transform of the charge spectrum [107].

The Light Sum Rule in this case gives

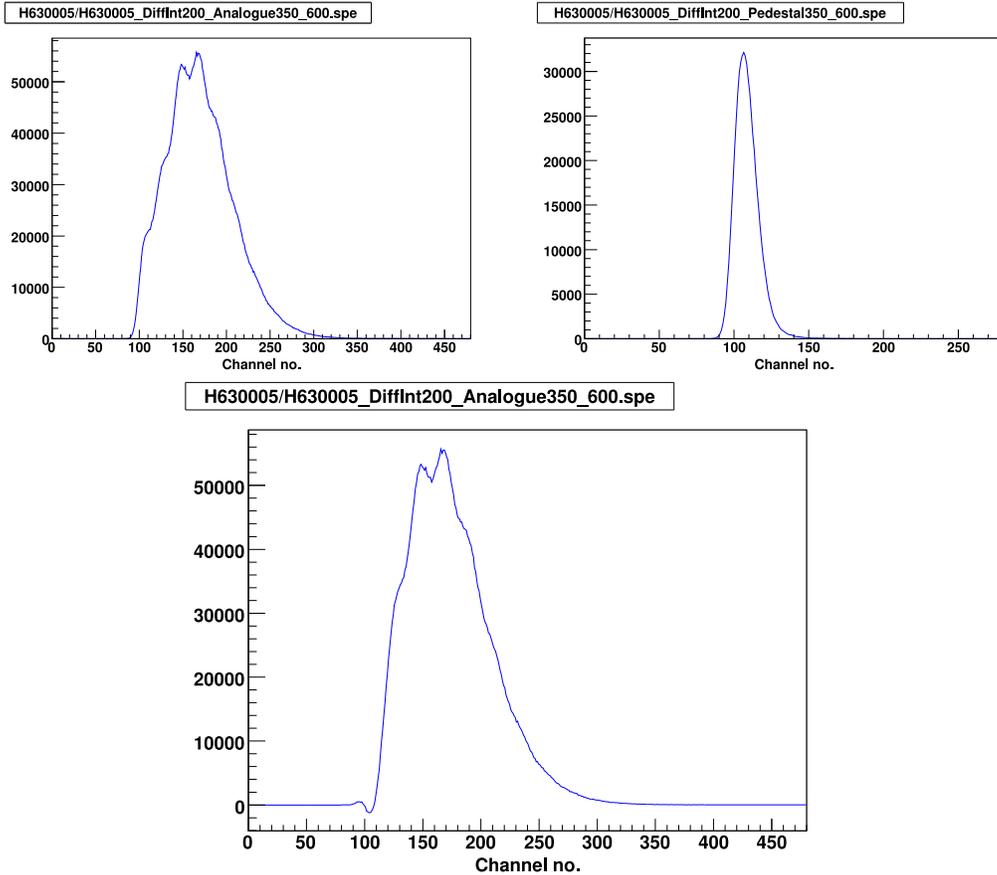
$$F(k) = \exp(\mu S(k)), \quad (3.5)$$

where  $\mu$  is the Poisson mean, and  $S(k)$  is the Fourier Transform of the SPR. The exponential arises as a power series formed by the addition in Fourier space of terms corresponding to each possible number of photoelectrons. The consequence of the Light Sum Rule is that one obtains a relatively simple expression for  $H(k)$ , and then one just applies the inverse Fourier Transform to obtain another relatively simple expression for  $h(x)$ , which now has an explicit dependence on  $\mu$ , the variable of interest here. Hence performing the fit in Fourier space is faster and more robust than a brute force method of building the energy deposition spectrum from the SPR and the noise function without making use of the Light Sum Rule.

### 3.4.5.3 Pedestal Subtraction

During data taking it was observed that the height of the pedestal in each spectrum was always too large, i.e. the pedestal did not fit the Poisson distribution followed by the photoelectron peaks. The reason for this is not known, although it could be that, due to its peak-sensing nature, the ADC is picking up some dark count events. To deal with the excessive pedestal size, it was decided to remove the pedestal from each spectrum and fit to the modified spectrum using a model that also has the pedestal removed. This way the pattern of the photoelectron peaks is preserved, and these can be used to deduce the Poisson mean of the distribution.

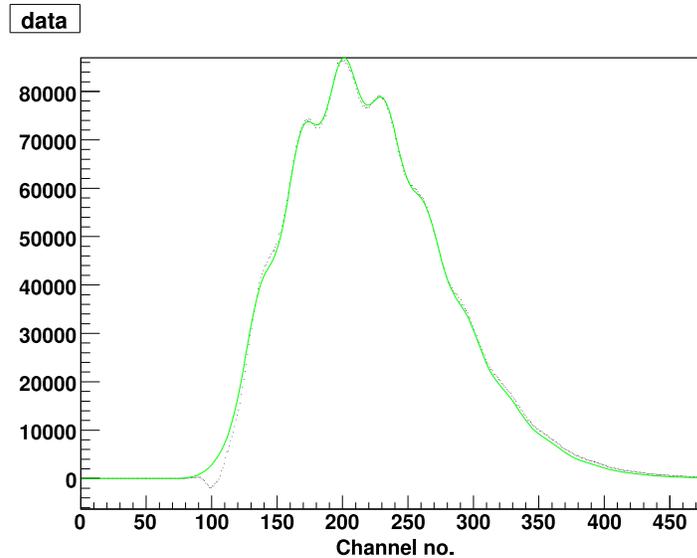
To subtract the pedestal from a data spectrum, first a pedestal spectrum is taken to accompany each analogue run. This is done simply by taking an analogue run with the LED disabled. Then the pedestal spectrum is scaled so that it contains approximately the same number of events that comprise the pedestal contribution to the spectrum taken for the full analogue run. To carry out this scaling, the peak channel of the pedestal spectrum is first found, and the total number of counts in this channel and the eight surrounding channels



**Figure 3.23:** Analogue data taken from the sensor backplane. The blue line in each figure is formed from the histogram of the number of counts in each channel by linear interpolation. Top left: a “raw” spectrum including the pedestal. Top right: the corresponding pedestal-only run. It can be seen that the shape of the pedestal is asymmetric, which is a further indication that dark count events are being picked up. Bottom: the spectrum formed by scaling the pedestal and then subtracting it from the raw spectrum.

(four channels each side) is calculated. The choice of taking four channels either side of the peak strikes a balance between the statistical fluctuations from using too few channels and dependence on the shape of the pedestal from using too many channels. The total number of counts in the corresponding channels of the full analogue spectrum is then calculated, and the scaling factor is found by taking the ratio of these two numbers. This method relies on the fact that the mapping from channels to the peaks in the spectrum does not (for a given HPD) change from run to run.

Finally, the scaled pedestal is subtracted from the full analogue spectrum to give a spectrum consisting of only the photoelectron peaks. This process is illustrated in Fig. 3.23. A modified fitting model is needed to cope with the pedestal-subtracted data. To subtract the pedestal from the fitting model,  $F(k) - 1$  is substituted for  $F(k)$  when constructing the fitting model (see Sec. 3.4.5.2).



**Figure 3.24:** Typical fit to a backpulse spectrum with the pedestal subtracted. The data points are in black, and the fitted PDF is shown by the green line.

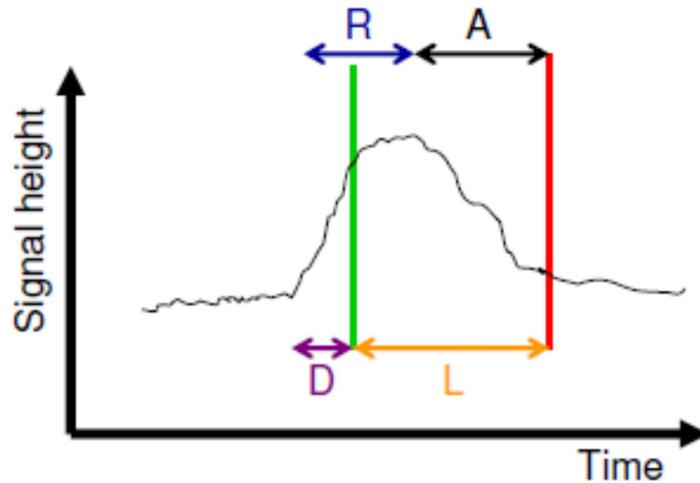
#### 3.4.5.4 Fit Region

The fit region is defined as the channels for which the  $\chi^2$  between the data and the fit PDF is calculated during the fit. Fits including the pedestal region were often found to not converge correctly. This is because the distribution in the pedestal region is very sensitive to the pedestal scaling factor, the calculation of which is not an exact process. In particular negative values can be obtained after subtraction. Also, the fit region does not include the higher channels in the spectrum as the number of signal counts there is very low. Hence background noise in the readout chain can significantly distort the shape of the spectrum there, affecting in turn the fit to the lower photoelectron peaks. By contrast a small number of background counts at the lower photoelectron peaks will not affect the fit significantly, since most of the photoelectron data can be found in this region. For these reasons, the fit to the pedestal-subtracted spectrum is performed starting at the first or second photoelectron peak, and finishing at around the sixth or seventh photoelectron peak<sup>5</sup>.

#### 3.4.5.5 Execution of Fit

A fit is then performed using the modified model,  $h(x)$ , for the charge spectrum (see Sec. 3.4.5.2). The minimisation is performed by the “Simplex” fitter from the GSL (GNU Scientific Library) package [108]. An example fit is shown in Fig. 3.24. There are four

<sup>5</sup>However the fitting function is still drawn covering the whole spectrum to demonstrate that it is still in broad agreement with the spectrum even in the areas where the fit is not being done.

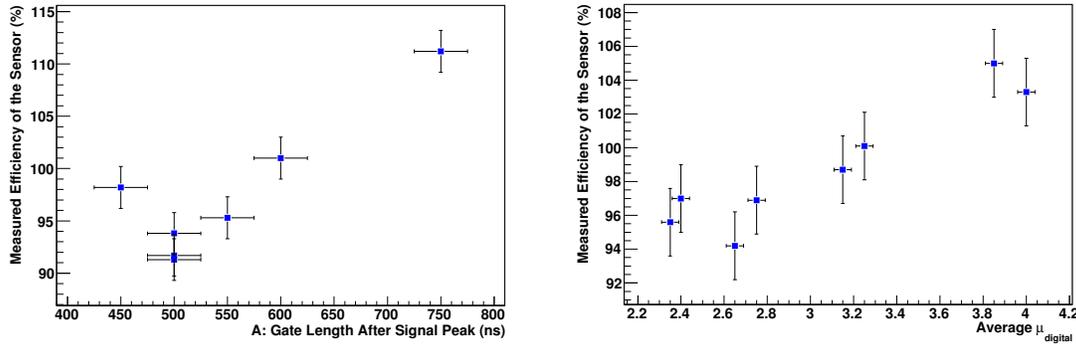


**Figure 3.25:** Sketch of the backplane signal and analogue gate setup. The green (red) vertical line represents the start (end) of the analogue gate. The signal rise time ( $R$ ) is independent of the gate setup. The sum of this and the length of time ( $A$ ) the gate extends after the peak is equal to the sum of the gate length ( $L$ ) and the delay ( $D$ ) between the start of the signal and the start of the gate.

parameters left floating in the fit: the location of the pedestal, the distance between photoelectron peaks, the width of the global noise Gaussian and the Poisson mean. As the location of the pedestal can be deduced within a few channels by inspecting the data, the fit is seeded with a value for the pedestal location that is very close to the full value. This prevents the fit from attempting to treat another photoelectron peak as the pedestal. The parameters of the SPR are fixed according to the results found from the fits to spectra from the prototype HPDs.

### 3.4.5.6 Dependency of Efficiency on Analogue Gate

Data were taken with a number of different setups for the analogue gate. A sketch of the backplane signal (as it appears on the oscilloscope, after passing through the amplifiers) and the analogue gate setup is shown in Fig. 3.25. The peak height of the signal outputted by the fast filter amplifier (see Sec. 3.4.4.2) is proportional to the total amount of charge deposited at the backplane. Hence the ADC is used in peak-sensing mode, and in principle the gate only has to include the signal peak in order to correctly sample the signal. However it was found that a strong dependence existed between the measured value for the efficiency and the length of time the analogue gate extended after the peak of the signal. The dependency of the measured  $\eta$  value on the gate setup for a fixed light input level (digital  $\langle n_{pe} \rangle \simeq 3.3$ ) is shown on the left of Fig. 3.26. Two separate effects can be seen.



**Figure 3.26:** Left: Dependence of the measured sensor efficiency ( $\eta$ ) on analogue gate setup. The x-axis is the length “A” shown in Fig. 3.25. Right: Dependence of the measured sensor efficiency on the light input level for an analogue gate with  $A = 600$  ns.

Firstly, when the gate extends less than 500 ns after the peak of the signal, the measured value of  $\eta$  rises. The size of the pedestal in the charge spectrum collected with such a gate is seen to increase sharply, resulting in domination of the spectrum by the pedestal. This disrupts the fit to the spectrum, causing it to return lower values for the analogue  $\langle npe \rangle$ . This translates into a higher measured value for the sensor efficiency. The reason for this effect is understood: the documentation of the ADC states that the sampling gate must extend for at least 500 ns after the peak of the signal to allow the height of the peak to be correctly sampled.

Secondly, when the analogue gate extends more than 500 ns beyond the peak of the signal, the measured value of  $\eta$  starts to rise. The increase in measured  $\eta$  values comes from an decrease in the measured analogue  $\langle npe \rangle$ , with the digital  $\langle npe \rangle$  (which is independent of the analogue gate setup) remaining constant. The question then arises as to which gate setup is giving the correct value for  $\eta$ . This question can be resolved by considering the dependence of the measured  $\eta$  values on the light input level (i.e. on the average digital  $\langle npe \rangle$ ) when the gate setup is fixed. For the gates extending significantly more than 500 ns after the signal peak,  $\eta$  is seen to depend on the digital  $\langle npe \rangle$ . An example of this is shown on the right of Fig. 3.26, showing data taken with an analogue gate that extends for 600 ns after the signal peak. The measured  $\eta$  values increase as the average digital  $\langle npe \rangle$  increases. This is a clear indication that these gate setups are not optimal. When the gate extends roughly 500 ns after the signal peak, however, the measured  $\eta$  values no longer depend on the digital  $\langle npe \rangle$  (see Fig. 3.28), which indicates that the efficiency is being correctly measured. Hence this was chosen as the optimal gate setup.

The reason for the drop in analogue  $\langle npe \rangle$  as the gate length increases is not fully understood. One possible cause is that the analogue signal is suffering some ringing after the main

signal peak, but that this ringing is not large enough to show itself above the noise on the oscilloscope. The longer gates include this ringing region of the signal, but the shorter ones do not. Another possible cause is an additional contribution from dark counts.

### 3.4.6 Results

Using the optimised analogue gate setup as described in Sec. 3.4.5.6, measurements were carried out on HPD H630005 using 25 ns and 50 ns digital gate lengths, and on HPD 708016 using 50 ns digital gate length. These particular HPDs were selected for this measurement as they exhibited excellent performance in the general PDTF tests. In particular, they have

- Very low leakage current ( $\mathcal{O}(10 \text{ nA})$ ) from the detector chip, which minimises the shot noise,
- Very low dark-count rate ( $\mathcal{O}(1 \text{ kHz/cm}^2)$ ), which reduces the number of background hits on the detector chip,
- No dead or noisy pixels, to eliminate the need to take these into account when calculating the number of photoelectrons.

Since the anode properties should vary little between HPDs, it is expected that no significant difference in  $\eta$  values will be seen between the two HPDs, and also that the  $\eta$  values obtained will be representative of the entire sample of HPDs.

#### 3.4.6.1 Errors on Digital $\langle npe \rangle$

The statistical error on each of  $\mu_{\text{dig}}^1$  and  $\mu_{\text{dig}}^2$  (the digital measurements before and after the analogue measurement) is simply the uncertainty on the Poisson mean returned by the PDTF software. This can be estimated by calculating the increase in the  $\chi^2$  of the Poisson fit when a different Poisson mean is assumed, and choosing the variation in the mean that causes the  $\chi^2$  to increase by 1. This variation is found to range from 0.011 for low values of  $\mu_{\text{dig}}^{1,2}$  ( $\simeq 2$ ) to 0.017 for high values ( $\simeq 6$ ). The statistical error on  $\bar{\mu}_{\text{dig}}$  (the average digital  $\langle npe \rangle$ ) is found by combining the errors on  $\mu_{\text{dig}}^{1,2}$  in quadrature, then halving the result. The typical (absolute) statistical error is then around 0.01 or less. As will be seen below, this is small compared to the systematic error, and will be neglected.

The main systematic error on  $\bar{\mu}_{\text{dig}}$  comes from possible temperature-dependent changes in the LED output rate over time. Even when the difference between  $\mu_{\text{dig}}^1$  and  $\mu_{\text{dig}}^2$  is small, the output rate may for example have risen between the first strobescan and the analogue run, then fallen again between the analogue run and the second strobescan. To take this

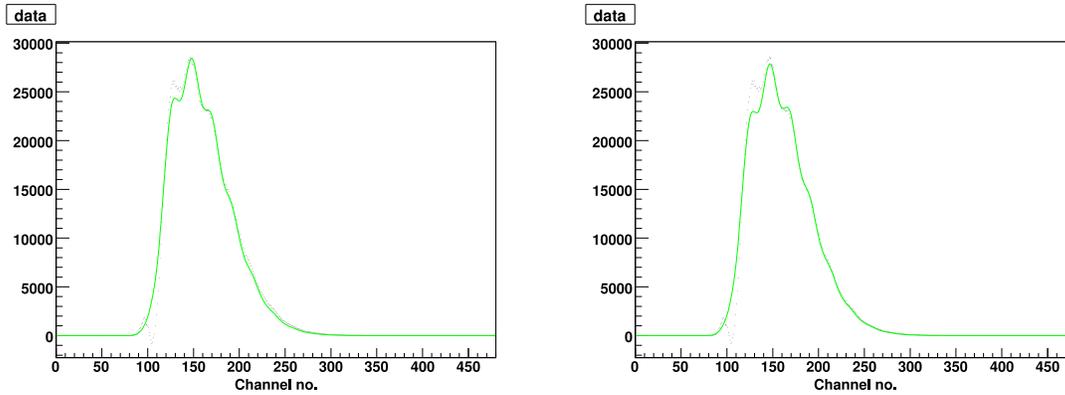
uncertainty into account, a systematic error is assigned to  $\bar{\mu}_{\text{dig}}$  as follows: the difference between  $\mu_{\text{dig}}^1$  and  $\mu_{\text{dig}}^2$  is calculated and halved. If the result is larger than 1% of  $\bar{\mu}_{\text{dig}}$ , it is taken as the systematic error. Otherwise the systematic error is set to 1%. The choice of 1% as the minimum systematic uncertainty on  $\bar{\mu}_{\text{dig}}$  is based on the typical variations in digital  $\langle npe \rangle$  that are seen during normal HPD testing, for example between consecutive long LED runs (see Sec. 3.2.3.2). Such variations can only be due to changes in the LED output rate.

This assignment leads to a systematic error of 1% for cases where  $\mu_{\text{dig}}^1$  and  $\mu_{\text{dig}}^2$  have very similar values, and of up to  $\simeq 2\%$  for cases where there was significant change between  $\mu_{\text{dig}}^1$  and  $\mu_{\text{dig}}^2$ . So the systematic error on  $\bar{\mu}_{\text{dig}}$  is always at least several times larger than the statistical error, which is therefore neglected.

### 3.4.6.2 Errors on Analogue $\langle npe \rangle$

The statistical error on the analogue  $\langle npe \rangle$  is the uncertainty on the  $\mu$  parameter as returned by the fit to the backpulse spectrum. This can be estimated by the same method used for the digital  $\langle npe \rangle$ , by measuring how much each parameter must change to increase the total  $\chi^2$  of the fit by 1. For the  $\mu$  parameter it is found that the change needed is of the order of  $10^{-3}$ . This value is so small because the statistics present in a given spectrum are huge, with each bin of the early photoelectron peaks (see Fig. 3.24) holding of order  $10^5$  events. Thus the fractional error on the bin height is well below the percent level. The error returned by the fit is correspondingly very small. It will be seen below that this statistical error is completely negligible when compared to the systematic error. Hence the statistical error will not be considered when calculating the overall error on the analogue  $\langle npe \rangle$ .

A significant systematic error on the analogue  $\langle npe \rangle$  arises because the model used to fit the backplane spectrum to extract  $\mu$  does not accurately describe the data over the whole distribution. The size of the error introduced by this can be estimated in two different ways. Firstly, the change in  $\mu$  caused by varying the region of the spectrum that is fitted to (i.e. the region where the  $\chi^2$  is calculated) can be studied. It is found that the largest change in  $\mu$  is induced by not including the region of the first photoelectron peak in the fit. For most spectra this causes a change in  $\mu$  of about 1%. But for spectra at low analogue  $\langle npe \rangle$ , the difference is more significant, at about 3%. A change of about 3% was also seen for the fits with the highest analogue  $\langle npe \rangle$  values. The effect of changing the fit region on the fit result for a spectrum at low analogue  $\langle npe \rangle$  is illustrated in Fig. 3.27. The fit on the left hand side is made including the first photoelectron peak in the fit region, while the fit region for the fit on the right hand side excludes the first photoelectron peak. It can be seen that the fit excluding the first photoelectron peak follows the data at the higher photoelectron peaks better, at the expense of a poorer fit at the first and second photoelectron peaks.



**Figure 3.27:** Effect of varying the fit region for a spectrum at low analogue  $\langle n_{pe} \rangle$ . The fit region for the left hand fit includes the first photoelectron peak, while the fit region for the right hand fit does not. The value of  $\mu$  returned by the fit is 2.75 for the left hand fit and 2.86 for the right hand fit.

The second way of examining the effect of the imperfect fit on  $\mu$  is to rerun the fit with one or more of the parameters that are normally floating being fixed, and examining how the fitted value for  $\mu$  changes as the value chosen for the fixed parameter changes. The parameters representing the separation between photoelectron peaks and the location of the pedestal are varied within the ranges found when varying the fit region. This range is typically a few channels in size. These variations are found to cause a typical change in  $\mu$  of around 1%. The size of this effect, unlike the effect of altering the fit region, does not significantly change from spectrum to spectrum.

The effect of choosing a different value for the backscattering probability in the SPR can also be considered. In light of the remarks about the backscattering probability in Sec. 3.4.3, 0.5% is chosen as an estimate of the uncertainty on the backscattering probability. Fitting with the probability fixed to be 0.5% either side of the 18% central value changes the fitted value of  $\mu$  by about 0.4%. The size of this effect is independent of the spectrum studied. An error of this size can be neglected when compared with the other systematic uncertainties described above.

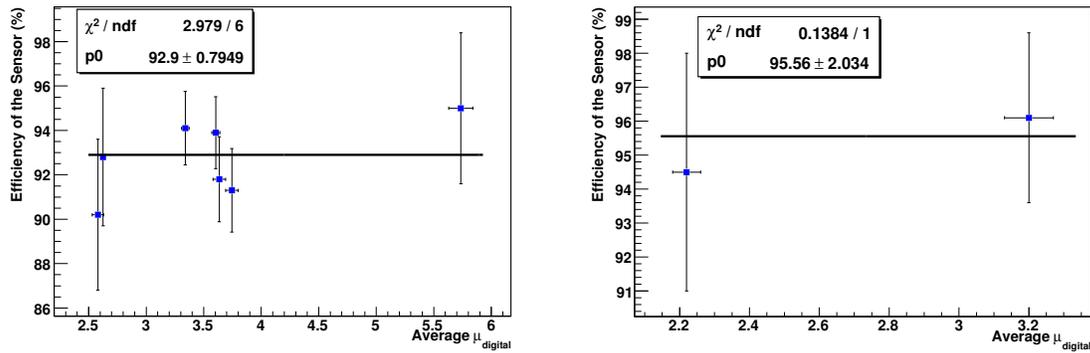
As the statistical error on  $\mu$  has also been seen to be negligible, the total error on  $\mu$  is found by adding the two dominant systematic uncertainties, from variation of the fit region and of the other parameters in the fit, in quadrature.

### 3.4.6.3 Measured Values of the Detection Efficiency

The errors on digital  $\langle n_{pe} \rangle$  and analogue  $\langle n_{pe} \rangle$  are independent of each other. Hence the error on a single  $\eta$  value can be found by making use of Eqn. 3.3 to combine the digital and

	Relative Error on Digital $\langle npe \rangle$	Relative Error on Analogue $\langle npe \rangle$	Relative Error on $\eta$
Statistical	$< 0.4\%$	$\mathcal{O}(0.1\%)$	$< 0.4\%$
Systematic	Between 1% and 2%, depending on $ \mu_{\text{dig}}^1 - \mu_{\text{dig}}^2 $	Between 1.4% and 3.2%, depending on the fit behaviour at first photoelectron peak	Between 1.7% and 3.7%
Combined	Between 1% and 2%	Between 1.4% and 3.2%	Between 1.7% and 3.7%

**Table 3.3:** Contributions to the total error on the sensor efficiency.



**Figure 3.28:** Results for  $\eta$  using digital gate length of 50 ns. The left hand plot shows results from HPD H630005, and the right hand plot shows results from HPD H708106.

analogue errors:

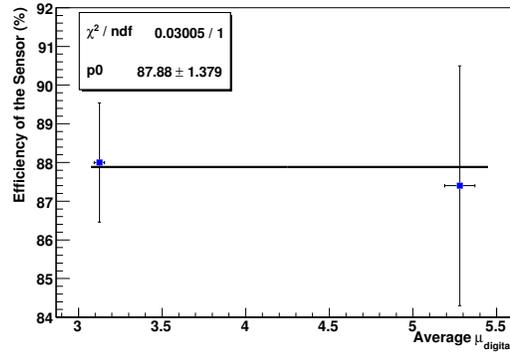
$$\left(\frac{\sigma_\eta}{\eta}\right)^2 = \left(\frac{\sigma_{\langle npe \rangle_{\text{digital}}}}{\langle npe \rangle_{\text{digital}}}\right)^2 + \left(\frac{\sigma_{\langle npe \rangle_{\text{analogue}}}}{\langle npe \rangle_{\text{analogue}}}\right)^2. \quad (3.6)$$

The contributions to the total error on  $\eta$  are summarised in Table 3.3. The results for 50 ns digital gate length are shown in Fig. 3.28, and those for 25 ns digital gate length in Fig. 3.29.

In each plot in Figs. 3.28 and 3.29, a constant fit has been made to give an average value for  $\eta$  from each plot. Performing a straight line fit shows that the gradient of  $\eta$  with respect to  $\bar{\mu}_{\text{dig}}$  is compatible with zero for all three plots, demonstrating that  $\eta$  is independent of  $\bar{\mu}_{\text{dig}}$  as desired.

Combining the results using digital gate length of 50 ns from the two HPDs allows overall results for both gate lengths to be stated. The errors are taken from the constant fits to each set of results. In Table 3.4 the overall results are given. The LHCb-RICH specification, which is governed by photoelectron backscattering effects and the expected performance of the pixel readout chip, is also given.

As expected, the detection efficiency when using a 25 ns gate is significantly lower than



**Figure 3.29:** Results for  $\eta$  using digital gate length of 25 ns. Results are from HPD H630005.

Digital Readout Window	Measured Value of $\eta$	LHCb-RICH Specification
50 ns	$(93.3 \pm 0.7)\%$	n/a
25 ns	$(87.9 \pm 1.4)\%$	85% typical

**Table 3.4:** Single photoelectron detection efficiency  $\eta$  of the manufactured HPDs, as measured at PDTF.

when using a 50 ns gate. It can also be seen that the production HPDs exceed the LHCb-RICH requirement that the single photoelectron detection efficiency with a 25 ns readout window should have a typical value of 85%. The results are in agreement with  $\eta$  values that were measured using preseries HPDs [109], and with indirect measurements from testbeams using preseries [110] and production [111] HPDs.

## 3.5 PDTF Test Results

In this section the distributions of the key parameters that were measured for all HPDs at the PDTFs will be presented, and the implications for the operation and performance of the HPDs within the RICHes will be discussed. Note that the following results cover testing of 557 HPDs. This is higher than the 550 required in the contract as it includes 7 HPDs that failed and were replaced with new HPDs by DEP.

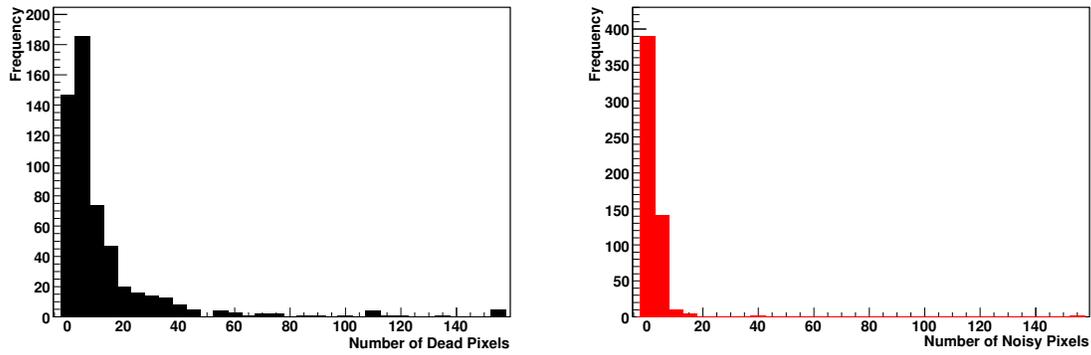
### 3.5.1 Performance

A summary of the performance of the HPD ensemble is given in Table 3.5. The distributions of key parameters are shown in Figs. 3.30 to 3.34<sup>6</sup>.

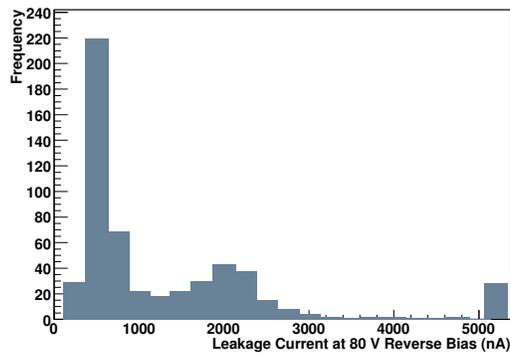
<sup>6</sup>Note that in each histogram in Figs. 3.30 to 3.34, the rightmost bin is an overflow bin that holds any HPDs lying above the range used for that histogram.

Property	Mean Value	RMS
Working Pixels	99.8%	0.4%
Leakage Current at 80 V Reverse Bias	1.49 $\mu\text{A}$	2.65 $\mu\text{A}$
Photoelectron Detection Threshold	1064e <sup>-</sup>	98e <sup>-</sup>
Photoelectron Detection Noise	145e <sup>-</sup>	11e <sup>-</sup>
Dark-Count Rate	2.49 kHz/cm <sup>2</sup>	5.82 kHz/cm <sup>2</sup>
Ion Feedback Probability	0.04%	0.15%
Photocathode Image Radius	6.61 mm	0.13 mm
Photocathode Image Displacement	0.39 mm	0.21 mm

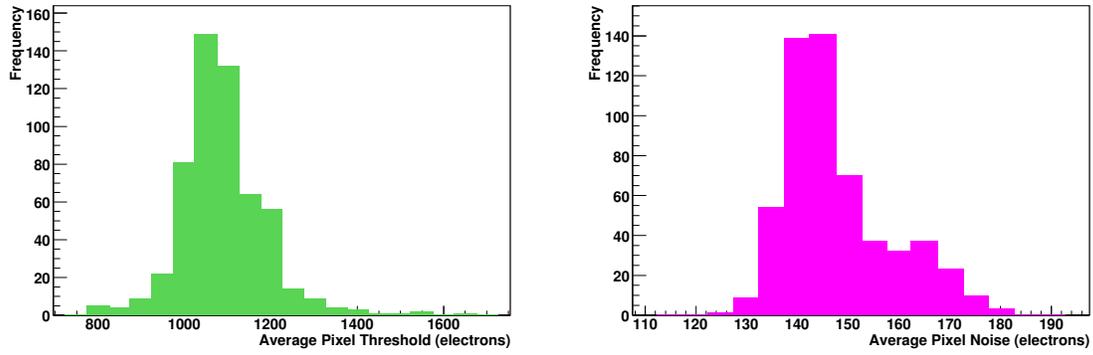
**Table 3.5:** Performance of the manufactured HPDs. Values given are the raw mean and RMS of the distribution.



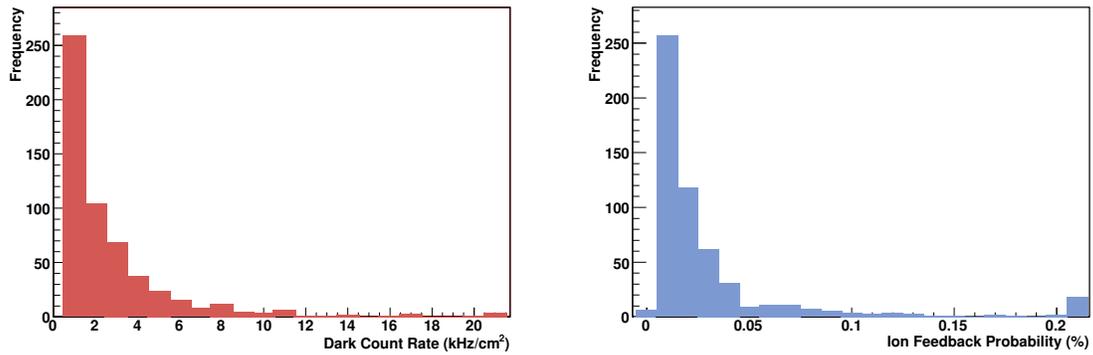
**Figure 3.30:** Distributions of the number of dead (left) and noisy (right) ALICE pixels (out of 8192).



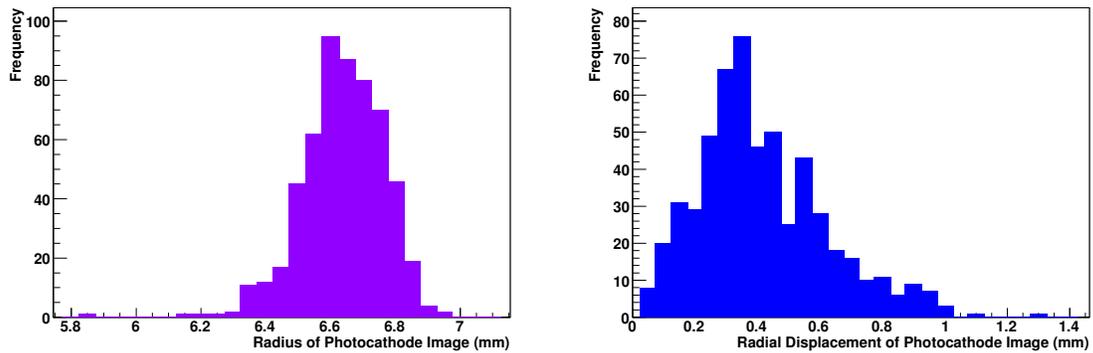
**Figure 3.31:** Distribution of leakage current at 80 V reverse bias.



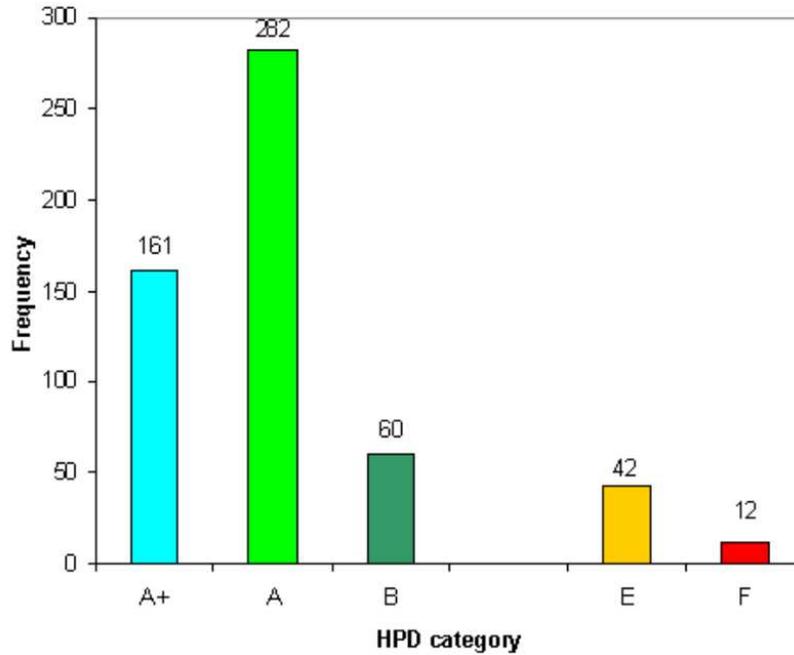
**Figure 3.32:** Distributions of average pixel thresholds (left) and noise (right). Values are calculated by averaging over the values from each individual pixel. The typical spread of these values in a single HPD is about 90 electrons for the pixel threshold, and about 20 electrons for the pixel noise.



**Figure 3.33:** Distribution of dark-count rate from tests over 5 million triggers (left) and ion feedback probability from strobescan (right).



**Figure 3.34:** Distributions of radius of the photocathode image (left) and radial displacement of the image with respect to the centre of the pixel chip (right).



**Figure 3.35:** Final classification of HPDs according to the tests performed at the PDTFs.

### 3.5.2 Discussion

After testing (including any retests to resolve issues found in the first test run), each HPD is classified as belonging to one of 5 categories: A<sup>+</sup>, A, B, E or F. The criteria for each category are detailed in Table 3.6. HPDs classified as A<sup>+</sup> or A are preferentially placed in the higher-occupancy regions of the RICH system (see Fig. 3.4). HPDs classified as B or E are placed in lower-occupancy regions, or kept as spares in case any HPDs fail after installation in the RICHes. HPDs classified as F are judged unusable, and are not placed in the RICH at all.

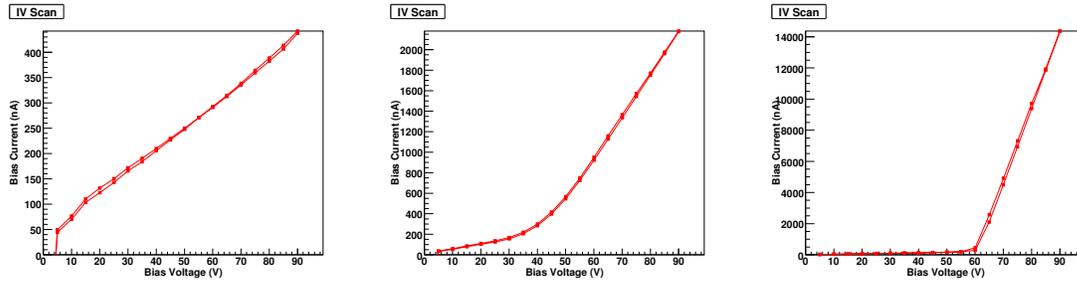
The distribution of HPDs across the categories is shown in Fig. 3.35. It can be seen that only 12 (or 2.2%) of HPDs fail the requirements in such a way as to be judged unusable in the RICH system. Some of these cases were due to the HPD performing poorly in a particular test and are described below. In other cases the HPD suffered a complete operational failure, for example failure of the pixel chip during the initial high voltage ramp-up. The justification of the HPD categories in light of the results shown in Sec. 3.5.1 will now be discussed.

Almost all HPDs have far fewer bad pixels than the maximum of 410 (5%) specified in the contract. This indicates that the quality and reliability of the bump-bonding process is excellent. A handful of HPDs were classified as E due to a high number of dead pixels.

In the leakage current distribution three distinct classes of behaviour can be distinguished. Examples of each of these are shown in Fig 3.36. The leakage current of the first class of HPD shows a linear dependence with bias voltage all the way up to 90 V. No breakdown

**Table 3.6:** Classification criteria for HPDs. To achieve categorisation as A<sup>+</sup>, A, B or E, an HPD must meet all of the listed criteria. Any HPD with one or more of the criteria given for the F category is categorised as F.

Category	Meaning	Specified Performance					
		IFB Rate	QE (270 nm)	Bad Channels	Dark-Count Rate	Leakage Current at 80 V	Photocathode Offset
A <sup>+</sup>	Meets all specifications and clearly exceeds certain specifications.	<1%	>30%	<20	Uniform with pixel rate $<2 \times 10^{-5}$	<3 $\mu$ A	<0.5 mm
A	Meets all specifications.	<1%	$\geq 20\%$	<80	<5 kHz/cm <sup>2</sup>	<5 $\mu$ A	<1 mm
B	Fails to meet one or more of the specifications but is still fully usable.	<1%	$\geq 20\%$	<200	<100 kHz/cm <sup>2</sup>	<5 $\mu$ A	<1 mm
E	Flagged with a particular issue that requires special attention, but is still fully usable.	<1%	$\geq 20\%$	<400	<100 kHz/cm <sup>2</sup>	<20 $\mu$ A	<2 mm
F	Fails to meet one or more of the specifications in such a way that it is unusable in the RICHes.	$\geq 1\%$	<20%	$\geq 400$	$\geq 100$ kHz/cm <sup>2</sup>	$\geq 20$ $\mu$ A	$\geq 2$ mm



**Figure 3.36:** Three different classes of leakage current curve for HPDs. Left, an HPD that shows no breakdown effect up to 90 V reverse bias, leading to low leakage current. Centre, an HPD that sees an increase in gradient beyond 40 V, leading to a moderate leakage current. Right, an HPD that displays a breakdown at around 50 V reverse bias, leading to very high leakage current at the operating point of 80 V reverse bias.

effects are observed, and only a very small residual current (typically below  $0.8 \mu\text{A}$ ) is observed. The second class of behaviour is similar to the first class at low voltages, but shows a clear increase in gradient above a certain bias voltage, typically around 40 or 50 V. The resulting leakage current at 80 V is usually between 1 and  $3 \mu\text{A}$ . This behaviour is believed to be caused by defects in the manufacturing process of the sensor, and often groups of HPDs consisting of sensors fabricated from the same wafer feature similar leakage current behaviour.

Examples of possible defects include scratches in the highly doped n+ layer at the sensor backplane, which may allow the intrinsic n bulk of the sensor to come into contact with the aluminum guard ring (this contact can also be caused if “spikes” are present in the guard ring causing the n+ layer to be pierced). A current sink then forms at the backplane once the bias voltage reaches a certain value. Another possible defect is the presence of fissures or cracks in the side edge of the sensor. If these extend into the region of non-zero electric field in the main part of the sensor, current can flow around the sensor edge. If the local electric fields generated by such defects is high enough, the electrons become sufficiently energetic to ionise silicon atoms, leading to an avalanche effect. A sudden breakdown is seen above a certain bias voltage, and the leakage current runs away to very high values (usually  $>5 \mu\text{A}$ ). This is the third class of behaviour. HPDs with this behaviour form the bulk of those classified as E.

However, for all except one HPD, the leakage current is low enough that the sensor still reaches full depletion at 80 V reverse bias, and so is still usable. The one HPD where full depletion was not attained (with leakage current of  $24 \mu\text{A}$ ) was judged unusable and classified as F. HPDs with similar leakage currents are placed on the same Level 0 board within the RICHes, as the leakage current affects the readout timing settings (see Sec. 3.2.3.1), and

these are set on each Level 0 board and not individually for each HPD.

Although three HPDs marginally exceed the contractual maximum pixel threshold of  $1500 e^-$ , this does not affect their classification, as they have normal noise values, so that the signal-to-noise ratio is still high enough that any effect on the performance of these HPDs will be minimal. No HPD exceeds the contractual typical noise of  $145 e^-$  by large enough margin to adversely affect its performance.

Although there are 63 HPDs (11%) that show dark-count rates above the contract specification of  $5 \text{ kHz/cm}^2$ , this will not have any significant effect on their performance in the RICHes. This is because the pattern recognition algorithm is not seriously affected by dark-count rates below  $5 \text{ MHz/cm}^2$ , i.e. one thousand times higher than the contract specification. However lower dark-count rates are still preferable, hence any HPDs with dark-count rates above  $5 \text{ kHz/cm}^2$  were classified as B. This amounted to 60 HPDs. Also, three HPDs exhibited unstable dark-count rates peaking between 50 and  $100 \text{ kHz/cm}^2$ . These HPDs were classified as F and rejected.

The vast majority of HPDs (521, or 94%) have ion feedback probabilities that are at least an order of magnitude lower than the contract specification of 1%. One HPD showed very high IFB probability ( $>10\%$ ) and was rejected.

Despite the excellent ion feedback probabilities that were measured by the PDTFs, many HPDs have been found to show considerably higher IFB rates following their installation inside the RICHes. Indeed the IFB rate of some HPDs has risen above the level that triggers an avalanche effect (see Sec 3.2.3.2), causing these HPDs to become unusable. These HPDs have been replaced in the RICHes with spares.

Studies of the time dependence of the IFB rate have shown an approximately linear increase of the IFB rate with time. This is the reason that the IFB rates measured at the PDTFs were not high enough to cause any concern. The cause of the vacuum degradation that is driving the increase in IFB rate is not yet known, however preliminary studies suggest that hydrogen is present inside the HPD. Due to these unresolved vacuum quality issues, the HPDs are not being considered as the first-choice RICH photodetector technology for the planned upgrade of LHCb.

Only 2 HPDs show a photocathode image that is displaced by more than 2 LHCb pixels (i.e. 1 mm). These HPDs were classed as E, because such a displacement at PDTF may translate to a larger displacement within the RICHes, due to the residual magnetic field (see discussion in Sec. 3.2.3.2). Hence there is a risk that part of the photocathode image may fall outside of the pixel chip. One HPD showed a smaller than normal photocathode image of radius 5.8 mm and was classified as E.

To summarise, the overall quality of the production HPDs has been seen to be excellent, with the vast majority of them performing well enough to be used in the RICHes without any concerns. The contractual specifications and LHCb requirements are met and often exceeded in key areas.

## 3.6 Conclusions

The LHCb RICH system uses arrays of Hybrid Photon Detectors (HPDs) to detect the Cherenkov photons produced by charged particles traversing the RICH. All 557 HPDs produced for LHCb have undergone thorough quality assurance testing and characterisation at dedicated test facilities. Their performance is excellent, with only 2.2% of HPDs found to be unusable for the RICHes. The LHCb requirements and the contractual specifications are met and often exceeded in key areas, for example the dead pixel number, ion feedback probability and dark-count rate. The quantum efficiency of the photocathode is found to be significantly higher than expectations, leading to increased photoelectron yield in the RICHes.

The single photoelectron detection efficiency  $\eta$  of the HPD anode has been measured using two different digital readout windows. The measured values are  $\eta_{50\text{ns}} = (93.3 \pm 0.7)\%$  and  $\eta_{25\text{ns}} = (87.9 \pm 1.4)\%$  respectively. These values are independent of the light input level. They confirm that the production HPDs exceed the LHCb-RICH requirement in this area, and are in agreement with previous measurements of  $\eta$ . The value for the 25 ns gate will be used to optimise the simulation of the RICH detectors.

The RICH detectors are now fully populated with the HPDs and are ready to participate in the study of the first collisions within LHCb.

# Chapter 4

## Measuring the Proper Time Resolution for Two-Body Hadronic $B$ Decays from Data

An introduction to the physics of two-body charmless hadronic  $B$  decays, in particular the decays known as  $B \rightarrow h^+h'^-$  decays, was given in Sec. 1.4. There are a number of interesting physics measurements that can be made by LHCb via the study of these decays. One example, described in detail in Sec. 1.4.4.2, is the measurement of the CKM angle  $\gamma$  using measurements of the time-dependent  $\mathcal{CP}$  asymmetries in the decays  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^+K^-$ . The “time” variable appearing in the expressions for decay rates and  $\mathcal{CP}$  asymmetries that were given in Secs. 1.2 and 1.4 is the lifetime of a particular meson in its own reference frame. This is known as the *proper time* of the meson, and is usually given the symbol  $\tau$ . It is measured experimentally by reconstructing the decay products of the meson, and performing the vector dot product of its decay length vector  $\vec{L}$  with its momentum vector  $\vec{p}$ . The proper time is then

$$\tau = m \frac{\vec{L} \cdot \vec{p}}{|\vec{p}|^2}, \quad (4.1)$$

where  $m$  is the mass of the meson.

Measuring the time-dependent  $\mathcal{CP}$  asymmetries in a given decay involves a comparison of the reconstructed proper time ( $\tau_{rec}$ ) distribution for events where the initial state of the  $B$  meson has been tagged as containing a  $b$  quark with the corresponding distribution for events where the initial state of the  $B$  meson has been tagged as containing a  $\bar{b}$  quark. Such *flavour-tagged* proper time distributions are affected by experimental factors — namely the finite proper time resolution of the detector, and the incorrect flavour tagging of a certain percentage of events (mistag). So to optimise an analysis aiming to measure time-dependent

$\mathcal{CP}$  asymmetries, knowledge of both the mistag rate and the proper time resolution from data is needed.

This chapter describes a method to extract the proper time resolution of the LHCb detector for  $B \rightarrow h^+h'^-$  decays. The method only uses information which will be present in data. The mistag rate is assumed to be known from measurements in control channels, which will be discussed. A proper time resolution model for  $B \rightarrow h^+h'^-$  is proposed, and then validated by showing that it accurately describes the proper time residual distributions for  $B \rightarrow h^+h'^-$  decays. The model used is a development of one presented in [112]. It is then shown that the values of the parameters of this model can be extracted by fitting to the reconstructed proper time distribution of flavour tagged events for the decay  $B_s \rightarrow K^-\pi^+$  (the reasons for this choice of decay are discussed in Sec. 4.1.3). The fit also makes use of the per-event errors ( $\sigma_{\tau_{rec}}$ ) on  $\tau_{rec}$ , but, crucially since the method is to be applied to data, it does not use any information on the true lifetime ( $\tau_{true}$ ) of the Monte Carlo simulated events. An earlier version of the study presented here is documented in [113].

Section 4.1 discusses how the flavour-tagged proper time distributions are affected by the experimental factors mentioned above. In Sec. 4.2, the model to describe the proper time resolution is introduced, and then validated via its application to four different  $B \rightarrow h^+h'^-$  decays. The determination of the parameters of the resolution model for  $B \rightarrow h^+h'^-$  decays from data is described in Sec. 4.3. Conclusions are given in Sec. 4.4.

## 4.1 Experimental Effects on Decay Rates

The theoretical expressions for the proper time distributions (decay rates) in neutral meson decays were given in Sec. 1.2.2.2. These expressions do not take into account experimental features such as the finite proper time resolution of the detector, and the imperfect flavour tagging process. The effect of these features on the decay rates and  $\mathcal{CP}$  asymmetries in neutral meson decays will now be discussed. For the sake of simplicity the discussion will assume that the width difference  $\Delta\Gamma$  is zero for both the  $B_d$  and  $B_s$  systems, as the effect from including non-zero  $\Delta\Gamma$  values compatible with the Standard Model expectations is negligible. The simplification  $|p/q| = 1$  will also be made for both  $B_d$  and  $B_s$ . While  $|p/q|_{B_s}$  (unlike  $|p/q|_{B_d}$ ) has yet to be well measured experimentally, it is (like  $|p/q|_{B_d}$ ) predicted to be very close to 1 in the SM.

### 4.1.1 Proper Time Resolution Effects

With the assumptions stated above, the decay rate for a neutral meson is (omitting time-independent factors such as  $|A_f|$ ):

$$\Gamma(N^0(\tau_{true}) \rightarrow f_{\mathcal{CP}}) \propto e^{-\Gamma\tau_{true}} [1 + \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta m\tau_{true}) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta m\tau_{true})] \quad (4.2)$$

for a decay to a  $\mathcal{CP}$  eigenstate (see Eqn. 1.51), and

$$\Gamma(N^0(\tau_{true}) \rightarrow f_{fs}) \propto e^{-\Gamma\tau_{true}} [1 + \cos(\Delta m\tau_{true})] \quad (4.3)$$

for a decay to a flavour-specific final state (see Eqn. 1.53). The variable  $\tau_{true}$  denotes the actual (true) proper time of the meson.

In the case of finite proper time resolution, the decay rate as a function of  $\tau_{rec}$ , the meson proper time as reconstructed (measured) by the detector, is found by convoluting the above theoretical expressions with the resolution function  $R(\Delta\tau)$ , where  $\Delta\tau \equiv \tau_{rec} - \tau_{true}$ . If  $R(\Delta\tau)$  is a Gaussian of fixed width  $\sigma_\tau$  and mean 0, then the measured decay rate becomes

$$\Gamma(N^0(\tau_{rec}) \rightarrow f_{\mathcal{CP}}) \propto \int_0^\infty e^{-\Gamma\tau_{true}} [1 + \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta m\tau_{true}) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta m\tau_{true})] e^{-\frac{1}{2}(\frac{\Delta\tau}{\sigma_\tau})^2} d\tau_{true} \quad (4.4)$$

for the  $\mathcal{CP}$  eigenstate case, and

$$\Gamma(N^0(\tau_{rec}) \rightarrow f_{fs}) \propto \int_0^\infty e^{-\Gamma\tau_{true}} [1 + \cos(\Delta m\tau_{true})] e^{-\frac{1}{2}(\frac{\Delta\tau}{\sigma_\tau})^2} d\tau_{true} \quad (4.5)$$

for the flavour-specific final state case. An analytical calculation of integrals of the above type shows [114] that the coefficient of each oscillating term ( $\cos(\Delta m\tau_{true})$  or  $\sin(\Delta m\tau_{true})$ ) is reduced by a factor  $D_{\sigma_\tau}$ , where

$$D_{\sigma_\tau} \equiv e^{-\frac{1}{2}(\Delta m\sigma_\tau)^2}. \quad (4.6)$$

This is the most significant change in the decay rate distribution for  $B \rightarrow h^+h'^-$  decays that is introduced when the proper time resolution is accounted for [114]. In principle, the proper time resolution also affects the observed distribution for small values of  $\tau_{rec}$ , as the reconstructed value can be negative if the true value is small enough. Such effects can be seen for decays where the selection is lifetime-unbiased, i.e. the selection does not include cuts (such as minimum cuts on the daughter particle impact parameters) which prevent  $B$  particles with small proper times from surviving the selection.

Examples of decays with lifetime-unbiased selections in LHCb are  $B_d \rightarrow J/\psi K^*$  and  $B^+ \rightarrow J/\psi K^+$ . Such a selection is used in order to allow the full angular distribution of

the daughter particles to be studied (these decays act as control channels for  $B_s \rightarrow J/\psi\phi$ , where the full angular distribution is needed to extract the  $\mathcal{CP}$  asymmetries, as the final state is not a  $\mathcal{CP}$  eigenstate). The  $\tau_{rec}$  distribution for these  $J/\psi X$  decays extends slightly into the unphysical negative region, by an amount dependent on the proper time resolution [112]. However for  $B \rightarrow h^+h'^-$  decays, impact parameter cuts are an important part of the selection (see Sec 4.2.3), so there are no selected events with very small proper times, and the shape of the  $\tau_{rec}$  distribution at the point where events start to pass the impact parameter cuts depends more strongly on these cuts than on the proper time resolution [56].

The dilution factor  $D_{\sigma_\tau}$  reflects the fact that the oscillations will be completely resolved (i.e.  $D_{\sigma_\tau} \simeq 1$ ) as long as  $\sigma_\tau \ll 1/(\Delta m)$ . It will be seen in Sec. 4.2 that the proper time resolution of the LHCb detector is approximately 40 fs. Assuming this value for  $\sigma_\tau$ , the dilution factor is

$$D_{\sigma_\tau} = e^{-\frac{1}{2}((0.507 \text{ ps}^{-1})(0.04 \text{ ps}))^2} = e^{-\frac{1}{2}(0.02)^2} = 0.9998 \quad \text{for } B_d, \quad (4.7)$$

$$D_{\sigma_\tau} = e^{-\frac{1}{2}((17.8 \text{ ps}^{-1})(0.04 \text{ ps}))^2} = e^{-\frac{1}{2}(0.71)^2} = 0.776 \quad \text{for } B_s. \quad (4.8)$$

So for  $B_d$  decays (and also for  $D^0$  decays, where  $\Delta m \simeq 0.02 \text{ ps}^{-1}$ ), the dilution factor is  $\simeq 1$  and can be neglected. However for  $B_s$  decays, the dilution factor is very different from 1, and has to be taken into account if the  $\mathcal{CP}$  asymmetries are to be measured correctly. A strategy to take account of the dilution factor will be outlined in Sec. 4.1.3.

### 4.1.2 Mistag Effects

The performance of the flavour tagging process is parameterised in terms of two variables: the tagging efficiency  $\epsilon_{\text{tag}}$  and the mistag rate  $\omega_{\text{tag}}$ . The tagging efficiency is the probability that the tagging algorithm provides a decision, tagging the event. The mistag rate is the probability (given that a tagging decision was made) that the tagging decision is incorrect. For a non-zero mistag rate, the observed decay rate for a neutral meson which is tagged as being (say) an  $N^0$  state at production becomes a superposition of the decay rates for  $N^0$  and  $\overline{N}^0$ :

$$\Gamma^{\text{tagged}}(N^0(\tau_{rec}) \rightarrow f) = \epsilon_{\text{tag}}((1 - \omega_{\text{tag}})\Gamma(N^0(\tau_{rec}) \rightarrow f) + \omega_{\text{tag}}\Gamma(\overline{N}^0(\tau_{rec}) \rightarrow f)). \quad (4.9)$$

This leads to [56] the coefficients of the oscillating terms suffering a tagging-related dilution of

$$D_{\text{tag}} \equiv (1 - 2\omega_{\text{tag}}), \quad (4.10)$$

regardless of whether the final state is a  $\mathcal{CP}$  eigenstate or flavour-specific. Since the expected mistag rates for  $B \rightarrow h^+h'^-$  decays are significant [75], the dilution factor from tagging is significantly different from 1 for both  $B_d$  and  $B_s$ :

$$D_{\text{tag}} = (1 - 2(0.37)) = 0.26 \quad \text{for } B_d, \quad (4.11)$$

$$D_{\text{tag}} = (1 - 2(0.34)) = 0.32 \quad \text{for } B_s. \quad (4.12)$$

Clearly the dilution from imperfect tagging needs to be taken into account to correctly measure  $\mathcal{CP}$  asymmetries in both  $B_d$  and  $B_s$  decays. A strategy to handle the dilutions from both tagging and proper time resolution will now be given.

### 4.1.3 Measuring Experimental Effects on Data

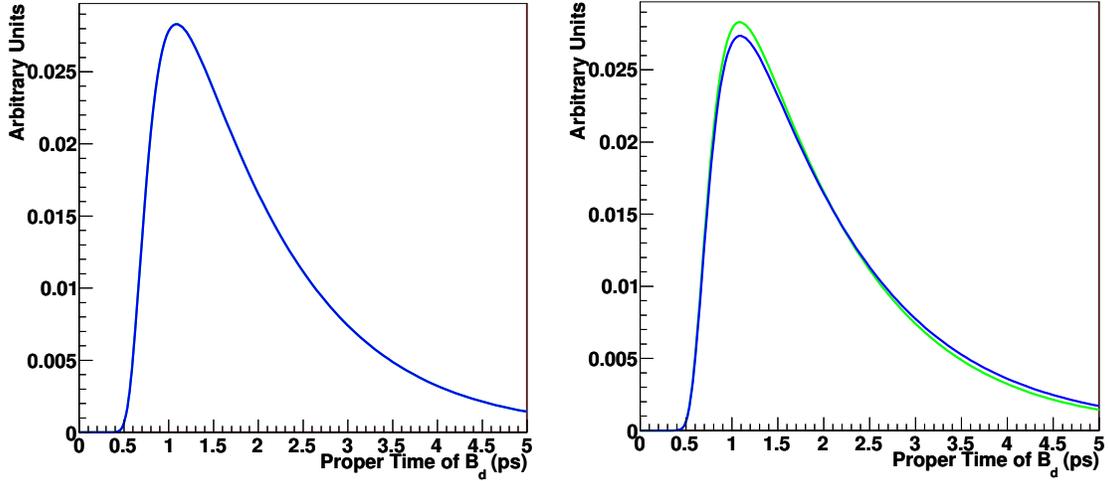
When both of the dilutions described in the previous sections are taken into account, the overall coefficients of the oscillating terms in the experimental neutral meson decay rates are  $D_{\sigma_\tau} D_{\text{tag}} \mathcal{A}_{\mathcal{CP}}^{\text{dir}/\text{mix}}$  for decays to  $\mathcal{CP}$  eigenstates, and  $D_{\sigma_\tau} D_{\text{tag}}$  for flavour-specific decays. The four “main”  $B \rightarrow h^+h'^-$  decays,  $B_d \rightarrow \pi^+\pi^-$ ,  $B_s \rightarrow K^+K^-$ ,  $B_s \rightarrow K^-\pi^+$  and  $B_d \rightarrow K^+\pi^-$ , were highlighted in Sec. 1.4.1 as being the  $B \rightarrow h^+h'^-$  decays with the highest branching ratios. The two modes with the  $K\pi$  final state are flavour-specific, while the other two decays are to  $\mathcal{CP}$  eigenstates. So if the dilution factors can be measured in the flavour-specific decays, this knowledge can be used to extract the  $\mathcal{CP}$  asymmetries in decays to  $\mathcal{CP}$  eigenstates (which is what is required to measure  $\gamma$ ).

#### 4.1.3.1 The $B_d$ Decays

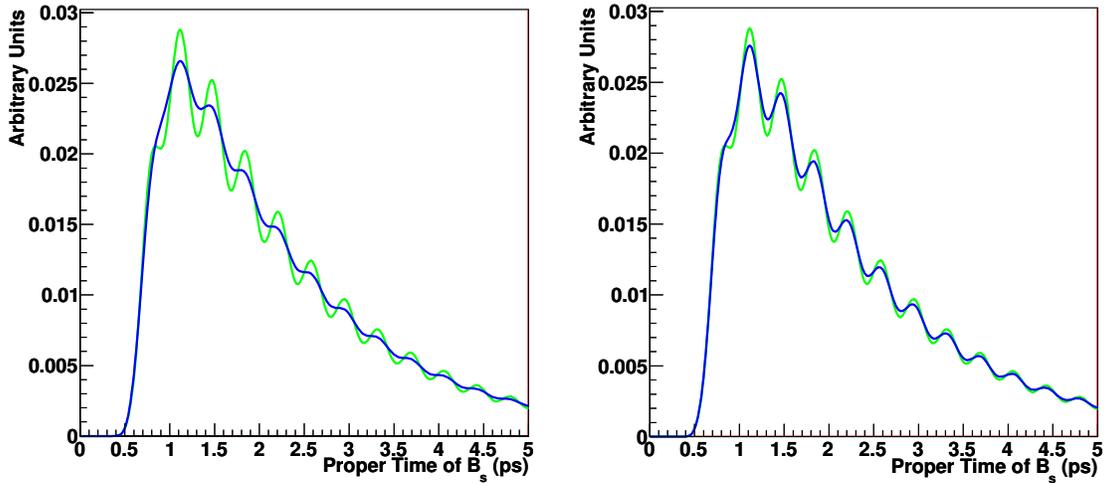
The previous sections have shown that for  $B_d$  decays, the dilution from the proper time resolution is negligible. So the only important dilution is from the mistag rate. This is illustrated in Fig 4.1, where a change in the proper time resolution has no effect on the flavour tagged  $\tau_{rec}$  distribution, but an increased mistag rate does have an effect. Since the mistag rates for  $B_d \rightarrow \pi^+\pi^-$  and  $B_d \rightarrow K^+\pi^-$  are expected to be almost identical [75], the tagging dilution factor as measured from  $B_d \rightarrow K^+\pi^-$  should be directly applicable to  $B_d \rightarrow \pi^+\pi^-$ .

#### 4.1.3.2 The $B_s$ Decays

For the  $B_s$  decays, the situation is more complicated, as the dilution from the proper time resolution cannot be neglected. This is illustrated in Fig. 4.2. Changes in both the proper time resolution and the mistag rate affect the depth of the oscillations for  $B_s \rightarrow K^-\pi^+$ . Hence the tagging dilution for the  $B_s$  channels cannot simply be measured from the flavour-specific



**Figure 4.1:** Illustration of the impact of the proper time resolution  $\sigma_\tau$  and the mistag rate  $\omega_{\text{tag}}$  on the flavour tagged  $\tau_{\text{rec}}$  distribution for  $B_d \rightarrow K^+\pi^-$ . In the left hand plot, the distribution shown in blue has  $\omega_{\text{tag}}=34\%$ , and a resolution model which is a Gaussian with fixed width  $\sigma_\tau=70$  fs. A distribution having  $\omega_{\text{tag}}=34\%$  and a Gaussian resolution model with fixed width 40 fs is also plotted, but is not visible as it overlaps completely with the distribution in blue. In the right hand plot, both distributions have  $\sigma_\tau=40$  fs, with the distribution shown in green (blue) having  $\omega_{\text{tag}}=34\%$  (40%).



**Figure 4.2:** Illustration of the impact of the proper time resolution  $\sigma_\tau$  and the mistag rate  $\omega_{\text{tag}}$  on the flavour tagged  $\tau_{\text{rec}}$  distribution for  $B_s \rightarrow K^-\pi^+$ . In the left hand plot, both distributions have  $\omega_{\text{tag}}=34\%$ , with the distribution shown in green (blue) having a resolution model which is a Gaussian with fixed width  $\sigma_\tau=40$  fs (70 fs). In the right hand plot, both distributions have  $\sigma_\tau=40$  fs, with the distribution shown in green (blue) having  $\omega_{\text{tag}}=34\%$  (40%).

decay  $B_s \rightarrow K^- \pi^+$ . Neither can the tagging dilution as measured from  $B_d \rightarrow K^+ \pi^-$  be used, as the mistag rates for  $B_d$  and  $B_s$  decays are not in general the same. This is due to the use of different same-side taggers for  $B_d$  and  $B_s$  decays, corresponding to the hadron that can be produced in conjunction with each type of meson. The  $B_d$  tagging uses a same side pion tagger, while the  $B_s$  tagging uses a same side kaon tagger. The same side kaon tagger has a lower mistag rate than the same side pion tagger, because there are fewer kaons than pions produced in a typical event, so the likelihood of the kaon reconstructed by the tagging algorithm as being associated with the production of the  $B_s$  meson being the correct kaon is higher than the equivalent likelihood for a pion. This causes the overall mistag rate for  $B_s$  mesons to be lower than that for  $B_d$  decays, even though they use the same opposite-side taggers.

One possible strategy to measure the mistag rate for  $B_s \rightarrow K^- \pi^+$  is to take the mistag rate from  $B_d \rightarrow K^+ \pi^-$ , and measure the decrease in rate due to the same-side differences described above using the differences in mistag rate seen in a pair of channels where this can be measured, such as  $B_s \rightarrow D_s^- \pi^+$  and  $B_d \rightarrow D^- \pi^+$ .

Once the mistag rate for  $B_s \rightarrow K^- \pi^+$  has been determined, the dilution due to the proper time resolution can be measured by studying the flavour-tagged  $\tau_{rec}$  distribution. A method to measure the proper time resolution in this way is the subject of Sec. 4.3. Both the tagging and proper time dilution factors can be applied to  $B_s \rightarrow K^- K^+$  — the mistag rates for  $B_s \rightarrow K^+ K^-$ ,  $B_s \rightarrow K^- \pi^+$  are expected to be very similar, and it will be shown in Sec. 4.2 that both channels have the same proper time resolution. This will allow the  $\mathcal{CP}$  asymmetries in  $B_s \rightarrow K^- K^+$  to be extracted.

## 4.2 Validation of the Resolution Model on Full Monte Carlo Simulation

This section introduces a proper time resolution model, and shows, making use of the Monte Carlo simulated truth information, that it gives an accurate description of the proper time residual distributions of the four main  $B \rightarrow h^+ h'^-$  decays. It will also be shown that the parameters of the proper time resolution model are similar between these  $B \rightarrow h^+ h'^-$  decays. This is necessary to allow the results from the fit to  $B_s \rightarrow K^- \pi^+$  data to be applied to  $B_s \rightarrow K^- K^+$ .

## 4.2.1 Form of the Proper Time Resolution Model

The model used here for  $B \rightarrow h^+h'^-$  decays is defined in Eqn 4.13. It is an adaptation of a model developed by members of the LHCb group at NIKHEF in the Netherlands. This model, and its application to the decays  $B_d \rightarrow J/\psi K^*$  and  $B^+ \rightarrow J/\psi K^+$ , is described fully in [112]. The model for  $B \rightarrow h^+h'^-$  decays takes the form of a Gaussian with mean  $M$  and width  $S$ :

$$R(\Delta\tau) = N e^{-\frac{1}{2}\left(\frac{\Delta\tau-M}{S}\right)^2}, \quad (4.13)$$

where  $N$  is a normalisation factor. Previous studies [113] used more complicated models with more parameters. However these models only provide a small improvement in the description of the residuals, at the expense of unreliable fits to the flavour-tagged proper time distribution, due to the fact that more parameters need to be fitted.

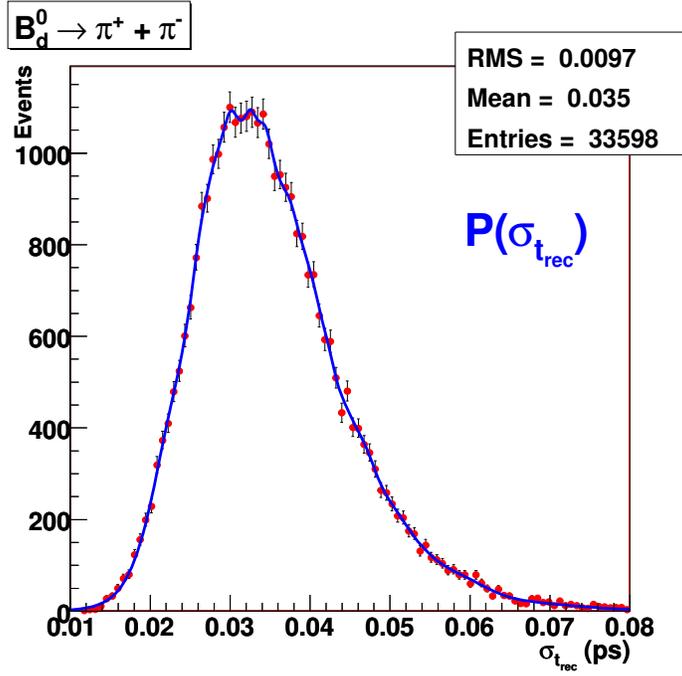
The model given above aims to describe the residuals of only one decay at a time (i.e. only signal events). This is not a problem when studying simulated data, as any background events which may be present can be eliminated by looking at the simulated truth information. The strategy for studying data (i.e. the reconstructed proper time distribution), where background events will be present, is described in Sec. 4.3.

## 4.2.2 Method for Validation of Model

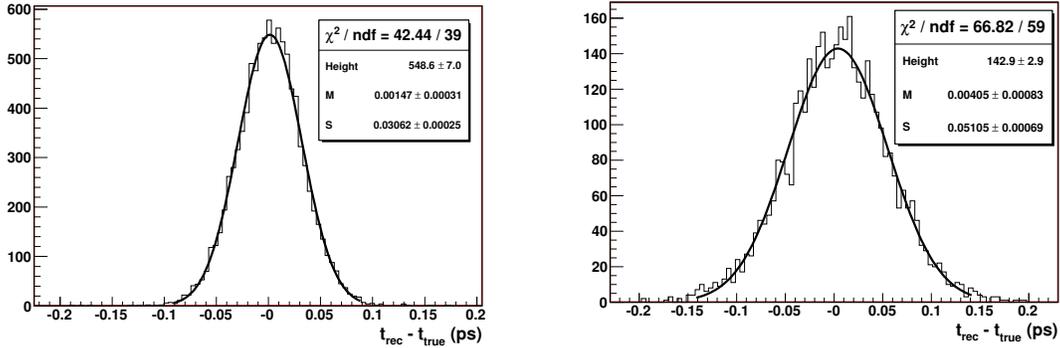
The method described here follows the method used in [112]. For each event, the reconstructed proper time  $\tau_{rec}$  has a per-event error  $\sigma_{\tau_{rec}}$  associated with it. This error is calculated by propagating the errors on the measurements which are combined to calculate the proper time — namely the momentum, flight distance and reconstructed mass of the  $B$ -meson (see Eqn. 4.1). In LHCb, a lifetime fitter tool has been developed [115] that calculates the  $\tau_{rec}$  and the  $\sigma_{\tau_{rec}}$  for each event. In most fully reconstructed  $B$  decay channels at LHCb, the  $\sigma_{\tau_{rec}}$  distributions are broadly similar to each other. As an example, the  $\sigma_{\tau_{rec}}$  distribution for  $B_d \rightarrow \pi^+\pi^-$  is shown in Fig. 4.3. The mean of this distribution, 35 fs, sets the scale for the expected values for the resolution model parameter  $S$ .

Although the  $\sigma_{\tau_{rec}}$  distribution does not follow an analytical formula, a PDF can be constructed that follows the distribution closely. This is done using the method of kernel estimation [116], where the PDF is built by a superposition of Gaussians, one for each data point in the distribution. An example of such a PDF is superimposed on the  $\sigma_{\tau_{rec}}$  distribution in Fig. 4.3.

For the model to be useful, its parameters have to be global, i.e. they have to describe the resolution model adequately for all signal events. However as signal events have a range of values of  $\sigma_{\tau_{rec}}$ , the same value of, for example,  $S$  is not expected to be able to describe the



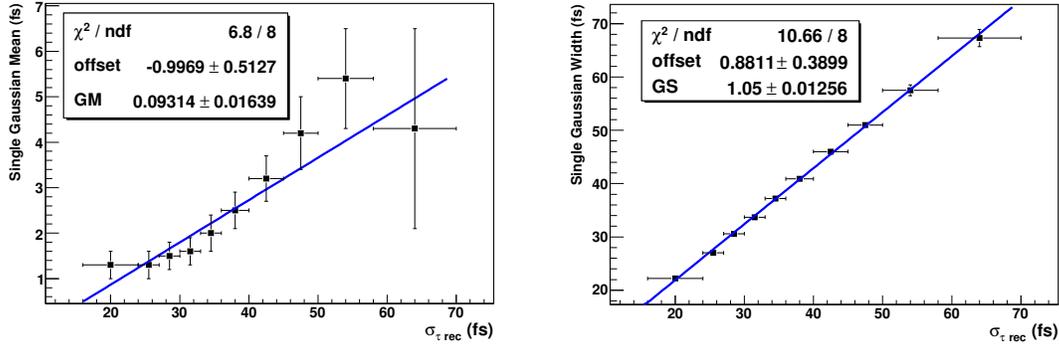
**Figure 4.3:** Proper time per-event error ( $\sigma_{\tau_{rec}}$ ) distribution (red points) for  $B_d \rightarrow \pi^+ \pi^-$ , fitted with a PDF (blue curve) built by kernel estimation.



**Figure 4.4:** Single Gaussian fits to proper time residual  $\Delta\tau$  in two bins of the proper time per-event error  $\sigma_{\tau_{rec}}$ . Left: Fit to events with  $\sigma_{\tau_{rec}} \in [27, 30]$  fs. Right: Fit to events with  $\sigma_{\tau_{rec}} \in [45, 50]$  fs.

width of the entire distribution. The goal is to determine a globally applicable model where the parameters of  $R(\Delta\tau)$  can be determined on an “event-by-event” basis according to some reconstructed parameter(s) of each event.

To search for a suitable parameterisation for  $M$  and  $S$ , the two parameters of  $R(\Delta\tau)$ , the  $\Delta\tau$  distribution is plotted in bins of some variable, then fitted with  $R(\Delta\tau)$ . The dependency of  $M$  and  $S$  on the binned variable can then be seen. Example fits to the the  $\Delta\tau$  distribution in two bins of the per-event proper time error,  $\sigma_{\tau_{rec}}$ , can be seen in Fig. 4.4.



**Figure 4.5:** Dependencies of the single Gaussian mean (left) and width (right) on the proper time per-event error  $\sigma_{\tau_{rec}}$ .

It is found that  $M$  and  $S$  can be well parameterised in a simple way, as linear functions of  $\sigma_{\tau_{rec}}$ . The dependencies of  $M$  and  $S$  on  $\sigma_{\tau_{rec}}$  are shown in Fig. 4.5. The scaling factor between  $M$  and  $\sigma_{\tau_{rec}}$  is named  $GM$ , and the scaling factor between  $S$  and  $\sigma_{\tau_{rec}}$  is named  $GS$ . A first-degree polynomial fit is made, allowing for some non-zero intercept as well as the gradient. The offset of the intercept from zero is very small ( $\simeq 1$  fs for both  $M$  and  $S$ ). It will be neglected from now on, as taking it into account would introduce two extra parameters (the offset for  $M$  and the offset for  $S$ ) into the global fit.

Using the dependencies as described, each event contributes a Gaussian with mean  $M = GM \cdot \sigma_{\tau_{rec}}$  and width  $S = GS \cdot \sigma_{\tau_{rec}}$  to the overall distribution. If there is no bias in the calculation of  $\tau_{rec}$ ,  $GM$  should be 0. If  $\sigma_{\tau_{rec}}$  is correctly calculated for each event,  $GS$  should be 1.

Using the dependencies of  $M$  and  $S$  on  $\sigma_{\tau_{rec}}$ ,  $R(\Delta\tau)$  can be made conditional on the  $\sigma_{\tau_{rec}}$  distribution. The aim is to obtain a resolution model that is a function of the per-event error. If  $P(\sigma_{\tau_{rec}})$  is a PDF built by kernel estimation from the  $\sigma_{\tau_{rec}}$  distribution, then the equation:

$$R(\Delta\tau) = \int_0^{\infty} R(\Delta\tau|\sigma_{\tau_{rec}})P(\sigma_{\tau_{rec}})d\sigma_{\tau_{rec}} \quad (4.14)$$

allows  $R(\Delta\tau)$  to be simultaneously fitted to the  $\Delta\tau$  distribution for all signal events, regardless of their  $\sigma_{\tau_{rec}}$  value. A fit can be performed to determine the values of  $GM$  and  $GS$  which give the best fit to the  $\Delta\tau$  distribution. The values of  $GM$  and  $GS$  obtained from this fit should be more reliable than those obtained from fitting single Gaussians to the  $\Delta\tau$  distribution in bins of  $\sigma_{\tau_{rec}}$  (see Fig. 4.5) as they do not depend on the choice of binning.

If, instead of the model described above, a simple Gaussian with some fixed global width (say the mean of the  $\sigma_{\tau_{rec}}$  distribution) was used for the proper time resolution model, the model would likely provide an adequate description of the  $\Delta\tau$  distribution around its peak, but would completely underestimate the tails of the distribution, due to the presence of events

with high values of  $\sigma_{\tau_{rec}}$ . If such a simple model were used in the analysis which studies the flavour tagged  $\tau_{rec}$  distributions to extract  $\mathcal{CP}$  asymmetries, extra systematic errors would most likely be introduced.

It should be noted that there are other methods which could possibly extract the proper time resolution model for the  $B \rightarrow h^+h'^-$  channels from data. One example involves the study of the reconstructed proper time distribution of prompt  $J/\psi \rightarrow \mu^+\mu^-$  events. The proper time residual for prompt  $J/\psi \rightarrow \mu^+\mu^-$  is the same as the reconstructed proper time, because the true lifetime is known to be (very close to) zero. The unphysical negative side of the reconstructed proper time distribution then gives information on the proper time resolution model for  $J/\psi \rightarrow \mu^+\mu^-$ . The open question is how similar this model would be to the model for the  $B \rightarrow h^+h'^-$  channels.

### 4.2.3 Selection of $B \rightarrow h^+h'^-$ Decays

The proper time resolution model is applied, using the method outlined in the previous section, to Monte Carlo simulated distributions of  $\Delta\tau$ . The method is applied to selected events in each of the four decays which have the highest branching ratios of the possible  $B \rightarrow h^+h'^-$  decays:  $B_d \rightarrow \pi^+\pi^-$ ,  $B_s \rightarrow K^+K^-$ ,  $B_s \rightarrow K^-\pi^+$  and  $B_d \rightarrow K^+\pi^-$ . To run the event selection and produce the  $\Delta\tau$  distribution, v19r10 of the LHCb physics analysis package DaVinci [117] is used. The selection used for this study will now be described, beginning with the “inclusive” selection for  $B \rightarrow h^+h'^-$  decays, which forms the basis for the selections used not only in this chapter, but also in chapters 5 and 6.

An inclusive selection has been developed within LHCb to select  $B \rightarrow h^+h'^-$  events. This selection is designed to efficiently select any decay of the  $B \rightarrow h^+h'^-$  type, while achieving excellent background rejection. For analyses where only one particular  $B \rightarrow h^+h'^-$  decay is of interest, and other  $B \rightarrow h^+h'^-$  decays are considered as backgrounds, the inclusive selection is turned into an exclusive selection by the addition of cuts on the invariant mass of the  $B$  and the PID likelihoods for the daughters. The exact mass and PID cuts that are applied are of course dependent on which  $B \rightarrow h^+h'^-$  mode is considered as the signal. Here the cuts that comprise the inclusive selection will be given and the motivation for each cut briefly described; a full discussion is given in [56]. An explanation of the process used to optimise the cut values for a previous inclusive  $B \rightarrow h^+h'^-$  selection for LHCb, very similar to the selection used here, can be found in [118]. The distributions of the cut variables for signal and background can also be found there.

The events that are mostly likely to form the main backgrounds to  $B \rightarrow h^+h'^-$  events, i.e. the events that an inclusive selection aims to suppress as far as possible, fall into two broad categories:

- **Physics background:** There are many decays of  $B$  hadrons where the final state (as seen by the detector) consists of three hadrons ( $\pi$ ,  $K$  or  $p$ ). These decays, known within LHCb as  $B \rightarrow hhh$  decays, often proceed via intermediate resonances (for example  $B \rightarrow \rho\pi$  or  $B \rightarrow K^*\pi$ ), but non-resonant three-body decays can also occur. When one of these final state hadrons is not correctly reconstructed, the resulting experimental signature can mimic that of a  $B \rightarrow h^+h'^-$  event. The probability of a hadron not being correctly reconstructed is far higher for neutral hadrons than for charged hadrons: for  $\pi^0$  particles that decay into two photons each having  $p_T > 200$  MeV, the reconstruction efficiency is around 50% [70], while for charged hadrons with  $p_T > 10$  GeV and with tracks formed using hits from both the VeLo and the tracking stations<sup>1</sup>, the reconstruction efficiency is around 94% [70]. This means that, for example, an event of the type  $B_d \rightarrow \pi^+\pi^-\pi^0$  is more likely to be reconstructed as a  $B \rightarrow h^+h'^-$  event than one of the type  $B^+ \rightarrow \pi^+\pi^-\pi^+$ , because the reconstruction efficiency for  $\pi^0$  is far lower than for  $\pi^+$ .

These “partially reconstructed” events have an invariant mass that is kinematically limited to be below the relevant  $B$  mass, if the correct mass hypothesis is used. However the mass hypothesis for certain  $B \rightarrow h^+h'^-$  decays can cause partially reconstructed decays to fall under the signal mass peak (this will be seen in Chapter 5 to be a particular issue for the decay  $B_d \rightarrow p\bar{p}$ ).

- **Combinatorial background:** This refers to backgrounds where the candidates for the final state particles (the  $h^+h'^-$  candidates in the current case) do not originate from a single  $B$  decay. They may be genuine charged hadrons originating from decays of two different  $B$  hadrons (which is “combinatorial background” in the truest sense), or there may be one genuine charged hadron from a  $B$  decay, and one fake or “ghost” track<sup>2</sup>. These two possibilities cover most of the combinatorial background that is found to pass loose preselections for  $B \rightarrow h^+h'^-$  events. In principle it is also possible for tracks from  $D$  decays, or even from the primary vertex, to pass the selection, but any contribution from these should be less significant than that from the two main types just described.

To suppress combinatorial background, cuts are placed on the  $h^+h'^-$  candidate tracks to select tracks which have the characteristics of originating from the decay of a  $B$  hadron. Due to the large masses of  $B$  hadrons, the  $B$  daughters typically have higher  $p_T$  values than other

---

<sup>1</sup>In LHCb such tracks are referred to as “long tracks”.

<sup>2</sup>In LHCb, a track is classified as a ghost track in simulation if less than 70% of the hits used to form the track originate from a single (true) particle.

particles, and due to the relatively long lifetimes of  $B$  hadrons, the  $B$  daughters will tend to have large impact parameters<sup>3</sup> (IP). The likelihood for an impact parameter to be non-zero, and hence for the track not to have originated from the primary vertex (PV) of the collision, is evaluated using the impact parameter significance (IPS), which is the IP divided by its error. So in the selection, minimum values are required for the  $p_T$  and IPS of both daughter tracks.

Having selected the daughters, the only other cuts to be made are those on the mother (the  $B$  candidate), because there are no intermediate resonances involved. This means that the total number of cuts in the selection is lower than in selections for decay modes involving intermediate resonances, so the individual cuts can be tighter (or “harder”) while maintaining a reasonable total efficiency for signal events. However there is a corresponding disadvantage, which is that there are fewer cuts for background events to survive, so it is more likely that a background event can fake the signature of a signal event (this is true of both physics backgrounds and combinatorial backgrounds).

The first cut applied when forming the mother candidate is a maximum cut on the  $\chi^2$  of the vertex of the daughter tracks, since the tracks should originate from the same mother. The long lifetimes of  $B$  hadrons justify a minimum cut on the flight distance significance (FDS) of the mother with respect to the primary vertex. The requirement that the  $B$  hadron originated at the PV translates into a maximum cut on the mother IPS (note the contrast with the *minimum* cut on the daughter IPS, which ensures that the daughters do *not* originate from the PV). Finally, since  $B$  hadrons are produced at larger opening angles than lighter particles (e.g.  $D$  hadrons), their  $p_T$  will on average be larger. So a minimum cut on the  $p_T$  of the  $B$  is applied. The complete set of selection cuts that comprise the inclusive  $B \rightarrow h^+h'^-$  selection is summarised in Table 4.1.

To form an exclusive selection for each  $B \rightarrow h^+h'^-$  decay considered here, cuts on particle identification (PID) and invariant mass are added to the inclusive selection. The PID variable that is cut on is the difference in log-likelihood (DLL) for two PID hypotheses. For example,  $\text{DLL}(K - \pi)$  is the log-likelihood for a kaon hypothesis, less the log-likelihood for a pion hypothesis. If  $\text{DLL}(K - \pi)$  is positive (negative), the particle is more likely to be a kaon (pion). The cuts that are added to the inclusive selection are given in Table 4.2.

The efficiencies of the relevant exclusive selection for each decay were calculated. The total selection efficiency, taking into account the geometrical efficiency of the LHCb detector, is found to be  $\simeq 4\%$  for all four decays considered. These efficiencies are very similar to those reported in [118]. From each signal sample, a few percent of the events passing the

---

<sup>3</sup>As stated in Sec. 2.2.1, the impact parameter of a track is defined as the distance of closest approach of a track to the primary vertex.

Cut Variable	Value
Max. $\chi^2$ of $B$ vertex	5.0
Min. $p_T$ of $B$	1.0 GeV
Max. IPS of $B$	2.5
Min. FDS of $B$	18.0
Min. $p_T$ for both daughters	1.0 GeV
Min. $p_T$ for (at least) one daughter	3.0 GeV
Min. IPS for both daughters	6.0
Min. IPS for (at least) one daughter	12.0

**Table 4.1:** List of cuts comprising the inclusive  $B \rightarrow h^+h'^-$  selection. Abbreviations are explained in the text.

Type of Cut	Value
Mass Window for $B$	$\pm 50$ MeV
Minimum DLL( $K - \pi$ ) for $K$ candidates	0.0
Maximum DLL( $K - \pi$ ) for $\pi$ candidates	0.0

**Table 4.2:** Mass and PID cuts for the exclusive  $B \rightarrow h^+h'^-$  selections.

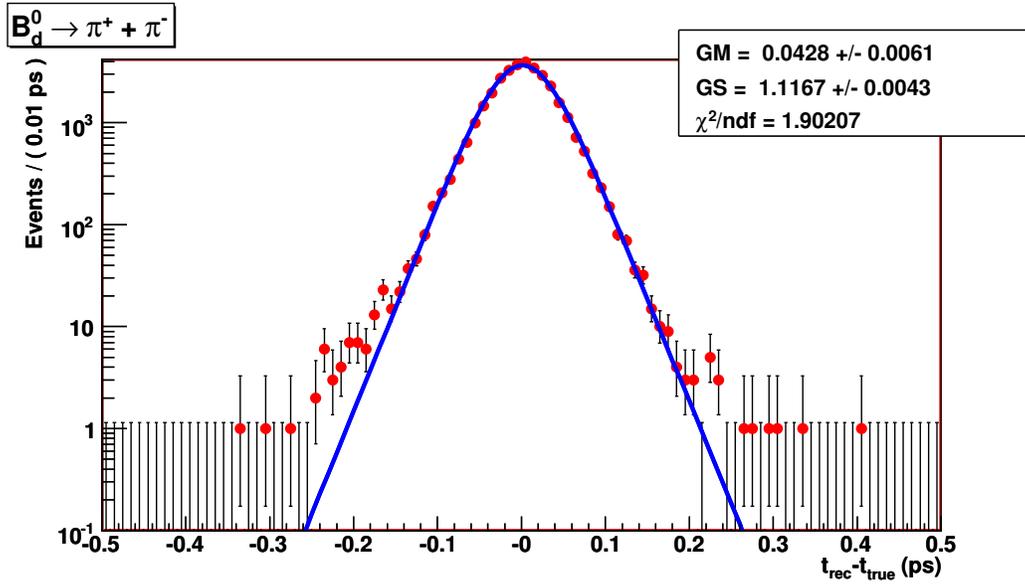
selection are found not to be signal events. Most of these non-signal events are formed using one reconstructed track from a true signal decay, and one ghost track. The non-signal events are removed from the set of selected events by applying a cut on the Monte Carlo simulated truth information for each event. This ensures that the validation of the proper time resolution model is carried out on true signal events only.

#### 4.2.4 Results of Model Validation

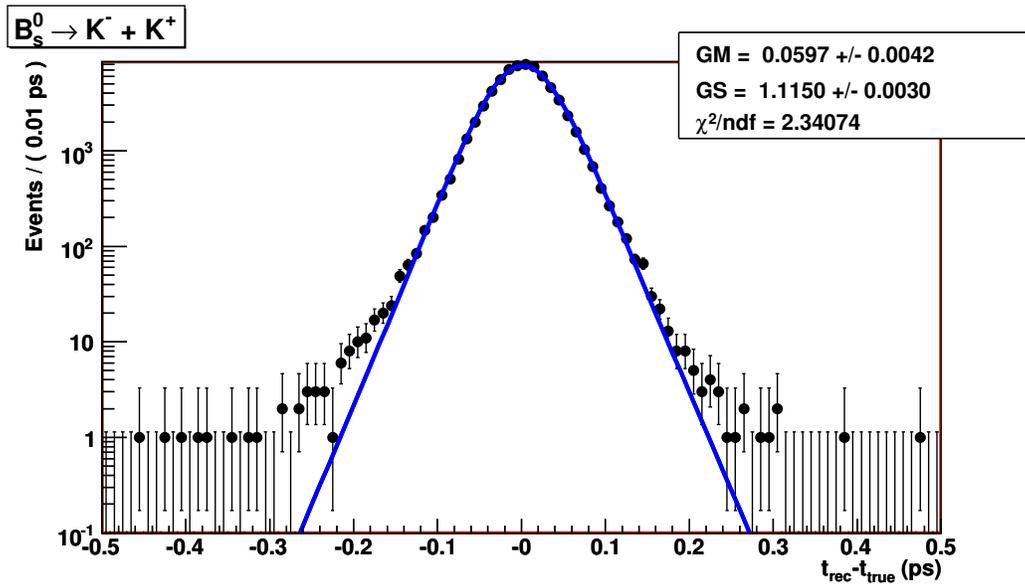
Once signal events for a particular  $B \rightarrow h^+h'^-$  decay have been selected as described above, the  $\Delta\tau$  distribution is fitted using  $R(\Delta\tau)$ , making use of the per-event errors (see Eqn. 4.14). Figures 4.6 to 4.9 show the fits to each  $B \rightarrow h^+h'^-$  decay in turn. The values of the parameters of the model are displayed in the top right corner of the graph, along with the  $\chi^2$  per degree of freedom of the fit.

For each decay, it can be seen that good fits are obtained over most values of  $\Delta\tau$ . However in the tails of the  $\Delta\tau$  distribution, the model does not fit the data well. Some excess events are present that are not described by the model.

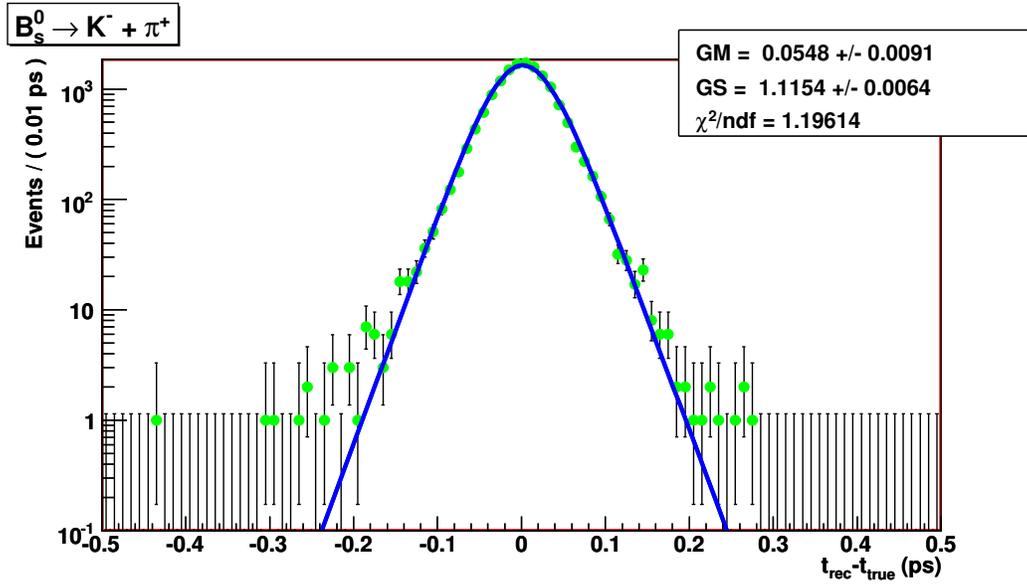
An explanation of the origin of the excess events with negative  $\Delta\tau$  values was proposed in [112], where a similar (but more significant) effect was seen in fits to the proper time



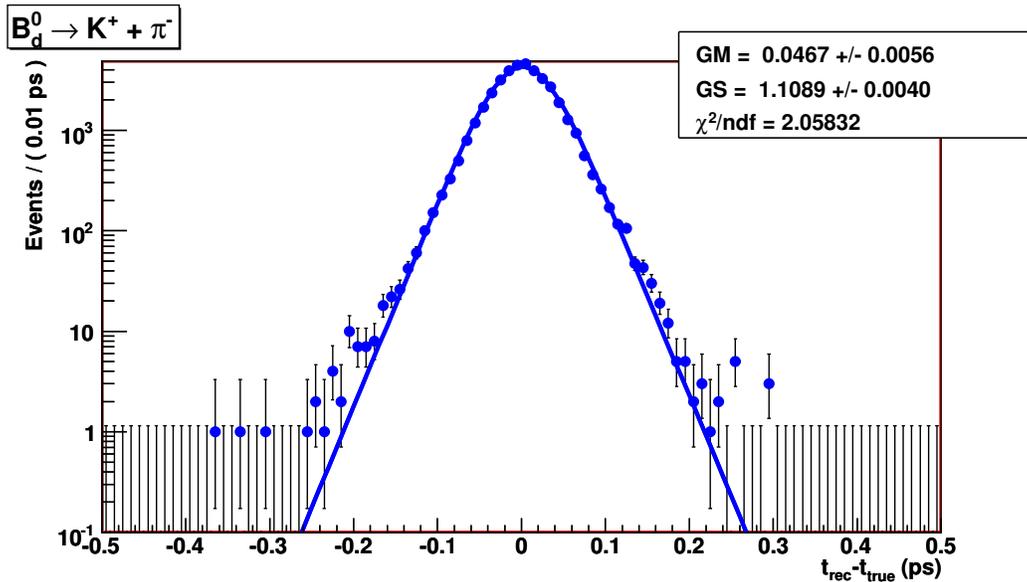
**Figure 4.6:** Resolution model fit to the proper time residual distribution for  $B_d \rightarrow \pi^+ \pi^-$  events.



**Figure 4.7:** Resolution model fit to the proper time residual distribution for  $B_s \rightarrow K^+ K^-$  events.



**Figure 4.8:** Resolution model fit to the proper time residual distribution for  $B_s \rightarrow K^- \pi^+$  events.



**Figure 4.9:** Resolution model fit to the proper time residual distribution for  $B_d \rightarrow K^+ \pi^-$  events.

residual distributions for the decays  $B_d \rightarrow J/\psi K^*$  and  $B^+ \rightarrow J/\psi K^+$ . This explanation claims that the excess events are due to a reconstruction bias effect that is seen in the simulation. This bias is due to events where the reconstructed primary vertex (PV) includes tracks from the decay of a  $B$  hadron, which could be the signal  $B$  or the other  $B$  in the event.<sup>4</sup> Such a reconstructed PV is “attracted” toward the  $B$  decay vertex, decreasing  $\tau_{rec}$ . This causes a small bias toward negative values of  $\Delta\tau$ . In [112] an attempt was made to remove this bias by removing from the PV any tracks that are reconstructed as decay products of a  $B$ , then refitting the PV and recalculating the proper time residual. The amount of excess events was found to decrease, but some remained. One possible origin of the remaining excess events is PVs that include tracks from the other  $B$  in the event, when the other  $B$  has not been identified. Hence the bias cannot be removed for these events. In any case the above mechanism can only cause excess events at negative  $\Delta\tau$ , so it cannot explain the excess events that are also seen at positive  $\Delta\tau$ .

Another possible explanation for the presence of outlying events with a very large residual is that there are issues with the reconstruction of these events. For example the momentum of one daughter particle can be poorly measured due to its track including a few hits from another particle (but not enough hits from another particle for the track to be classified as a ghost). This would cause a large momentum residual, which can translate into a large proper time residual.

It is also possible that a more complicated parameterisation of the dependencies of  $M$  and  $S$  on  $\sigma_{\tau_{rec}}$  would alter the model to take account of the excess events. In particular the  $M$  dependence may feature a quadratic or higher order component (see Fig. 4.5).

The disagreements between the model and the data in both of the regions described above can be removed (or at least greatly reduced) by adding additional terms, beyond the single Gaussian, to the resolution model  $R(\Delta\tau)$  [113]. However this introduces extra parameters into the model, and it was found that these extra parameters cannot be reliably extracted from a fit to the reconstructed proper time distribution. This shows that the reconstructed proper time distribution is sensitive to the gross features of the resolution model, but not to its fine details. So the small discrepancies between the model and the data are tolerated, because they only appear for a tiny fraction of the total events (so neglecting the discrepancies should not introduce significant biases), and also because removing the discrepancies requires a more complicated resolution model whose parameters cannot be extracted from data.

The fitted values for the resolution model parameters  $GM$  and  $GS$  from the fits to the four  $B \rightarrow h^+ h'^-$  decays are summarised in Table 4.3. Note that the lower statistics available

---

<sup>4</sup>The PV reconstruction necessarily takes place before any  $B$  hadrons can be identified, as one of the key characteristics used to identify  $B$  hadrons is their large flight distance from the PV.

Parameter	$B_d \rightarrow \pi^+\pi^-$	$B_s \rightarrow K^+K^-$	$B_s \rightarrow K^-\pi^+$	$B_d \rightarrow K^+\pi^-$
$GM$	$0.043 \pm 0.006$	$0.060 \pm 0.004$	$0.055 \pm 0.009$	$0.047 \pm 0.006$
$GS$	$1.117 \pm 0.004$	$1.115 \pm 0.003$	$1.115 \pm 0.006$	$1.109 \pm 0.004$
$\chi^2/\text{ndf of fit}$	1.90	2.34	1.19	2.05

**Table 4.3:** Comparison of resolution model parameters from fits to proper time residual distributions for  $B \rightarrow h^+h'^-$  decays.

for the  $B_s \rightarrow K^-\pi^+$  decay lead to the fit errors being higher there than for the other three decays. The lower statistics also lead to a lower  $\chi^2$  per degree of freedom for  $B_s \rightarrow K^-\pi^+$ . This is because none of the fits describe the distribution perfectly, and for the decays with more statistics the discrepancy between the data and the model becomes statistically more significant, increasing the  $\chi^2$  per degree of freedom.

The values of  $GS$  for all four decays are compatible with each other, within the fit errors. The fact that the values of  $GS$  is greater than unity shows that the lifetime fitter slightly underestimates the value of the per-event error  $\sigma_{\tau_{rec}}$ . The true proper time resolution for a typical  $B \rightarrow h^+h'^-$  event in LHCb with  $\sigma_{\tau_{rec}} \approx 35$  fs (see Fig 4.3) is therefore  $\sigma_\tau \approx GS \cdot 35\text{fs} \approx 39\text{fs}$ .

The values of  $GM$  in the  $B_s$  decays are slightly larger than in the  $B_d$  decays, however this is not an issue as the main aim of this study is to extract the resolution model parameters from  $B_s \rightarrow K^-\pi^+$  and apply them to  $B_s \rightarrow K^+K^-$ , and the  $GM$  values in these decays agree within the fit errors. The fact that  $GM$  values are non-zero shows that there is a bias in the measurement of the proper time. Since  $GM > 0$ , this bias is toward larger values of  $\tau_{rec}$ . The size of the bias is very small,  $\approx GM \cdot 35\text{fs} \approx 2\text{fs}$ . The origin of this bias is not yet understood.

The fact that the parameters for  $B_s \rightarrow K^-\pi^+$  and  $B_s \rightarrow K^+K^-$  agree with each other means that the parameters for  $B_s \rightarrow K^-\pi^+$  that are measured from data can be applied to  $B_s \rightarrow K^+K^-$ . The method used to find the parameter values for  $B_s \rightarrow K^-\pi^+$  from data will now be described.

### 4.3 Determining the Resolution Model Parameters from Data

The following sections will demonstrate that a fit to the flavour tagged  $\tau_{rec}$  distribution of the  $B_s \rightarrow K^-\pi^+$  decay can correctly recover the parameters of the proper time resolution model. The previous section has demonstrated that these parameters can be applied to  $B_s \rightarrow K^+K^-$ , which will allow the  $\mathcal{CP}$  asymmetries in that channel to be measured.

	Signal	Specific Background	Combinatorial Background
Resolution Model $GM$	0.05	0.0	0.0
Resolution Model $GS$	1.10	1.0	1.0
Average Lifetime	1.47 ps	1.53 ps	1.01ps
Mistag Rate ( $\omega_{\text{tag}}$ )	34%	37%	50%
Oscillation Frequency ( $\Delta m$ )	17.8 ps <sup>-1</sup>	0.507 ps <sup>-1</sup>	n/a
Width Difference ( $\Delta\Gamma$ )	-0.07 ps <sup>-1</sup>	0 ps <sup>-1</sup>	n/a
Bkg-to-Signal (B/S) Ratio	n/a	0.52	1.94

**Table 4.4:** Inputs for the proper time distribution of the signal and background components of the toy data.

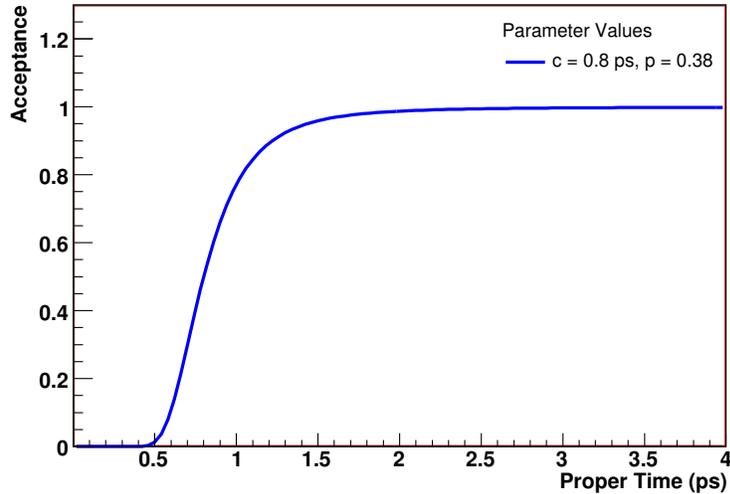
### 4.3.1 Construction of Toy Data $\tau_{rec}$ Distribution

To provide input for the fit studies, samples of simulated “toy data” were generated. The toy data consist of a given number of events which follow a Probability Density Function (PDF) which resembles as closely as possible the distribution of reconstructed proper time ( $\tau_{rec}$ ) that will be seen in data for selected and flavour tagged as  $B_s \rightarrow K^- \pi^+$  events. This distribution will have contributions from signal events, specific background events and combinatoric background events. The construction of the flavour tagged  $\tau_{rec}$  distributions for signal, specific background and combinatoric background will now be described in turn.

The first step in constructing the signal ( $B_s \rightarrow K^- \pi^+$ ) contribution to the flavour tagged  $\tau_{rec}$  PDF for the toy data is to choose the values for the parameters of the proper time resolution model. The values of  $GM$  and  $GS$  found in the fit to the  $\Delta\tau$  distribution for  $B_s \rightarrow K^- \pi^+$  are taken, and slightly rounded off for simplicity. The resulting resolution model is then used as one of the inputs to a PDF which is designed to describe the flavour tagged  $\tau_{rec}$  distribution, taking into account physics parameters and detector resolution<sup>5</sup>. The values used to construct the signal and background distributions are given in Table 4.4. The mistag values are taken from [75]. The values used for  $\tau_{B_{d,s}}$  and  $\Delta m_{d,s}$  are taken from the current world average experimental measurements [15]. The values used for  $\Delta\Gamma_{d,s}$  are the Standard Model expectations (see e.g. [43]). The choice of values for the remaining parameters will be discussed below.

A PDF,  $P(\sigma_{\tau_{rec}})$ , is now formed from the  $\sigma_{\tau_{rec}}$  distribution for  $B_s \rightarrow K^- \pi^+$ , using kernel estimation. The signal PDF described above is then made conditional on  $P(\sigma_{\tau_{rec}})$ . This step defines the proper time resolution model for the toy data, because its parameters

<sup>5</sup>The PDF object used is of the RooFit class RooBDecay.



**Figure 4.10:** Proper time acceptance function used in the toy data. See the text for details.

are functions of  $GM$ ,  $GS$  and  $P(\sigma_{\tau_{rec}})$ . The final step in constructing the signal PDF is to multiply this conditional PDF by a proper time acceptance function, to take account of the effect of the impact parameter cuts in the selection. The acceptance function used has the form [56]

$$\epsilon(\tau) = \frac{1}{2} \left( 1 - \text{Erf} \left( \frac{c - \tau}{p\tau} \right) \right), \quad (4.15)$$

where  $\text{Erf}(x)$  is the error function:

$$\text{Erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (4.16)$$

The acceptance function is 0 at  $\tau = 0$ , and rises to  $\approx 1$  for  $\tau \gg c$ . The parameters  $c$  and  $p$  are found from a fit to the  $\tau_{rec}$  distributions from the full simulation. These fits find values of  $c$  and  $p$  that are consistent across the  $B \rightarrow h^+h'^-$  decays. The values used here are the average values [56], which are  $c = 0.80$  ps and  $p = 0.38$ . The acceptance function with these values of  $c$  and  $p$  is shown in Fig. 4.10. The backgrounds should have very similar acceptances to the signal as they have to pass the same impact parameter cuts. So the same acceptance function is used here for the signal and background distributions.

To make the total toy data PDF, the contributions from specific and combinatoric background now need to be added to the signal contribution. The composition and relative sizes used for the backgrounds to  $B_s \rightarrow K^- \pi^+$  are taken from the selection study in [118]<sup>6</sup>. It

<sup>6</sup>Since this study was carried out, an updated selection study has been published [56]. However the signal and background yields found in the two studies are very similar to each other.

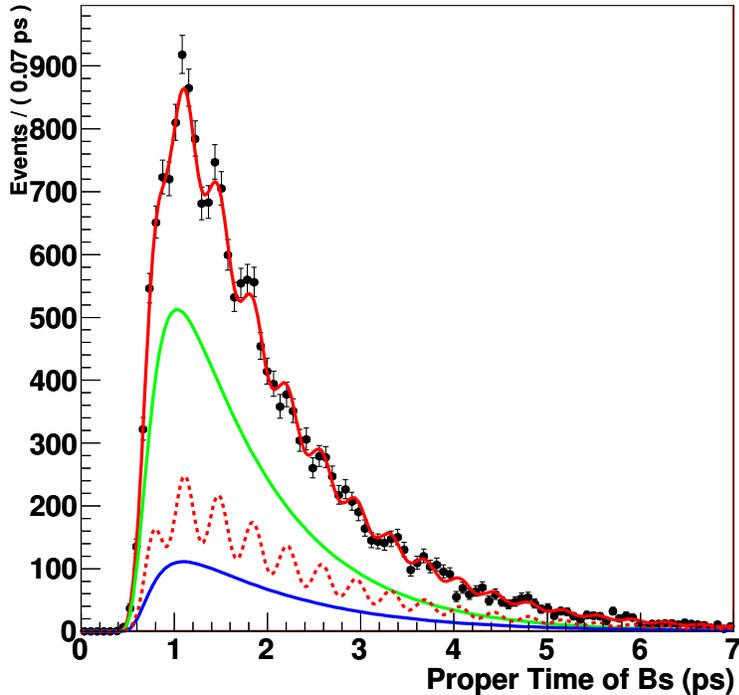
is assumed that only  $B$  events can pass the selection. This is because any short-lived background should be removed by the cuts which are made on the impact parameter significance of the final state particles and flight distance significance of the  $B$ . This assumption can be tested during early data taking, when data will be taken with a random trigger to allow the nature of minimum bias events to be studied in detail. This leaves two types of background to consider — specific background and inclusive combinatoric background (see Sec. 4.2.3).

The specific backgrounds arise from other  $B \rightarrow h^+h'^-$  decays, as any partially reconstructed  $B \rightarrow hhh$  decays (see Sec. 4.2.3) will have masses which fall below the  $B_s$  mass. The dominant  $B \rightarrow h^+h'^-$  background for  $B_s \rightarrow K^-\pi^+$  is the  $B_d \rightarrow K^-\pi^+$  decay, because this background cannot be reduced by cutting on the particle identification (PID) of the detected hadrons, since the background and signal final states are identical. The only discriminating variable available is the invariant mass of the  $h^+h'^-$  pair. The expected background to signal ratio from specific background has been calculated to be 0.5 [118]. The level of contamination is significant because the expected yield of selected  $B_d \rightarrow K^-\pi^+$  events is more than 10 times the expected yield of  $B_s \rightarrow K^-\pi^+$  events (135,000 versus 9,800 in  $2 \text{ fb}^{-1}$ ). This factor arises due to both the higher probability for a  $b$  quark to hadronise into a  $B_d$  meson rather than a  $B_s$  meson, and the higher branching ratio for  $B_d \rightarrow K^-\pi^+$  compared to  $B_s \rightarrow K^-\pi^+$ .

Since the backgrounds are insensitive to the details of the resolution model, a “default” resolution model, with  $GM = 0$  and  $GS = 1$ , is used. Then a distribution similar to the signal distribution, but with different physics parameters (see Table 4.4), is created.

As well as specific background, there will also be combinatoric background. The expected background to signal ratio from combinatoric background has been calculated to be 1.9 [118]. This value is higher than for the other  $B \rightarrow h^+h'^-$  channels, due to the lower yield of  $B_s \rightarrow K^-\pi^+$  events.

The proper time distribution for combinatoric background from  $B$  decays should not be too different from that for the  $B \rightarrow h^+h'^-$  channels, because their proper time acceptance behaviour should be similar as discussed above, and the combinatoric background will show an exponential decay (but not necessarily with the same lifetime as the signal proper time distribution). The proper time distribution for combinatoric events passing the selection has been fitted [118]. The inputs for combinatoric background in Table 4.4 reflect the result of this fit. In data there will be a large and pure sample of combinatoric events in the upper sidebands of the invariant mass distribution (which are free of partially reconstructed background events). These events can be used to precisely determine the proper time distribution for combinatoric background. An alternative method, which measures the proper time distribution for combinatoric background to  $B \rightarrow h^+h'^-$  decays directly in the signal mass region



**Figure 4.11:** Flavour tagged  $\tau_{rec}$  distribution for an example toy dataset (black points). The total number of events is 20,300. The solid red line shows the total PDF (from which the toy data is generated). The broken red line, blue line and green line show the contributions to the total PDF from  $B_s \rightarrow K^- \pi^+$  signal events, specific background events and combinatoric background events respectively.

using background subtraction, has recently been developed [119].

The total PDF for the toy data is now formed by adding the signal PDF and the two background PDFs together, with the background-to-signal (B/S) ratios fixed to values close to those found in [118]. The toy data are generated following this total PDF.

An example toy dataset is shown in Fig. 4.11. This dataset contains 20,300 events, 6,000 of which are signal events. This corresponds to what is expected after  $2 \text{ fb}^{-1}$  of LHCb data taking [118].

### 4.3.2 Construction of $\tau_{rec}$ Distribution for Fit

The  $\tau_{rec}$  distribution for the PDF that will be used to fit the toy data distribution described in the previous section will have floating values for the resolution model parameters, which should converge during the fit to the values used in the toy data generation.

For the signal contribution to the fit PDF, the form of the resolution model is the same as

Parameter	Seed Interval	Float Interval
$GM$	[-2.0,2.0]	[-100,100]
$GS$	[0.5,2.0]	[0.001,100]

**Table 4.5:** Intervals used to generate the fit PDF and perform the fit.

in the toy data, but the parameters are seeded using a value uniformly chosen from some seed interval for each parameter. During the fit  $GM$  and  $GS$  are allowed to float in intervals larger than the seed intervals. The seed and float intervals for each parameter are given in Table 4.5. These seed intervals aim to represent a reasonable range of values for the parameter to take. The float intervals should be as large as possible while still being physical (e.g. the width cannot be negative).

This resolution model is then used as an input to a flavour tagged  $\tau_{rec}$  distribution with the same physics parameters and mistag as the toy data, but with the parameters of the resolution model floating rather than being fixed. The values of the mistag rate and the physics parameters are assumed to have been already well measured in various external analyses. The mistag rate can be measured as described in Sec. 4.1.3.2, and the  $B$  meson lifetimes (including the width differences  $\Delta\Gamma_q$ ) will be measured by studying untagged  $\tau_{rec}$  distributions for both high-yield decays such as the  $B \rightarrow D\pi$  channels [120] and for  $B \rightarrow h^+h'^-$  decays [119].

As with the toy data, the signal PDF is made conditional on  $P(\sigma_{\tau_{rec}})$ . This allows the PDF to be sensitive to the values of  $GM$  and  $GS$ . Finally the signal PDF is multiplied by the same proper time acceptance function that was used in the toy data. The acceptance function will be measured on data as part of the lifetime measurement analyses mentioned above. Although those analyses relate to untagged  $\tau_{rec}$  distributions, the acceptance function for tagged events (which is in any case expected to be very similar to that for untagged events) can be measured on data by applying the same analysis to a sample of tagged events where both tags are considered together, so that no oscillations are present in the distribution. As the acceptance function should not depend on the flavour of the tag ( $B$  or  $\bar{B}$ ), this will give the acceptance function to be used in a tagged analysis such as the current one.

The specific and combinatoric background contributions to the fit PDF are the same as those used in the toy data. The physics parameters and mistag for the backgrounds can be determined on data using the same methods as described above for the signal. The default proper time resolution model used for the backgrounds in the toy data is re-used for the fit PDF, since the backgrounds are not sensitive to the resolution model parameters.

The total fit PDF is then made by summing the signal and background contributions. The background-to-signal (B/S) ratios used are the same as in the toy data, as these ratios can be

accurately determined by a fit to the invariant mass distribution for tagged events.

The fit is an unbinned log-likelihood fit, carried out using the RooFit fitting package [121], which operates within the ROOT data analysis framework [122]. The maximisation of the log-likelihood is carried out by the Minuit package [123]. The results from fitting the PDF described above to the toy data to extract the resolution model parameters will now be described.

### 4.3.3 Results from Fit to Toy Data

#### 4.3.3.1 Setup for Toy Monte Carlo Simulation Study

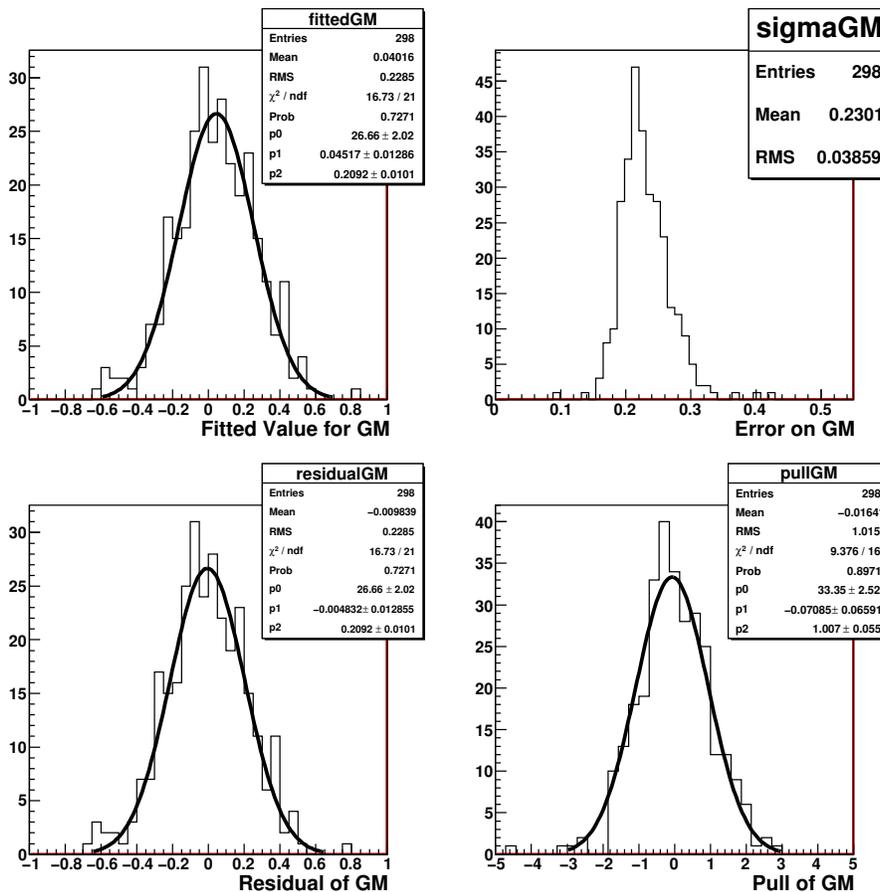
A number of toy datasets are created, in order to build up distributions of the fitted parameters to check that the values returned are unbiased and have correctly estimated errors. The toy data PDF is the same for each toy dataset, but the generation of the toy data points from this PDF is done using a different seed each time. For the fit PDF the background PDFs are the same each time, while the signal PDF has different resolution model parameters each time, uniformly drawn from the seed interval (see previous section).

Since the number of selected  $B_s \rightarrow K^- \pi^+$  events per nominal year of LHCb running (equivalent to  $2 \text{ fb}^{-1}$  integrated luminosity) is expected to be  $\simeq 10,000$  [118], and the tagging efficiency is around 60% [75], around 6,000 tagged  $B_s \rightarrow K^- \pi^+$  events are expected per nominal year. It follows that around 30,000 tagged  $B_s \rightarrow K^- \pi^+$  events are expected after  $10 \text{ fb}^{-1}$  (the total integrated luminosity that will be collected during the data collection period of the current LHCb detector).

As the purpose of extracting the resolution model parameters is to use them as input to the fit for  $\gamma$ , a fit which requires a considerable amount of data to produce a significant result, it is not worthwhile to use the method described here on very few events. Two separate sets of toy experiments are run - one corresponding to  $2 \text{ fb}^{-1}$  of data, and one corresponding to  $10 \text{ fb}^{-1}$  of data.

At each integrated luminosity 300 sets of toy data are generated, and each set is fitted using a PDF with different seeds as described above. For each of the fit parameters,  $GM$  and  $GS$ , the distributions for the following quantities are plotted:

- The value of the parameter as returned by the fit.
- The error assigned by the fit to this value.
- The residual of the parameter, i.e. the fitted value minus the value in the toy data.
- The pull of the parameter, i.e. the residual divided by the error.



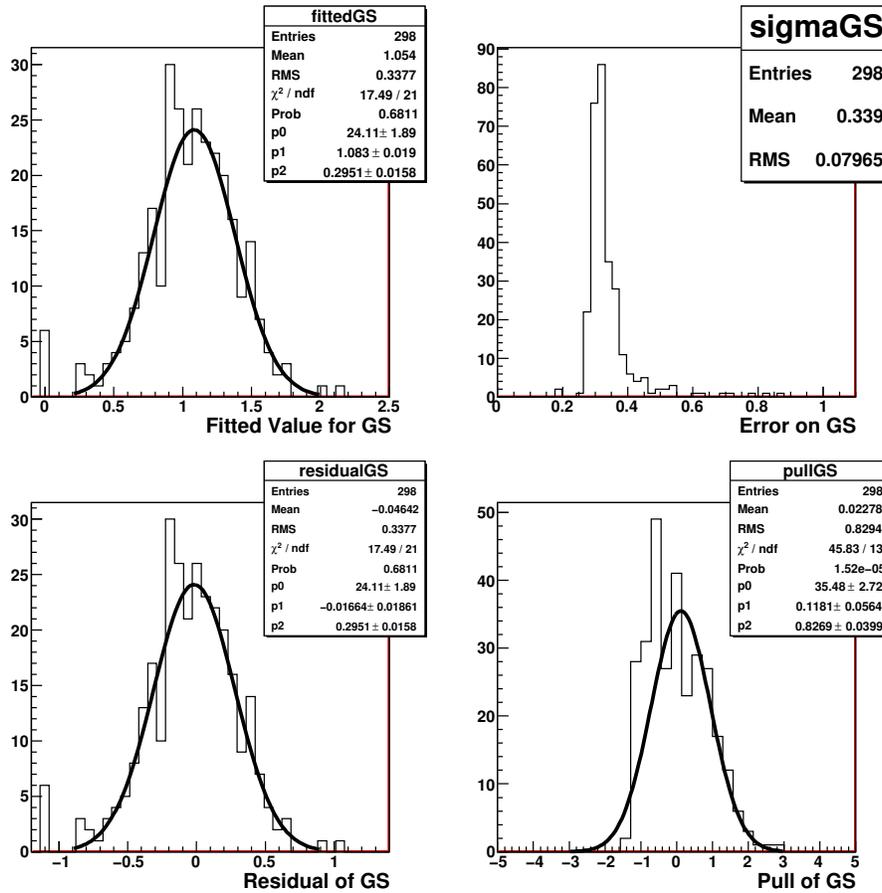
**Figure 4.12:** Fit results for the parameter  $GM$  with  $2 \text{ fb}^{-1}$  of data. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

To monitor how the fit has performed, single Gaussians are fitted to the fitted values, residuals and pull distributions of each parameter. The reliability of the fit result can also be monitored by studying the correlations between the different variables in the fit. For each variable, the minimisation package Minuit calculates a “global correlation”, which combines the correlations of that variable with each other variable in the fit to estimate the overall independence of that variable in the fit. In all plots, only fits that returned a full and accurate covariance matrix are shown.

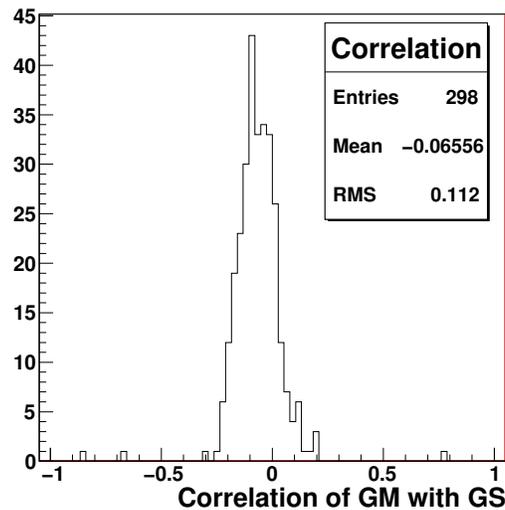
#### 4.3.3.2 Results With $2 \text{ fb}^{-1}$ of Data

From the 300 fits to the  $2 \text{ fb}^{-1}$  datasets, almost all (298) returned a full and accurate covariance matrix. The fit results for  $GM$  and  $GS$  are shown in Figs. 4.12 and 4.13, and the distribution of the correlation coefficient between  $GM$  and  $GS$  is shown in Fig. 4.14.

It can be seen that the distributions of the fitted values are both approximately Gaussian, and there are no significant biases on the fitted values. However there are six events where



**Figure 4.13:** Fit results for the parameter GS with  $2 \text{ fb}^{-1}$  of data. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.



**Figure 4.14:** Correlation coefficient between GM and GS in the fits with  $2 \text{ fb}^{-1}$  of data.

$GS$  has been fitted to be very close to zero. These fits have very large errors on the  $GS$  value (between 5 and 10), but have normal values for  $GM$  and its error. The presence of these events indicates that there is a small probability ( $\simeq 6/300 = 2\%$ ) of the fit converging to a local minimum at  $GS=0$ . Should the fit to data happen to converge to this local minimum, it would be identified as such, as  $GS=0$  is not physical (it would imply perfect proper time resolution). Careful steps would then have to be taken to coax the fit away from the local minimum and make it return a sensible result.

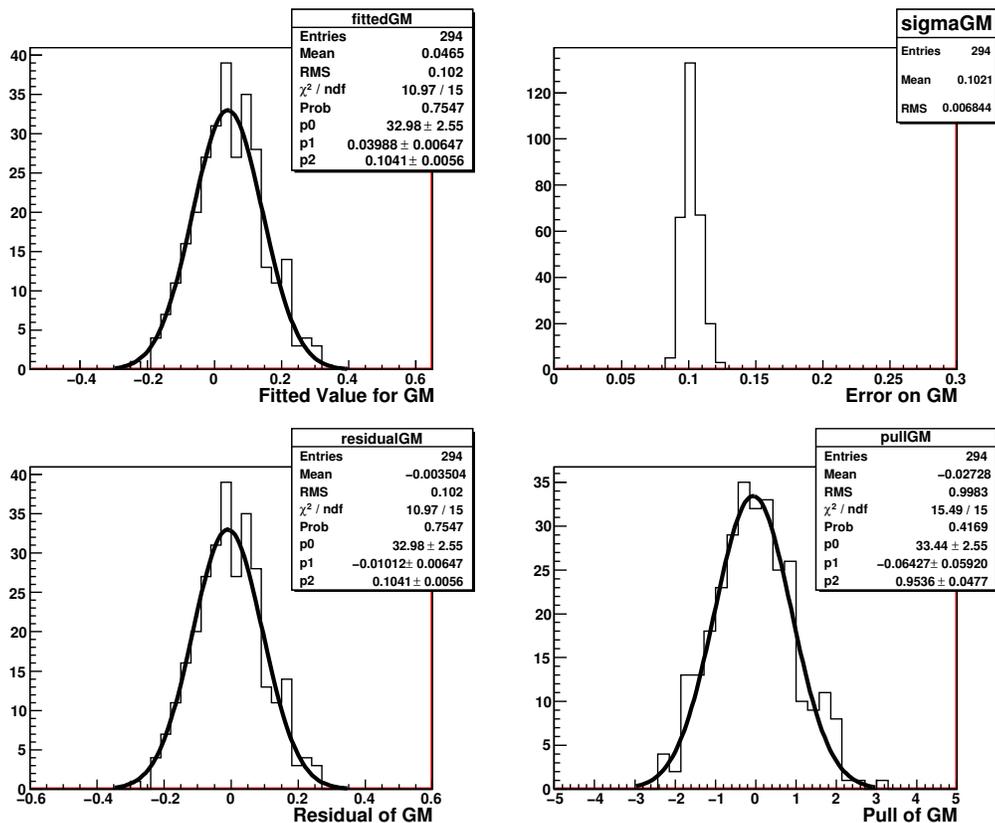
The pull width for  $GM$  is compatible with 1, indicating that the error on  $GM$  is correctly estimated. However the  $GS$  pull distribution shows a small but significant bias, with mean  $0.12 \pm 0.06$ , and a width of  $0.83 \pm 0.04$ , significantly below 1. Part of the reason for this is that the events at  $GS=0$  are causing some excess events to appear close to zero on the negative side of the pull distribution, since their pull values are around  $(0-1.2)/8 = -0.15$ . In general, the fits with lower fitted values for  $GS$  have larger errors on  $GS$  than the rest of the fits, which is reflected in the presence of a tail in the distribution of fitted errors. This tendency for larger errors for small  $GS$  values causes a deficit of fits with pull values in  $[-3.0, 1.5]$ , which distorts the fit. Refitting the  $GS$  pull distribution with the six fits at  $GS=0$  excluded and fitting only to the region  $[-1.5, 3.0]$  causes the pull width to increase to  $0.88 \pm 0.05$ , and decreases the bias to  $0.08 \pm 0.06$ . The dependence of the size of the  $GS$  fit error on the  $GS$  fitted value explains why its pull distribution shows a small bias, while its fitted values distribution does not.

The correlation between  $GM$  and  $GS$  in the fit is small, with the magnitude of the correlation being less than 0.2 for almost all of the fits. This demonstrates that  $GM$  and  $GS$  are fairly independent in the fit, so there is no problem in fitting for both of them simultaneously.

#### 4.3.3.3 Results With $10 \text{ fb}^{-1}$ of Data

The above study, corresponding to  $2 \text{ fb}^{-1}$  of data, was repeated with toy datasets corresponding to  $10 \text{ fb}^{-1}$  of data. From the 300 fits to the  $10 \text{ fb}^{-1}$  datasets, 294 returned a full and accurate covariance matrix. This is compatible with the number of successful fits for the  $2 \text{ fb}^{-1}$  case. The fit results for  $GM$  and  $GS$  using the  $10 \text{ fb}^{-1}$  datasets are shown in Figs. 4.15 and 4.16, and the distribution of the correlation coefficient between  $GM$  and  $GS$  is shown in Fig. 4.17.

As with the  $2 \text{ fb}^{-1}$  fits, the distributions of the fitted values are approximately Gaussian, and the means of the fitted values are unbiased. The key change is that the fitted errors (and width of the fitted values distributions) are now smaller by factors of around  $\sqrt{5} \approx 2.2$ , which is the expected error reduction from increasing the dataset from  $2 \text{ fb}^{-1}$  to  $10 \text{ fb}^{-1}$ . The smaller error value means that the  $GS$  parameter never approaches zero when it is



**Figure 4.15:** Fit results for the parameter  $GM$  with  $10 \text{ fb}^{-1}$  of data. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

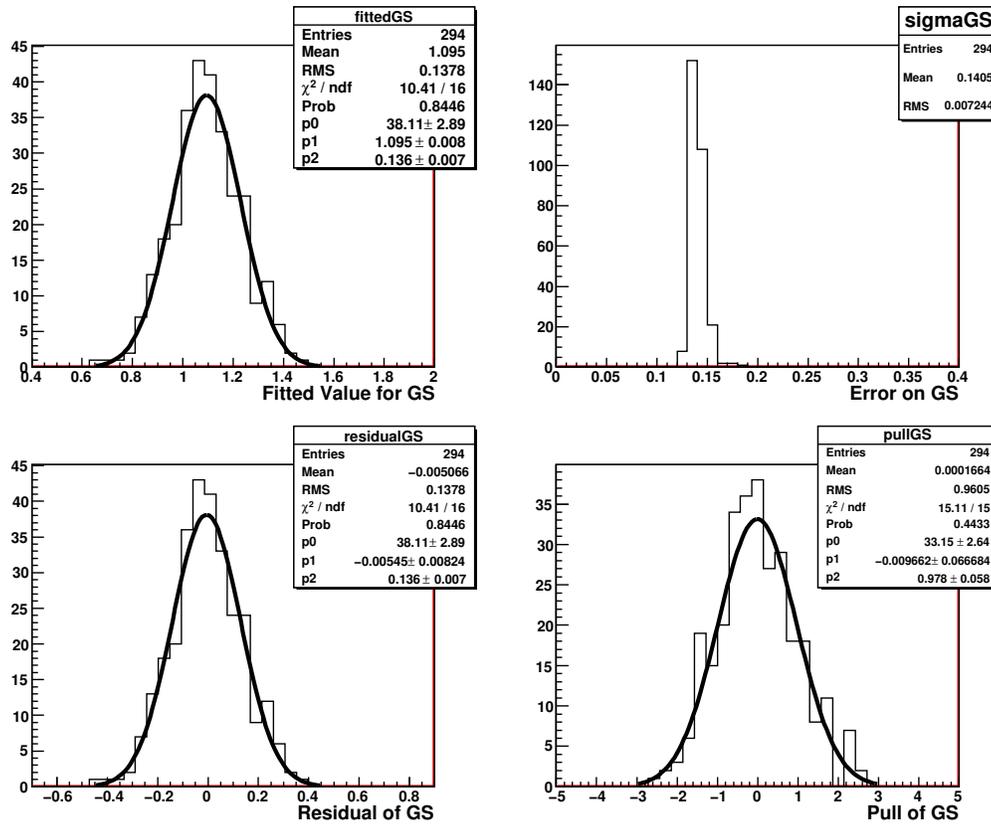
floating during the fit, hence no fits are returned with the local minimum result  $GS=0$ . Without such events, the pull distribution for  $GS$  is now Gaussian, with unbiased mean and width compatible with 1. Also there are no longer any outliers in the fitted error distribution for  $GS$ . The  $GM$  pull distribution is well-behaved, as it was for the  $2 \text{ fb}^{-1}$  fits. The correlation between  $GM$  and  $GS$  in the  $10 \text{ fb}^{-1}$  fits is similar to the  $2 \text{ fb}^{-1}$ , except that the spread of values is smaller, due to the higher statistics involved.

The results for the fits to data with the different dataset sizes are summarised in Table 4.6. The results shown are the mean and width of the Gaussians fitted to the distribution of the fitted values and the pulls distribution of each parameter.

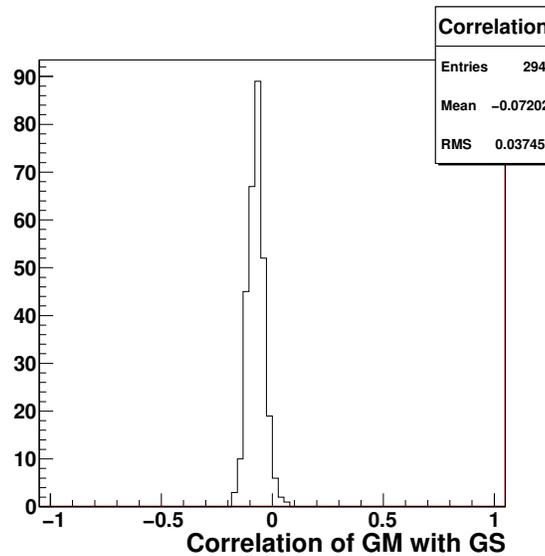
#### 4.3.3.4 Sources of Systematic Error

There are several sources of systematic error in this analysis, mainly related to uncertainties in the values of the physics parameters and experimental quantities that are fixed in the fit. If the fixed values assumed in the fit are incorrect, this will introduce a systematic error.

The most significant source of systematic error is likely to be an incorrect determination



**Figure 4.16:** Fit results for the parameter GS with  $10 \text{ fb}^{-1}$  of data. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.



**Figure 4.17:** Correlation coefficient between GM and GS in the fit with  $10 \text{ fb}^{-1}$  of data.

	Variable	Value with $2 \text{ fb}^{-1}$	Value with $10 \text{ fb}^{-1}$	Input
$GM$	Fitted values mean	$0.05 \pm 0.01$	$0.04 \pm 0.01$	0.05
	Fitted values width	$0.21 \pm 0.01$	$0.10 \pm 0.01$	n/a
	Pull mean	$-0.07 \pm 0.07$	$-0.06 \pm 0.06$	n/a
	Pull width	$1.01 \pm 0.06$	$0.95 \pm 0.05$	n/a
$GS$	Fitted values mean	$1.08 \pm 0.02$	$1.10 \pm 0.01$	1.10
	Fitted values width	$0.30 \pm 0.02$	$0.14 \pm 0.01$	n/a
	Pull mean	$0.12 \pm 0.06$	$-0.01 \pm 0.07$	n/a
	Pull width	$0.83 \pm 0.04$	$0.98 \pm 0.06$	n/a
	Fits with $GS=0$	6	0	n/a

**Table 4.6:** Comparison of fit results with  $2 \text{ fb}^{-1}$  of data and  $10 \text{ fb}^{-1}$  of data.

of the mistag rate  $\omega_{\text{tag}}$ , as for a  $B_s$  decay, changes in  $\omega_{\text{tag}}$  affect the oscillations in the flavour tagged  $\tau_{\text{rec}}$  distribution in exactly the same way as changes in the proper time resolution do (see Fig. 4.2). To evaluate the size of the systematic error associated with the mistag determination, the studies of the previous two sections were repeated, but with the mistag rate for the signal in the toy data set to a different value from the mistag rate used in the fit PDF (recall from Table 4.4 that the toy data and fit PDF in the studies discussed above both used  $\omega_{\text{tag}}=34\%$  for the signal). As the expected precision on the mistag rates is about 1% [75], this was chosen as the difference between the two rates. Fits were made with  $\omega_{\text{tag}}=35\%$  in the toy data and  $\omega_{\text{tag}}=34\%$  in the fit PDF, and also with  $\omega_{\text{tag}}=34\%$  in the toy data and  $\omega_{\text{tag}}=35\%$  in the fit PDF. As before, 300 fits were performed each time, with datasets corresponding to  $2 \text{ fb}^{-1}$  of data, and again with  $10 \text{ fb}^{-1}$  of data. As with the previous fits, all but a handful of the 300 fits return good covariance matrices. The results are summarised in Table 4.7.

From Table 4.7 it can be seen that the  $GM$  parameter has not been affected — it is still recovered without bias and with correctly estimated errors. On the other hand, the  $GS$  parameter is subject to a clear bias, although the size of the errors on  $GS$  do not change significantly. If the mistag value used in the fit is smaller than the “true” value in the toy data, the fit compensates for the increased dilution of the oscillations by returning a higher value for  $GS$ , implying a poorer proper time resolution. The opposite is true when the mistag value in the fit is larger than the true value. In both cases the size of the bias is the same,  $\simeq 0.13$ . The bias does not depend on the size of data sample, as is expected for a systematic error.

Note also that the lower fitted values for  $GS$  seen when the mistag rate is overestimated caused more fits to fall into the local minimum at  $GS=0$ , with a correspondingly worse

	Variable	$\omega_{\text{tag}}(\text{toy})=35\%, \omega_{\text{tag}}(\text{fit})=34\%$		$\omega_{\text{tag}}(\text{toy})=34\%, \omega_{\text{tag}}(\text{fit})=35\%$		Input
		$2 \text{ fb}^{-1}$	$10 \text{ fb}^{-1}$	$2 \text{ fb}^{-1}$	$10 \text{ fb}^{-1}$	
$GM$	Fitted values mean	$0.05 \pm 0.01$	$0.04 \pm 0.01$	$0.04 \pm 0.01$	$0.05 \pm 0.01$	0.05
	Fitted values width	$0.24 \pm 0.01$	$0.10 \pm 0.01$	$0.21 \pm 0.01$	$0.10 \pm 0.01$	n/a
	Pull mean	$0.06 \pm 0.06$	$-0.11 \pm 0.06$	$-0.06 \pm 0.06$	$0.01 \pm 0.06$	n/a
	Pull width	$0.97 \pm 0.04$	$0.97 \pm 0.05$	$1.00 \pm 0.05$	$0.98 \pm 0.05$	n/a
$GS$	Fitted values mean	$1.22 \pm 0.02$	$1.23 \pm 0.01$	$0.95 \pm 0.02$	$0.96 \pm 0.01$	1.10
	Fitted values width	$0.33 \pm 0.02$	$0.14 \pm 0.01$	$0.31 \pm 0.02$	$0.15 \pm 0.01$	n/a
	Pull mean	$0.48 \pm 0.06$	$0.97 \pm 0.07$	$-0.21 \pm 0.06$	$-0.93 \pm 0.06$	n/a
	Pull width	$0.95 \pm 0.04$	$1.06 \pm 0.06$	$0.73 \pm 0.04$	$0.90 \pm 0.05$	n/a
	Fits with $GS=0$	0	0	15	1	n/a

**Table 4.7:** Comparison of fit results with incorrect mistag value assumed in the fit PDF.

distortion of the pull distribution for  $GS$ . In contrast, the larger  $GS$  fitted values seen when the mistag rate is overestimated reduce the number of fits in the local minimum, leading to better pull distributions.

#### 4.3.3.5 Summary of Toy Data Results

It has been shown that the parameters of the resolution model are recovered correctly from a fit to data. For both dataset sizes, the pull distributions for the parameter  $GM$  are well-behaved. For the  $GS$  parameter in the fits with  $2 \text{ fb}^{-1}$  of data, there is a correlation between the fitted value and the error that causes the pull distribution to be slightly non-Gaussian, with a small bias in the mean from 0 and a width less than 1. However the size of the bias is small, only 10% of the size of the statistical error. There is also a  $\simeq 2\%$  probability for the fit to converge to a local minimum with  $GS=0$ .

For the fits with  $10 \text{ fb}^{-1}$  of data, the errors on  $GS$  are small enough that the fitted values never approach zero, and there is no significant correlation with the  $GS$  value and its error. This results in well-behaved pull distributions for  $GS$ .

The systematic error on  $GS$  arising from the uncertainty on the value of the mistag rate for the signal, which is expected to be the dominant systematic, has been estimated as  $\simeq 0.13$ . Combining the statistical uncertainties from Secs 4.3.3.2 and 4.3.3.3 with this systematic gives the overall sensitivity to the resolution model parameters ( $\oplus$  represents addition in quadrature):

$$\begin{aligned} \sigma_{GM} &= 0.21(\text{stat}), & \sigma_{GS} &= 0.30(\text{stat}) \oplus 0.13(\text{syst}) = 0.33, & (2 \text{ fb}^{-1}) \\ \sigma_{GM} &= 0.10(\text{stat}), & \sigma_{GS} &= 0.14(\text{stat}) \oplus 0.13(\text{syst}) = 0.19. & (10 \text{ fb}^{-1}) \end{aligned}$$

So the determination of the resolution model parameters will not become systematics-limited during the initial LHC data-taking period. A full study of the systematic effects from the proper time resolution model on the measurement of  $\gamma$  using  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^-K^+$  has yet to be made<sup>7</sup>. However it is expected that, given the magnitude of the errors, the fit to the  $B_s \rightarrow K^-\pi^+$  proper time distribution will start to give useful input to the study of the proper time resolution model for  $B_s \rightarrow K^-K^+$  (and hence to the measurement of  $\gamma$ ) with around  $2 \text{ fb}^{-1}$  of data. By the time  $10 \text{ fb}^{-1}$  of data has been collected, the fit will be able to provide important constraints to the study of  $B_s \rightarrow K^-K^+$ .

## 4.4 Conclusions

The proper time resolution for the decay  $B_s \rightarrow K^-K^+$  has to be measured in order to correctly extract the  $\mathcal{CP}$  asymmetries for that decay, which are needed for the measurement of  $\gamma$  using  $B \rightarrow h^+h'^-$  decays. This can be done on data by measuring the proper time resolution for the flavour-specific decay  $B_s \rightarrow K^-\pi^+$ , via a fit to its flavour-tagged reconstructed proper time distribution.

A proper time resolution model has been developed that accurately describes the proper time residual ( $\Delta\tau$ ) distributions in Monte Carlo simulated data for four  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  channels:

$$R(\Delta\tau) = N e^{-\frac{1}{2}\left(\frac{\Delta\tau-M}{S}\right)^2}, \quad (4.17)$$

with the mean  $M$  and width  $S$  being determined event-by-event according to

$$\begin{aligned} M &= GM \cdot \sigma_{\tau_{rec}}, \\ S &= GS \cdot \sigma_{\tau_{rec}}, \end{aligned} \quad (4.18)$$

where  $\sigma_{\tau_{rec}}$  is the per-event error on the reconstructed proper time. The two parameters of the model,  $GM$  and  $GS$ , are found to be compatible with each other for all four decays.

The parameters obtained from the fits to the full Monte Carlo residuals have been used to guide a toy Monte Carlo study. This study, which requires that certain properties of the data, such as the mistag rate and the  $B$  meson lifetime, have already been well measured by other methods, demonstrated that the parameters of the resolution model for  $B_s \rightarrow K^-\pi^+$  can be determined from the fit to data, using information from the per-event proper time error. The statistical uncertainty on the parameters of the model with  $2 \text{ fb}^{-1}$  of data is found to be  $\sigma_{GM(\text{stat})} = 0.21$  and  $\sigma_{GS(\text{stat})} = 0.30$ . With  $10 \text{ fb}^{-1}$  of data the statistical errors are lower by

---

<sup>7</sup>However a study into the proper time resolution systematics for the measurement of  $\delta\gamma$  using  $B_s \rightarrow J/\psi\phi$  has been made [51]. Although no significant bias on  $\delta\gamma$  was found, the extraction of the mistag rate was significantly biased if an incorrect resolution model was assumed.

the expected factor of  $\simeq 2.2$ . The dominant systematic uncertainty is expected to arise from an incorrect determination of the mistag rate. This uncertainty affects  $GS$  but not  $GM$ , and has been estimated to be  $\sigma_{GS(\text{syst})} = 0.13$ , which is similar in size to the statistical uncertainty on  $GS$  after  $10 \text{ fb}^{-1}$  of data.

With one nominal year of LHCb data, the fit can be expected to constrain the resolution model for  $B_s \rightarrow K^- K^+$  sufficiently to provide useful input for the extraction of its time-dependent  $\mathcal{CP}$  asymmetries, and hence for the measurement of  $\gamma$ .

# Chapter 5

## Two-Body Charmless Baryonic $B$ Decays

This chapter presents a study into the feasibility of discovering new baryonic decay modes of  $B$  mesons at LHCb, with particular focus on the  $B_d \rightarrow p\bar{p}$  mode. This study is also reported on in [124]. Section 5.1 summarises the theoretical predictions for the SM branching ratios of relevant two-body charmless baryonic  $B$  decays arising from different calculational models, and also gives the current best experimental upper limits on these branching ratios. Section 5.2 describes the adaption of the inclusive  $B \rightarrow h^+h'^-$  selection described in Sec. 4.2.3 to make it suitable for selecting  $B_d \rightarrow p\bar{p}$  events. Section 5.3 discusses the possible sources of background, and how they are suppressed by the selection. Section 5.4 gives the performance of the selection, along with the resulting background-to-signal ratios and signal significance as a function of the  $B_d \rightarrow p\bar{p}$  branching fraction. Section 5.5 lays out the strategy for measuring the  $B_d \rightarrow p\bar{p}$  branching ratio once a clear signal has been seen. Trigger mass window issues for  $B_d \rightarrow p\bar{p}$  are considered in Sec. 5.6. Section 5.7 briefly discusses the prospects for the observation of other two-body charmless baryonic  $B$  decays. Conclusions are given in Sec. 5.8.

### 5.1 Current Status of Theory and Experiment

Observations of  $B$  mesons decaying into two light (charmless) mesons have been made in several modes. However, despite searches with the CLEO detector and at the SLAC and KEK B-factories, no charmless two-body baryonic  $B$  decay has yet been observed [125, 126, 127]. This situation is in contrast with the observation of a multitude of three-body baryonic  $B$  decays. Empirically, the suppression of two-body compared to three-body baryonic  $B$  decays

is well established; well-known examples are (B.R. stands for the branching ratio):

$$\begin{aligned} \text{B.R.}(B^+ \rightarrow p\bar{p}K^+) &\gg \text{B.R.}(B_d \rightarrow p\bar{p}), \\ \text{B.R.}(B^+ \rightarrow \Lambda_c^- p\pi^+) &\gg \text{B.R.}(B_d \rightarrow \Lambda_c^- p). \end{aligned} \quad (5.1)$$

Assuming a suppression factor in charmless decays (first example) similar to that observed in  $b \rightarrow c$  transitions (second example), of order 0.1, the simplest two-body baryonic  $B$  decay,  $B_d \rightarrow p\bar{p}$ , may be expected to have a branching ratio of order  $10^{-7}$ , and therefore be observable at LHCb.

Theoretical predictions of the branching ratios for such two-body baryonic decays within the Standard Model (SM) vary depending on the method of calculation used. However, the predicted branching ratios are normally of order  $10^{-6}$  or lower. The LHCb experiment is very well placed to make an observation of these rare decays, thanks to its excellent vertexing and particle identification (PID) abilities, which enable background to be suppressed while efficiently retaining signal events.

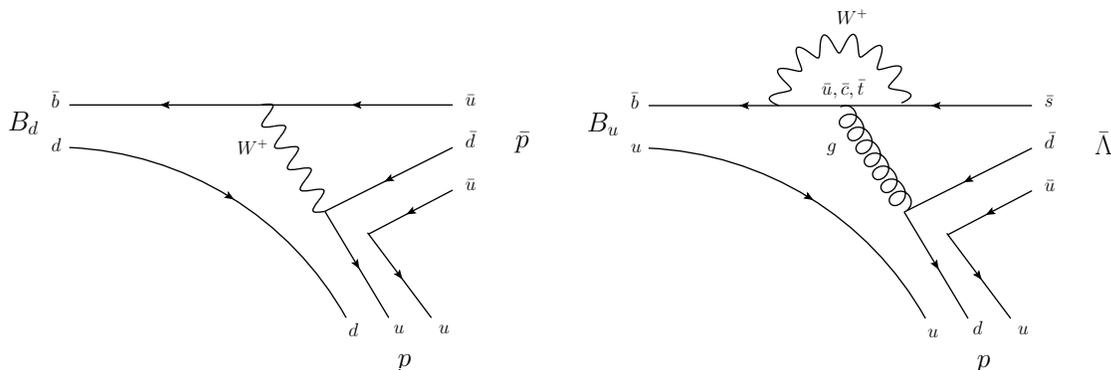
This study will concentrate on the particular channel  $B_d \rightarrow p\bar{p}$ , which as a  $B \rightarrow h^+h'^-$  channel is topologically identical to decays such as  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^+K^-$ , which have been the subject of much study within LHCb.

### 5.1.1 Theoretical Predictions for Branching Ratios

A theoretical calculation of the branching ratio for  $B$  meson decays involving baryons is challenging: the final state involves six quarks (rather than four as in a decay to two mesons), and the amplitudes typically require the production of a quark-antiquark pair out of the vacuum. This involves recoil effects which are not yet fully understood and hence are difficult to account for. The dominant decay amplitudes (see below) are non-factorisable, and hence more difficult to evaluate.

Most two-body charmless baryonic  $B$  decays are dominated by a single amplitude, which is either a tree amplitude featuring a  $b \rightarrow u$  transition, or a gluonic penguin amplitude featuring a  $b \rightarrow s$  or  $b \rightarrow d$  transition. Other amplitudes such as  $W$  annihilation,  $W$  exchange, penguin annihilation and electroweak penguin processes can also occur, but their contribution is expected to be small [128]. Figure 5.1 shows an example of a tree-dominated decay,  $B_d \rightarrow p\bar{p}$ , and an example of a penguin-dominated decay,  $B^+ \rightarrow p\bar{\Lambda}$ . The diagrams are similar to the corresponding diagrams for  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  decays (see Fig. 1.4), except that an additional  $q\bar{q}$  pair is created out of the vacuum.

Predictions have been published using several different approaches, including the use of QCD sum rules [129], the diquark model [130] and the pole model [131, 132, 128]. These models differ in how they break down the decay amplitude in order to calculate it. For



**Figure 5.1:** Feynman diagrams for the tree-level process contributing to  $B_d \rightarrow p \bar{p}$  (left), and for the gluonic penguin process contributing to  $B^+ \rightarrow p \bar{\Lambda}$  (right).

Decay Channel	QCD Sum Rules	Diquark Model	Pole Model	
			Harmonic Oscillator Model	MIT Bag Model
$B_d \rightarrow p \bar{p}$	$1.2 \times 10^{-6}$	$2.9 \times 10^{-6}$	$7.0 \times 10^{-6}$	$1.1 \times 10^{-7}$
$B_d \rightarrow \Lambda \bar{\Lambda}$			$2 \times 10^{-7}$	0
$B^+ \rightarrow p \bar{\Lambda}$	$\lesssim 3 \times 10^{-6}$			$2.2 \times 10^{-7}$

**Table 5.1:** Theoretical predictions for the branching ratios of different baryonic two-body  $B$  decays. A blank entry indicates that the branching ratio was not calculated using that particular model.

example, the pole model considers the decay amplitude as a ‘‘pole amplitude’’, where the  $B$  undergoes a strong decay (evaluated as a 3-point effective interaction) to a charmless baryon and an intermediate  $b$ -baryonic state (such as  $\Sigma_b^{(*)}$  or  $\Lambda_b^{(*)}$ ), which then decays weakly to give the other charmless baryon. Meanwhile the diquark model considers the decay amplitude as a weak decay into a diquark and an antidiquark, followed by the  $q\bar{q}$  pair creation. The  $q$  ( $\bar{q}$ ) then combines with the diquark (antidiquark) to form the baryon (antibaryon).

A summary of predictions using these different frameworks is given in Table 5.1. It can be seen that the predictions of different theoretical approaches do not agree, even in order of magnitude. Within the pole model framework two different methods of calculating the hadronic matrix element are used: the harmonic oscillator model [131, 132], and the MIT bag model [128]. The MIT bag model calculations only take the parity-conserving matrix element (corresponding to intermediate  $b$ -baryon states with positive parity) into account, as calculating the parity-violating matrix element in the bag model poses additional theoretical challenges. The bag model values shown in Table 5.1 are the parity-conserving-only values.

Decay Channel	Belle UL [125]	Babar UL [126]	CLEO UL [127]
$B_d \rightarrow p\bar{p}$	$1.1 \times 10^{-7}$	$2.7 \times 10^{-7}$	$7.0 \times 10^{-6}$
$B_d \rightarrow \Lambda\bar{\Lambda}$	$3.2 \times 10^{-7}$		$3.9 \times 10^{-6}$
$B^+ \rightarrow p\bar{\Lambda}$	$3.2 \times 10^{-7}$		$2.6 \times 10^{-6}$

**Table 5.2:** Experimental upper limits on the branching ratios of different two-body charmless baryonic  $B$  decays. Limits shown correspond to a 90% confidence level.

### 5.1.2 Experimental Limits on Branching Ratios

Charmless two-body baryonic  $B$  decays have been searched for by the CLEO, Babar and Belle experiments. Their current best upper limits on the branching ratios of several baryonic modes are listed in Table 5.2; they are of the order of  $10^{-7}$ .

The experimental 90% confidence level (C.L.) upper limit for the  $B_d \rightarrow p\bar{p}$  branching ratio,  $1.1 \times 10^{-7}$ , is dominated by the latest Belle search [125], which uses  $414 \text{ fb}^{-1}$  of data. It is interesting to note that this result has already ruled out the predictions of the QCD sum rules and diquark models, as well as the harmonic oscillator method within the pole model. Indeed, pressure is starting to be applied to the prediction of the MIT bag model in the pole model. Updated results from Belle with a larger data sample – to date Belle has accumulated over  $700 \text{ fb}^{-1}$  at the  $\Upsilon(4S)$  resonance – would have the potential, if no signal were to be seen, to exclude all theoretical calculations to date.

Considering the inconsistency among the several theoretical predictions for the  $B_d \rightarrow p\bar{p}$  branching ratio and the experimental data, it is clear that an early measurement of this branching ratio by LHCb could have an impact on the theoretical understanding of the dynamics involved in (two-body) baryonic  $B$  decays, while providing the first observation of a charmless two-body baryonic  $B$  decay.

## 5.2 Selection of $B_d \rightarrow p\bar{p}$ Events

The decay  $B_d \rightarrow p\bar{p}$  is topologically identical to the  $B \rightarrow \{\pi, K\}^+ \{\pi, K\}^-$  modes that are considered as core channels for the LHCb physics programme. The inclusive selection for  $B \rightarrow h^+ h'^-$  modes was described in Sec. 4.2.3. This set of cut variables, optimised for the selection of the “standard”  $B \rightarrow \{\pi, K\}^+ \{\pi, K\}^-$  decays, served as a starting point for the selection of  $B_d \rightarrow p\bar{p}$  candidates. Indeed, since that set of cuts efficiently selects signal events in these channels while keeping background from so-called  $b\bar{b}$  inclusive (i.e. generic  $b$ -decays) and minimum bias events down to an acceptable level, it is reasonable to expect that it will perform well also for  $B_d \rightarrow p\bar{p}$  events. The selection described below will be

Type of cut	Value
$B$ invariant mass window	$m_{B_d} \pm 50$ MeV
Min. DLL( $p - \pi$ ) for both daughters	5.0
Min. DLL( $p - K$ ) for both daughters	0.0
Max. $\chi^2/n_{\text{DoF}}$ for both daughter tracks	3.0

**Table 5.3:** List of extra cuts applied in the selection of  $B_d \rightarrow p\bar{p}$  events.

equally applicable to the decay  $B_s \rightarrow p\bar{p}$ , which is discussed in Sec. 5.7.1.

To reduce the background level further, several extra cuts have been added to the inclusive  $B \rightarrow h^+h'^-$  selection. To select protons and reject charged pions and kaons, particle identification (PID) cuts on the difference between the log-likelihood of two particle hypotheses,  $\text{DLL}(p - \pi) > 5$  and  $\text{DLL}(p - K) > 0$ , were employed. A tighter proton-pion separation cut  $\text{DLL}(p - \pi)$  is necessary as a typical event contains more pions than kaons. In addition, a cut on the track quality,  $\chi^2/n_{\text{DoF}} < 3$ , was applied to each of the daughter tracks in order to reduce the background due to ghost (fake) tracks.

The extra cuts that are added to the inclusive selection (see Table 4.1 for the inclusive selection cuts) are given in Table 5.3. The background suppression resulting from these extra cuts is discussed in Sec. 5.3.2.

## 5.3 Background studies

### 5.3.1 Sources of Background

As stated in Sec. 5.2, the inclusive  $B \rightarrow h^+h'^-$  selection enables high suppression of background from  $b\bar{b}$  inclusive and minimum bias events. This leads to an expectation that the selection described above should suppress such backgrounds similarly well for  $B_d \rightarrow p\bar{p}$  events. Particular care is nevertheless needed, given the low  $B_d \rightarrow p\bar{p}$  branching ratio.

In addition to these inclusive backgrounds, two other potential sources of background were investigated: background from misidentified  $B \rightarrow h^+h'^-$  final states and from partially reconstructed three-body  $B \rightarrow hhh$  decays.

Given their same signature and considerably higher branching ratios,  $B \rightarrow h^+h'^-$  decays are a potentially dangerous source of background. The following decay modes were considered:  $B_d \rightarrow K^+\pi^-$ ,  $B_s \rightarrow K^+K^-$  and  $\Lambda_b \rightarrow pK^-$ . These are the  $B \rightarrow h^+h'^-$  decay modes that most likely to contribute to the background to  $B_d \rightarrow p\bar{p}$ , since they all have both a relatively high branching ratio,  $\mathcal{O}(10^{-5})$ , and at least one proton or kaon in the final state.

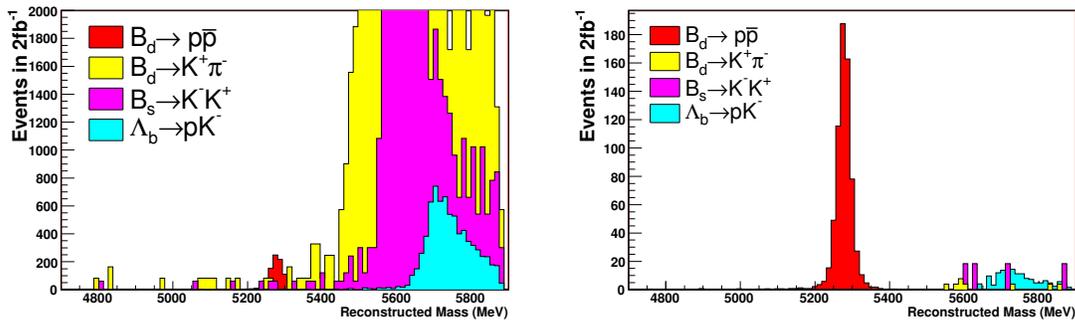
Partially reconstructed  $B \rightarrow hhh$  decays are typically not a dangerous source of back-

ground in the analysis of the standard  $B \rightarrow h^+h'^-$  modes because their reconstructed mass tends to fall significantly below the signal mass peak, as one track is not reconstructed. However, in the present analysis these decays have the potential to populate the  $B_d \rightarrow p\bar{p}$  mass peak region as the reconstructed mass of most two-track pairs from these decays will shift significantly upward under a proton mass hypothesis. The  $B \rightarrow hhh$  decays considered in the present study are:  $B_d \rightarrow \pi^+\pi^-\pi^0$ ,  $B_d \rightarrow K_S \pi^-\pi^+$ ,  $B^+ \rightarrow \pi^+\pi^-\pi^+$ ,  $B^+ \rightarrow \pi^+\pi^-K^+$ ,  $B^+ \rightarrow \pi^+K^-K^+$ ,  $B^+ \rightarrow p\bar{p}\pi^+$ ,  $B^+ \rightarrow p\bar{p}K^+$  and  $B^+ \rightarrow K^+K^-K^+$ .

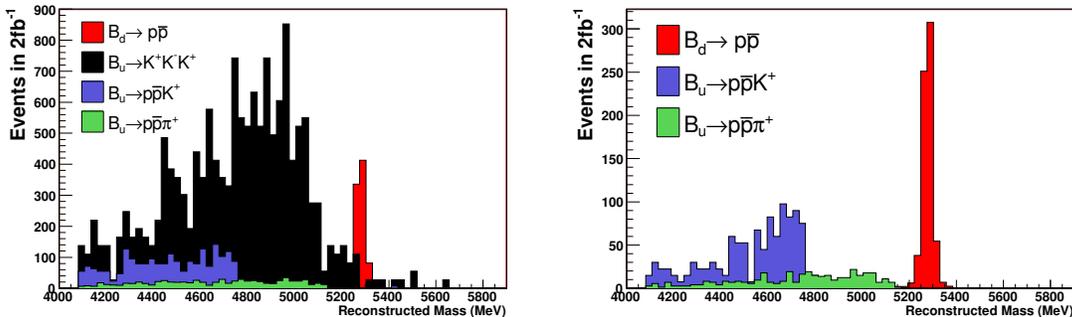
### 5.3.2 Background Suppression

To examine the role of PID and invariant mass cuts in reducing the specific background, the invariant mass distributions for signal and background before applying PID cuts can be compared with those after the PID cuts are applied. The following four figures, Figs. 5.2 – 5.5, show the signal and specific background events from  $2 \text{ fb}^{-1}$  of data which are expected to pass the inclusive  $B \rightarrow h^+h'^-$  selection with a very wide mass window of  $m_{B_d} \pm 1200 \text{ MeV}$ . In each figure, the left-hand plot shows the distributions before any PID cuts, and the right-hand plot shows the distributions after the application of the PID cuts given in Sec. 5.2. The efficiency of these PID cuts on the signal is 73%.

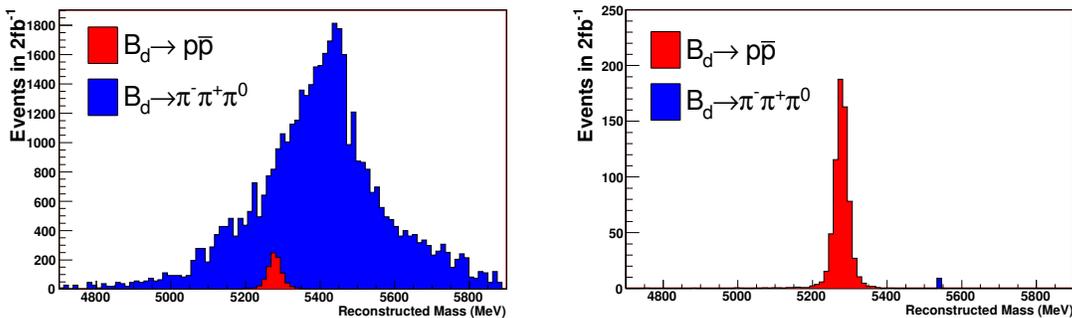
Figure 5.2 shows the mass distribution for  $B_d \rightarrow p\bar{p}$  alongside those for the  $B \rightarrow h^+h'^-$  channels considered as backgrounds here. Figure 5.3 shows the mass distribution for  $B_d \rightarrow p\bar{p}$  alongside those for  $B \rightarrow hhh$  channels where the reconstructed mass tends to fall well below  $m_{B_d}$ . Finally, Figures 5.4 and 5.5 compare the signal mass distribution with those for the  $B \rightarrow hhh$  channels where the reconstructed mass distribution overlaps significantly with the signal.



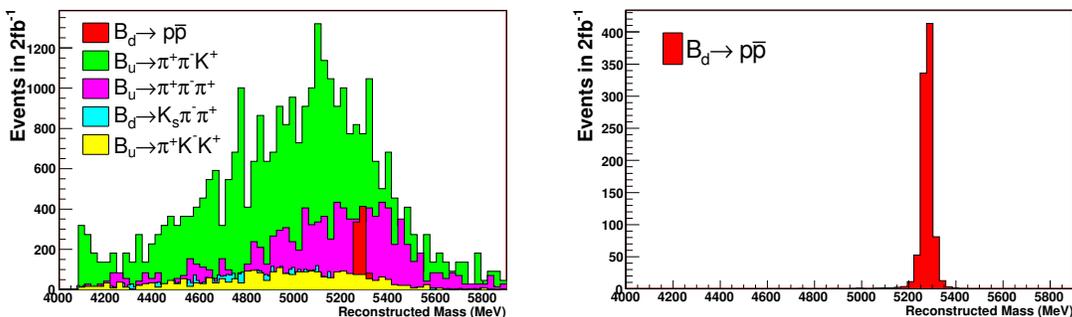
**Figure 5.2:** Reconstructed mass distributions before (left) and after (right) PID cuts for  $B_d \rightarrow p\bar{p}$  (red),  $B_d \rightarrow K^+\pi^-$  (yellow),  $B_s \rightarrow K^+K^-$  (purple) and  $\Lambda_b \rightarrow pK^-$  (cyan) events.



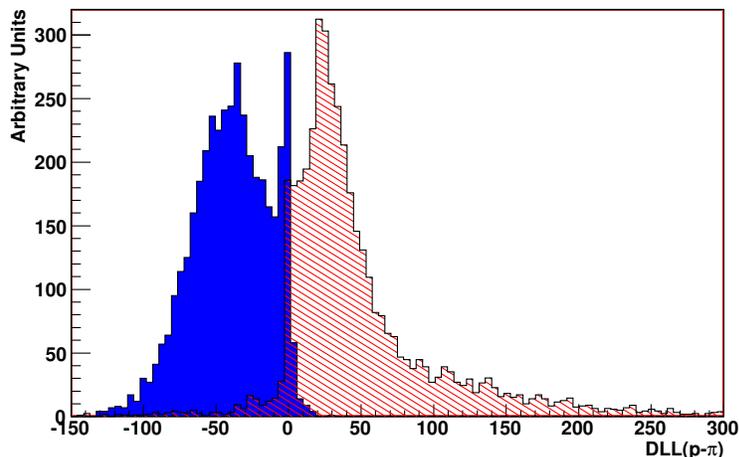
**Figure 5.3:** Reconstructed mass distributions before (left) and after (right) PID cuts for  $B_d \rightarrow p\bar{p}$  (red),  $B^+ \rightarrow K^+K^-K^+$  (black),  $B^+ \rightarrow p\bar{p}K^+$  (violet) and  $B^+ \rightarrow p\bar{p}\pi^+$  (dark green) events.



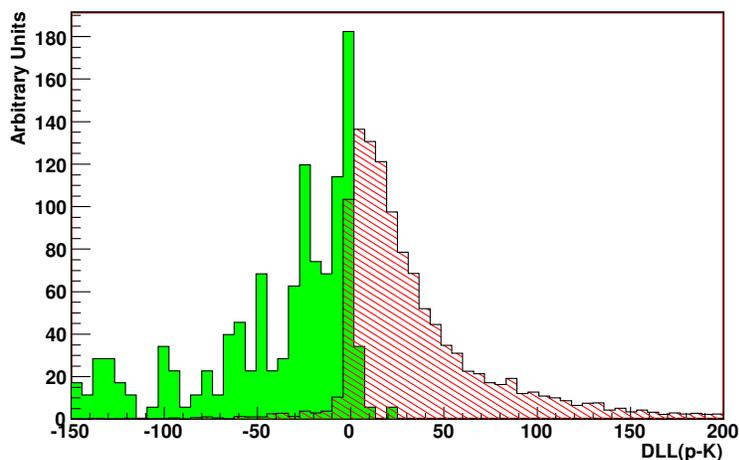
**Figure 5.4:** Reconstructed mass distributions before (left) and after (right) PID cuts for  $B_d \rightarrow p\bar{p}$  (red) and  $B_d \rightarrow \pi^+\pi^-\pi^0$  (blue) events.



**Figure 5.5:** Reconstructed mass distributions before (left) and after (right) PID cuts for  $B_d \rightarrow p\bar{p}$  (red),  $B^+ \rightarrow \pi^+\pi^-K^+$  (green),  $B^+ \rightarrow \pi^+\pi^-\pi^+$  (pink),  $B_d \rightarrow K_S\pi^-\pi^+$  (cyan) and  $B^+ \rightarrow \pi^+K^-K^+$  (yellow) events.



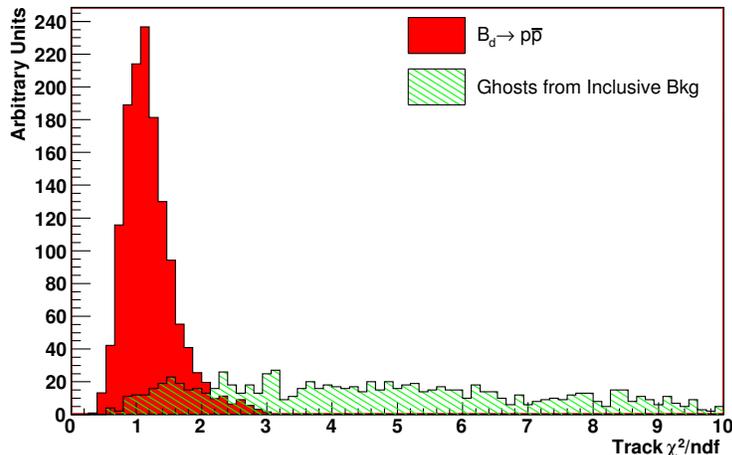
**Figure 5.6:** Distributions of  $DLL(p-\pi)$  for  $B_d \rightarrow p\bar{p}$  (red, hatched) and  $B_d \rightarrow \pi^+\pi^-\pi^0$  (blue) events.



**Figure 5.7:** Distributions of  $DLL(p-K)$  for  $B_d \rightarrow p\bar{p}$  (red, hatched) and  $B^+ \rightarrow \pi^+\pi^-K^+$  (green) events.

From Figures 5.2 and 5.3 it is clear that a mass window of 50 MeV around the  $B_d$  mass removes almost all of the background from the  $B \rightarrow h^+h'^-$  channels and some of the  $B \rightarrow hhh$  channels, even before PID cuts being applied. However, Figures 5.4 and 5.5 show that this is not the case for the remaining  $B \rightarrow hhh$  channels, where PID cuts play a crucial rôle in reducing the background level.

Figures 5.6 and 5.7 illustrate why the PID cuts are so effective against the specific background. Figure 5.6 compares the distributions of  $DLL(p-\pi)$  for true protons from  $B_d \rightarrow p\bar{p}$  and true pions from  $B_d \rightarrow \pi^+\pi^-\pi^0$ . There is excellent separation: the cut at  $DLL(p-\pi) = 5$  removes almost all of the  $B_d \rightarrow \pi^+\pi^-\pi^0$  events whilst having very high signal efficiency.



**Figure 5.8:** Distributions of track  $\chi^2/n_{\text{DoF}}$  for  $B_d \rightarrow p\bar{p}$  (red) and ghost tracks from  $b\bar{b}$  inclusive (green, hatched) events. See text for details.

Figure 5.7 compares the distributions of  $\text{DLL}(p - K)$  for true protons from  $B_d \rightarrow p\bar{p}$  and true kaons from  $B^+ \rightarrow \pi^+\pi^-K^+$ . Again, there is very good separation: the cut at  $\text{DLL}(p - K) = 0$  removes a large majority of the background with minimal effect on the signal.

The PID cuts are in fact even more powerful than Figures 5.6 and 5.7 at first suggest, because they are applied to both daughter particles, so that an event without a true proton must have both daughters in the tail of the DLL distribution, which is very unlikely. Of course a few specific backgrounds ( $\Lambda_b \rightarrow pK^-$ ,  $B^+ \rightarrow p\bar{p}K^+$  and  $B^+ \rightarrow p\bar{p}\pi^+$ ) do contain a true proton, but the invariant mass distributions for these channels do not overlap with the signal peak (see Figures 5.2 and 5.3). To summarise, it is clear that the  $B \rightarrow hhh$  background which is present in the signal mass region can be very effectively reduced using PID information.

Finally, the effectiveness of the cut on the quality of the track fit is illustrated in Fig. 5.8, which compares the track  $\chi^2/n_{\text{DoF}}$  distributions from signal tracks with that for ghost tracks from  $b\bar{b}$  inclusive events. The  $b\bar{b}$  inclusive events in Fig. 5.8 have only been passed through a very loose preselection, so that enough  $b\bar{b}$  inclusive events survive to make the  $\chi^2/n_{\text{DoF}}$  distribution for ghosts tracks apparent. A cut of  $\chi^2/n_{\text{DoF}} < 3$  removes many ghost tracks, while having almost no effect on signal tracks. This cut removes the few  $b\bar{b}$  inclusive events which survive the other selection cuts.

## 5.4 LHCb Sensitivity to $B_d \rightarrow p\bar{p}$

### 5.4.1 Selection Performance

The standard  $B \rightarrow h^+h'^-$  selection was run on simulated data samples of signal and specific background events, with a typical size of order 100,000 events per channel. An inclusive  $b\bar{b}$  sample consisting of 980,000 stripped events<sup>1</sup> was used, corresponding to around 27 million unstripped events. Finally, around 5.5 million minimum bias events, which passed the first-level (L0) trigger, were used. The physics analysis program `DaVinci` [117] version `v19r14` was employed.

A mass window of 50 MeV around the nominal  $B_d$  mass was applied for the signal and specific background sources, while for the  $b\bar{b}$  inclusive and minimum bias samples the mass window was widened to 600 MeV, to increase the available statistics. Their yields were later scaled down by a factor of 12 to account for this widened window.

The selection described in Sec. 5.2 is found to have an efficiency ( $\epsilon_{\text{sel}}$ ) of 9.8% on signal. No background event of any category was found to survive this selection. The specific background events are rejected mainly by a combination of PID and invariant mass cuts, whilst the  $b\bar{b}$  inclusive and minimum bias events are rejected mainly by the standard selection cuts such as impact parameter significances and transverse momentum  $p_T$  cuts.

Given that no background events are selected, an upper limit on the selection efficiency  $\epsilon_{\text{sel}}$  at 90% C.L. can be found using the Feldman-Cousins statistical approach [133]. For zero selected events, it attributes an upper limit on the number of selected events of 2.44 at 90% C.L., which can be converted into a 90% C.L. upper limit on  $\epsilon_{\text{sel}}$ .

Table 5.4 gives the size of each background sample<sup>2</sup>, and the upper limit on  $\epsilon_{\text{sel}}$  for each. It also gives the geometrical efficiency ( $\epsilon_{\text{geo}}$ ) of each channel, i.e. the percentage of events that pass the generator level cut which requires that all decay products of the signal  $B$  lie within the LHCb acceptance<sup>3</sup>. The branching ratio of each channel is also given. For all but one channel, the branching ratio used is the mean value given in the Particle Data Group 2008 review [15]. The exception is  $\Lambda_b \rightarrow pK^-$ , the branching ratio of which was recently measured [134] for the first time by the CDF Collaboration.

---

<sup>1</sup>Stripped events are events passing at least one of a set of looser pre-selection cuts typically designed for specific  $B$ -decays.

<sup>2</sup>For the inclusive  $b\bar{b}$  sample, the stripping efficiency has been accounted for to give an effective sample size.

<sup>3</sup>The  $B_s \rightarrow K^+K^-$  sample uses a looser cut, which requires instead that the signal  $B$  lies within the LHCb acceptance.

Decay Channel	Size of Sample	Upper Limit on $\epsilon_{\text{sel}}$	$\epsilon_{\text{geo}}$ (%)	Branching Ratio
$B_d \rightarrow \pi^+\pi^-\pi^0$	139,322	$1.8 \times 10^{-5}$	18.3	$2.5 \times 10^{-5}$
$B_d \rightarrow K_S \pi^-\pi^+$	457,135	$5.3 \times 10^{-6}$	21.2	$2.2 \times 10^{-5}$
$B^+ \rightarrow \pi^+\pi^-\pi^+$	61,138	$4.0 \times 10^{-5}$	17.9	$1.6 \times 10^{-5}$
$B^+ \rightarrow \pi^+\pi^-K^+$	65,110	$3.7 \times 10^{-5}$	18.3	$5.5 \times 10^{-5}$
$B^+ \rightarrow \pi^+K^-K^+$	63,141	$3.9 \times 10^{-5}$	18.7	$5.0 \times 10^{-6}$
$B^+ \rightarrow p\bar{p}\pi^+$	68,078	$3.6 \times 10^{-5}$	20.0	$1.6 \times 10^{-6}$
$B^+ \rightarrow p\bar{p}K^+$	46,737	$5.2 \times 10^{-5}$	20.6	$5.9 \times 10^{-6}$
$B^+ \rightarrow K^+K^-K^+$	67,118	$3.6 \times 10^{-5}$	19.3	$3.4 \times 10^{-5}$
$B_d \rightarrow K^+\pi^-$	296,700	$8.2 \times 10^{-6}$	20.2	$1.9 \times 10^{-5}$
$B_s \rightarrow K^+K^-$	99,048	$2.5 \times 10^{-5}$	34.6	$3.3 \times 10^{-5}$
$\Lambda_b \rightarrow pK^-$	64,579	$3.8 \times 10^{-5}$	21.1	$5.0 \times 10^{-6}$
Inclusive $b\bar{b}$	26.94M	$9.1 \times 10^{-8}$	43.7	n/a
L0-yes minimum bias	5.55M	$4.4 \times 10^{-7}$	n/a	n/a

**Table 5.4:** Channel-specific values used to evaluate signal and background yields. Limits shown correspond to a 90% confidence level upper limit.

## 5.4.2 Signal and Background Yields

The yield in  $2 \text{ fb}^{-1}$  of integrated luminosity (defined as a “nominal LHCb year”, amounting to  $10^7$  s of data taking at a nominal average luminosity) for signal and for a specific  $B$  background is given by

$$\text{Yield} = \epsilon_{\text{sel}} \times \epsilon_{\text{trig}} \times \epsilon_{\text{geo}} \times f_B \times \text{B.R.} \times \int \mathcal{L} dt \times 2 \times \sigma_{b\bar{b}}. \quad (5.2)$$

Here,  $\epsilon_{\text{trig}}$  is the efficiency of the LHCb trigger on events of that type which have passed the selection,  $f_B$  is the probability for a  $b$  quark to form the relevant  $B$  hadron,  $\mathcal{L}$  is the average luminosity at the LHCb interaction point, and  $\sigma_{b\bar{b}}$  is the beauty cross section at 14 TeV centre-of-mass energy. The factor of 2 accounts for the production of both a  $b$  and a  $\bar{b}$ .

The yields for  $b\bar{b}$  inclusive background and minimum bias background are, respectively,

$$\text{Yield}_{b\bar{b}} = \epsilon_{\text{sel}} \times \epsilon_{\text{trig}} \times \epsilon_{\text{geo}} \times \int \mathcal{L} dt \times \sigma_{b\bar{b}} \times \frac{1}{12} \quad (5.3)$$

and

$$\text{Yield}_{\text{MB}} = \epsilon_{\text{sel}} \times \epsilon_{\text{trig}} \times \text{Rate}_{\text{L0}} \times \text{Time}_{2\text{fb}^{-1}} \times \frac{1}{12}, \quad (5.4)$$

where these yields have been scaled down by a factor 12 to account for the wide mass window used in their selection (see Sec. 5.4.1). In Eqn. 5.4,  $\text{Rate}_{\text{L0}}$  is the output rate of the L0 trigger,

Parameter	Value used
Luminosity $\mathcal{L}$	$2 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$
$\sigma_{b\bar{b}}$	$500 \mu\text{b}$
Rate <sub>L0</sub>	$10^6 \text{ Hz}$
Time <sub>2fb<sup>-1</sup></sub>	$10^7 \text{ s}$
$f_B(B_d)$	$0.40 \pm 0.01$
$f_B(B^+)$	$0.40 \pm 0.01$
$f_B(B_s)$	$0.11 \pm 0.01$
$f_B(\Lambda_b)$	$0.07 \pm 0.02$

**Table 5.5:** General constants used to evaluate signal and background yields. The values for  $f_B$  are taken from [15].

Decay channel	Size of sample	$\epsilon_{\text{sel}}$ (%)	$\epsilon_{\text{geo}}$ (%)	Branching ratio	$2 \text{ fb}^{-1}$ yield
$B_d \rightarrow p\bar{p}$	46707	9.8	21.9	$1.1 \times 10^{-7}$	678

**Table 5.6:** Values used to evaluate signal yield.

and Time<sub>2fb<sup>-1</sup></sub> is the time taken to accumulate  $2 \text{ fb}^{-1}$  of data at the nominal LHCb luminosity. Values for these quantities, along with  $\sigma_{b\bar{b}}$  and  $f_B$ , are given in Table 5.5. Note that the  $\sigma_{b\bar{b}}$  value is a working assumption, based on theoretical predictions from QCD.

The efficiency of the trigger with respect to offline selected  $B \rightarrow h^+h'^-$  events has been studied in [56]. It was found that for the main  $B \rightarrow h^+h'^-$  channels,  $\epsilon_{\text{trig}}$  was always in the range [35%, 38%]. In the current study, a value of  $\epsilon_{\text{trig}} = 36\%$  is assumed for the signal and the  $B \rightarrow h^+h'^-$  backgrounds. Although the precision was limited by available statistics in the simulation of background events,  $\epsilon_{\text{trig}}$  for offline selected background events (both physics backgrounds like  $B \rightarrow hhh$  and  $b\bar{b}$  inclusive background) was found to be broadly similar, although somewhat lower, than  $\epsilon_{\text{trig}}$  for the signals. This is expected, as the background events that have passed the selection are in some sense “signal-like”. Here a conservative assumption is made, taking  $\epsilon_{\text{trig}} = 36\%$  for specific background and  $b\bar{b}$  inclusive background. The value for the L0-yes minimum bias sample is higher (68%) as these events have already passed the L0 trigger, which has an efficiency of  $\simeq 53\%$  on the selected events.

The signal yield per  $2 \text{ fb}^{-1}$  was calculated assuming the current experimental limit of  $1.1 \times 10^{-7}$  for the  $B_d \rightarrow p\bar{p}$  branching ratio; it is found to be 678 events per  $2 \text{ fb}^{-1}$ . The values used to calculate the signal yields are given in Table 5.6. Table 5.7 shows the upper limits on the background yields and the resulting background-to-signal (B/S) ratios. Each upper limit is calculated using the relevant yield equation as given above, taking the relevant upper limit on  $\epsilon_{\text{sel}}$  from Table 5.4.

Decay Channel	Upper limit on $2 \text{ fb}^{-1}$ yield	Upper limit on B/S
$B_d \rightarrow \pi^+ \pi^- \pi^0$	23	0.034
$B_d \rightarrow K_S \pi^- \pi^+$	7	0.011
$B^+ \rightarrow \pi^+ \pi^- \pi^+$	33	0.048
$B^+ \rightarrow \pi^+ \pi^- K^+$	109	0.160
$B^+ \rightarrow \pi^+ K^- K^+$	10	0.015
$B^+ \rightarrow K^+ K^- K^+$	69	0.101
$B^+ \rightarrow p \bar{p} \pi^+$	3	0.005
$B^+ \rightarrow p \bar{p} K^+$	18	0.027
$B_d \rightarrow K^+ \pi^-$	9	0.013
$B_s \rightarrow K^+ K^-$	22	0.033
$\Lambda_b \rightarrow p K^-$	2	0.003
Inclusive $b\bar{b}$	1,188	1.75
Minimum Bias	249,129	367.4

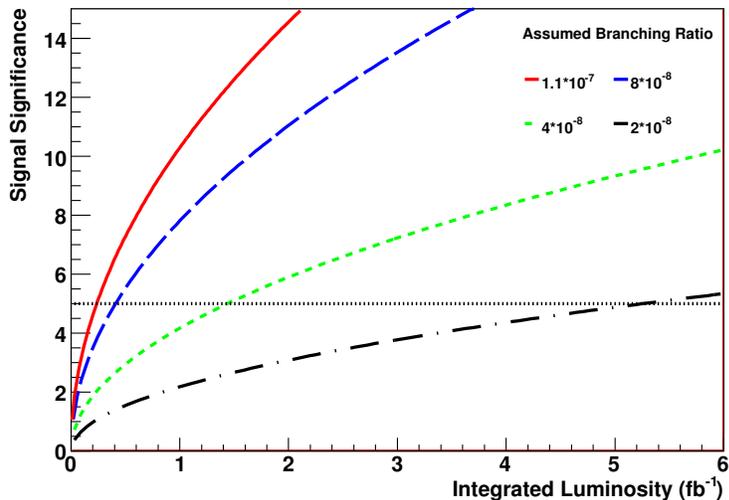
**Table 5.7:** Upper limits on background yields and resulting background-to-signal (B/S) ratios. Limits correspond to a 90% confidence level.

The constraint placed on the minimum bias yield is severely limited by the quantity of available simulated minimum bias events. It is assumed in what follows that the amount of minimum bias events passing the trigger and the final offline selection will be negligible, so that any possible background contribution from minimum bias events can be neglected. This assumption is expected to be valid because selection cuts which suppress background from  $B$  decays should be even more effective against background from charm and lighter quarks. The assumption can be tested using the large amounts of minimum bias events that will be collected once data taking begins.

With this condition, the upper limit (at 90% confidence level) on the total background-to-signal ratio is:

$$\begin{aligned}
 B/S_{\text{total}} &= B/S_{B \rightarrow h+h'^-} + B/S_{B \rightarrow hhh} + B/S_{b\bar{b}\text{inclusive}} \\
 &< 0.05 + 0.40 + 1.75 && \text{(at 90\% C.L.)} \\
 &= 2.20.
 \end{aligned}$$

Note that there are two conservative assumptions used in the calculation of this upper limit. Firstly, simply summing a number of 90% confidence levels from different sources overestimates the total upper limit. Secondly, it is assumed that there is no overlap between the inclusive and specific background sources.



**Figure 5.9:** Significance of the  $B_d \rightarrow p\bar{p}$  signal as a function of integrated luminosity, assuming a centre-of-mass energy of  $E_{CM} = 14$  TeV. Each curve assumes a different  $B_d \rightarrow p\bar{p}$  branching ratio:  $1.1 \times 10^{-7}$  (red, solid),  $8 \times 10^{-8}$  (blue, long dashes),  $4 \times 10^{-8}$  (green, short dashes) and  $2 \times 10^{-8}$  (black, dash-dot).

### 5.4.3 Signal Significance

In the following calculation of the expected signal significance, the upper limit on the background-to-signal ratio calculated above is conservatively taken as the central value. The significance of the  $B_d \rightarrow p\bar{p}$  signal can be expressed as a function of integrated luminosity and its (currently unknown) branching fraction. The significance is defined as

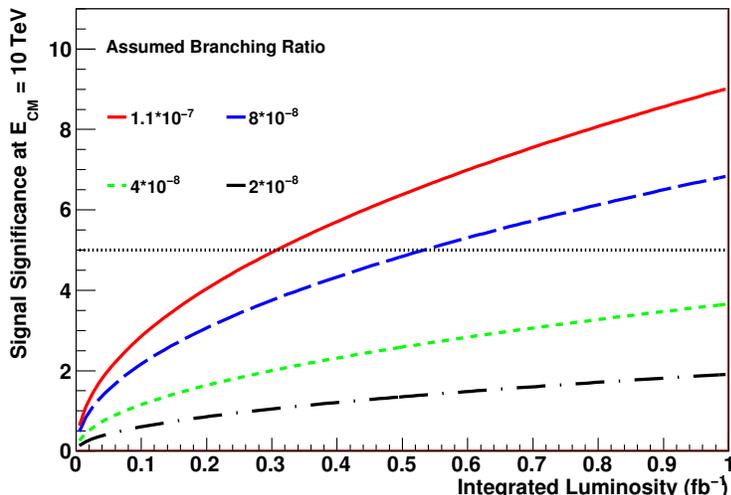
$$\text{Significance} = \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{\text{bkg}}}}, \quad (5.5)$$

where  $N_{\text{sig(bkg)}}$  is the number of signal (background) events.

Figure 5.9 shows, for four possible values for the branching ratio of  $B_d \rightarrow p\bar{p}$ , how the significance evolves with integrated luminosity. If the true branching ratio is close to the current experimental upper limit, a discovery at the  $5\sigma$  level can be made with early LHCb data. Taking the current upper limit, a  $5\sigma$  significance is achieved with of order  $0.25 \text{ fb}^{-1}$  of data. Even if the true branching ratio is a factor of 5 below the current limit, which would contradict all existing theoretical predictions, LHCb can make a discovery with  $5.3 \text{ fb}^{-1}$ .

### 5.4.4 Prospects for 2010

The initial physics run of the LHC, starting in late 2009 and finishing in late 2010, will feature a lower than nominal centre-of-mass energy. In this section the discovery prospects for  $B_d \rightarrow p\bar{p}$  in 2010 are examined, assuming a centre-of-mass energy of  $E_{CM} = 10$  TeV.



**Figure 5.10:** Significance of the  $B_d \rightarrow p\bar{p}$  signal as a function of integrated luminosity, assuming  $E_{\text{CM}} = 10$  TeV. Each curve assumes a different  $B_d \rightarrow p\bar{p}$  branching ratio:  $1.1 \times 10^{-7}$  (red, solid),  $8 \times 10^{-8}$  (blue, long dashes),  $4 \times 10^{-8}$  (green, short dashes) and  $2 \times 10^{-8}$  (black, dash-dot).

The lower  $E_{\text{CM}}$  value will lower the beauty cross section  $\sigma_{b\bar{b}}$ . The Monte Carlo-based program Pythia [135], used to simulate high-energy  $pp$  collisions, predicts (version 6.2) that at  $E_{\text{CM}} = 10$  TeV,  $\sigma_{b\bar{b}}$  will be a factor  $\simeq 0.77$  less than at  $E_{\text{CM}} = 14$  TeV<sup>4</sup>. Hence, assuming that the reconstruction and trigger performances are similar at  $E_{\text{CM}} = 10$  TeV, the signal and background<sup>5</sup> yields in Table 5.7 will all be scaled down by the same factor. Hence the background-to-signal ratio for a given  $B_d \rightarrow p\bar{p}$  branching ratio will not change. However the signal significance will change as it is not a linear function of the signal and background yields.

The signal significance for  $B_d \rightarrow p\bar{p}$  as a function of integrated luminosity for data taken with  $E_{\text{CM}} = 10$  TeV is shown in Fig 5.10. It can be seen that  $0.32 \text{ fb}^{-1}$  of data at  $E_{\text{CM}} = 10$  TeV is required to achieve  $5\sigma$  significance, assuming a branching ratio close to the current experimental limit. Such an amount of data is unlikely to be collected in the initial physics run of the LHC. However it should be recalled that the background-to-signal ratio assumed in Fig. 5.10 is the upper limit at 90% confidence level. Should the background-to-signal ratio on data be significantly below this upper limit, an observation may still be made

<sup>4</sup>Very recently it was announced that the initial stages of running in 2010 will take place at  $E_{\text{CM}} = 7$  TeV, with the energy rising to 10 TeV over the course of the run. At  $E_{\text{CM}} = 7$  TeV,  $\sigma_{b\bar{b}}$  will be a factor  $\simeq 0.58$  less than at  $E_{\text{CM}} = 14$  TeV.

<sup>5</sup>The minimum bias yield will not scale with the same factor, but since minimum bias background has been neglected this is not relevant here.

in the initial physics run. For illustration, assuming that  $0.2 \text{ fb}^{-1}$  of data are collected at  $E_{\text{CM}} = 10 \text{ TeV}$  (this is considered feasible for the initial physics run), and again assuming a branching ratio close to the current experimental limit, a background-to-signal ratio of 1.09 would lead to a  $5\sigma$  discovery of  $B_d \rightarrow p\bar{p}$ .

## 5.5 Measurement of the $B_d \rightarrow p\bar{p}$ Branching Ratio

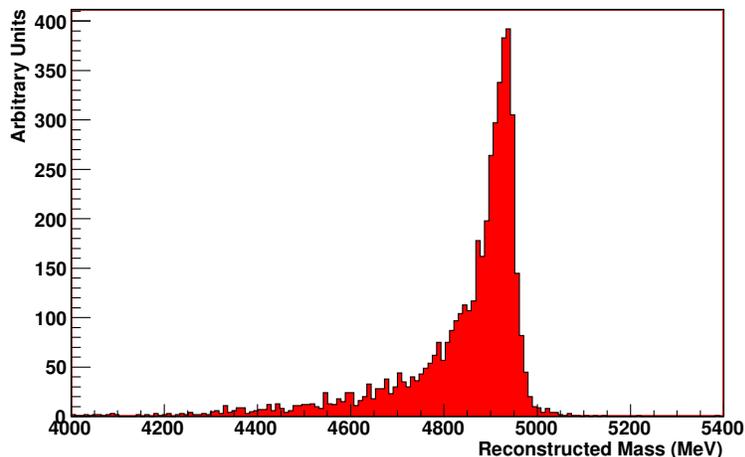
Once a  $5\sigma$  signal for  $B_d \rightarrow p\bar{p}$  has been observed, its branching ratio can be determined. This can be done by normalising the branching ratio to that of  $B_d \rightarrow K^+\pi^-$ , which is the  $B \rightarrow h^+h'^-$  channel with the highest branching ratio and smallest experimental error. Its branching ratio has been precisely measured by the B-factories to be  $(1.94 \pm 0.06) \times 10^{-5}$ . The branching ratio for  $B_d \rightarrow p\bar{p}$  can be obtained from this in the following way.

Firstly, the yield in each channel will be measured. This is done by fitting to the peak in the invariant mass distribution after all relevant cuts have been applied, taking the invariant mass distributions of the various backgrounds into account. For  $B_d \rightarrow K^+\pi^-$ , the fit to the invariant mass distribution will be done simultaneously with the other main  $B \rightarrow h^+h'^-$  channels, as their distributions overlap significantly (the fitting method is described in [56]). PID cuts will not have been applied at this stage. For  $B_d \rightarrow p\bar{p}$  the only other signal mass distribution in that region will be  $B_s \rightarrow p\bar{p}$ , which is expected to make a negligible contribution to the mass peak (see Sec. 5.7.1). PID cuts will need to be applied before the fit to the  $B_d \rightarrow p\bar{p}$  mass peak can be made, as otherwise the signal will be swamped by specific backgrounds (see Sec. 5.3.2) and  $b\bar{b}$  inclusive background.

Then the efficiencies of the trigger and selection need to be taken into account. The efficiency  $\epsilon_{\text{sel}}$  of a full offline selection for a  $B \rightarrow h^+h'^-$  channel can be broken down using  $\epsilon_{\text{sel}} = \epsilon_{\text{top/kin}} \times \epsilon_{\text{PID}}$ , where  $\epsilon_{\text{top/kin}}$  accounts for the topological and kinematic cuts and  $\epsilon_{\text{PID}}$  accounts for the PID cuts. Then  $\epsilon_{\text{top/kin}}$  should be almost the same for  $B_d \rightarrow K^+\pi^-$  and  $B_d \rightarrow p\bar{p}$ , while  $\epsilon_{\text{PID}}$  will be different. However  $\epsilon_{\text{PID}}$  can be estimated using control channels which will be used to calibrate the PID performance.

For calibration of the proton PID, the decay  $\Lambda \rightarrow p\pi^-$  is used [136]. Studies aiming to calibrate the kaon and pion PID using  $D^{*+} \rightarrow D^0(K\pi)\pi^+$  decays [56] have shown very promising results, hence the proton PID calibration can also be expected to perform well, especially considering the higher production rate of  $\Lambda$  with respect to  $D^{*+}$ . Hence the estimation of  $\epsilon_{\text{PID}}$  should not be a major source of systematic error.

The trigger efficiency is expected to be very similar in the two channels (assuming that the trigger mass window is wide enough, see Sec. 5.6), as they will have to pass the same cuts there. Finally, the ratio of the geometrical efficiencies for the two channels can be accurately



**Figure 5.11:** Reconstructed mass distribution for offline selected  $B_d \rightarrow p\bar{p}$  events taking the pion mass hypothesis for the  $B$ -daughters.

estimated from simulation.

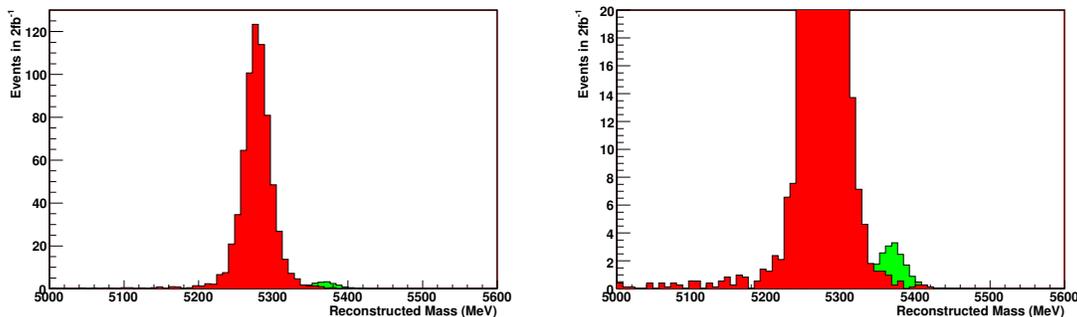
Under these assumptions, Eqn. 5.2 can be used to obtain an expression for the  $B_d \rightarrow p\bar{p}$  branching ratio ( $\epsilon_{\text{PID}}$  does not appear for  $B_d \rightarrow K^+\pi^-$  as the  $B_d \rightarrow K^+\pi^-$  yield will be calculated before PID cuts are applied):

$$\text{BR}(B_d \rightarrow p\bar{p}) = \text{BR}(B_d \rightarrow K^+\pi^-) \times \frac{\text{Yield}_{p\bar{p}} \times \epsilon_{\text{geo}(K\pi)}}{\text{Yield}_{K\pi} \times \epsilon_{\text{PID}(p\bar{p})} \times \epsilon_{\text{geo}(p\bar{p})}} \quad (5.6)$$

All of the terms on the right-hand side of Eqn. 5.6 can be obtained as described above to give a value for  $\text{BR}(B_d \rightarrow p\bar{p})$ .

## 5.6 Trigger Mass Window

In the high-level trigger selection of  $B \rightarrow h^+h'^-$  channels, a pion mass hypothesis is assumed for all  $B$ -daughters. As a result, the reconstructed mass for a  $B_d \rightarrow p\bar{p}$  event will be far below the nominal  $B_d$  mass, as shown in Fig. 5.11. It follows that in order to select  $B_d \rightarrow p\bar{p}$  events in the trigger, the mass window must extend at least  $\simeq 500$  MeV below the nominal  $B_d$  mass. A mass window of  $m_{B_d} \pm 600$  MeV for example would select 87% of offline-selected signal events. Having the trigger mass window similar to this would, as well as losing some signal events, also add an extra systematic effect in the calculation of the  $B_d \rightarrow p\bar{p}$  branching ratio, as the percentage of events lost due to the mass window would have to be estimated using a Monte Carlo simulation of the invariant mass shape. Hence, from the point of view of  $B_d \rightarrow p\bar{p}$ , the best mass window for the trigger to use would be one which started far below the  $B_d \rightarrow p\bar{p}$  peak — a window lower limit around 4400 MeV.



**Figure 5.12:** Reconstructed mass distributions for  $B_d \rightarrow p\bar{p}$  (red) and  $B_s \rightarrow p\bar{p}$  (green) events. The right-hand plot is a detail of the left-hand plot.

Recently, the above trigger issues for  $B_d \rightarrow p\bar{p}$  have caused the mass window used in the HLT2 trigger (see Sec. 2.2.7.2) for  $B \rightarrow h^+h'^-$  events to be widened. The studies in [56] assumed a trigger mass window of [5000, 5800] MeV, which would select almost no  $B_d \rightarrow p\bar{p}$  events. The window has now been widened to [4800, 5800] MeV [137], which would select 76% of offline-selected  $B_d \rightarrow p\bar{p}$  events.

## 5.7 Other Two-Body Charmless Baryonic $B$ Decays

### 5.7.1 Prospects for the Observation of $B_s \rightarrow p\bar{p}$

The decay  $B_s \rightarrow p\bar{p}$  is expected to be rather suppressed relative to  $B_d \rightarrow p\bar{p}$ , because in the Standard Model it can only occur via annihilation diagrams, given that none of the quarks in the final state are those of the initial  $B_s$  meson. Despite the lack of any firm theoretical predictions, it is legitimate to search for this rare mode given its identical final state. LHCb has the potential to improve the present experimental branching ratio upper limit of  $\approx 5.9 \times 10^{-5}$  by several orders of magnitude.

Figure 5.12 shows the mass distributions of  $B_d \rightarrow p\bar{p}$  and  $B_s \rightarrow p\bar{p}$  events as expected after applying the full selection (including PID cuts) to  $2 \text{ fb}^{-1}$  of data, in the absence of background. Note that the core width of the mass distribution is 17 MeV. The assumed branching ratios are  $1.1 \times 10^{-7}$  for  $B_d \rightarrow p\bar{p}$  and  $1.1 \times 10^{-8}$  for  $B_s \rightarrow p\bar{p}$ . Far fewer  $B_s \rightarrow p\bar{p}$  events are present due to the lower assumed branching ratio, and the fact that  $f_B(B_d)/f_B(B_s) \approx 4$ .

Figure 5.12 illustrates that the much larger  $B_d \rightarrow p\bar{p}$  peak will form a significant “background” to  $B_s \rightarrow p\bar{p}$ . Assuming that the branching ratios are as given above, with  $2 \text{ fb}^{-1}$  there will be, within  $\pm 50 \text{ MeV}$  of  $m_{B_s}$ , 24  $B_d \rightarrow p\bar{p}$  events and 18  $B_s \rightarrow p\bar{p}$  events, leading to a background-to-signal ratio of 1.33 from  $B_d \rightarrow p\bar{p}$  events alone. This value will increase

if the mass resolution in data is poorer than in the simulation. Combined with the presence of the background sources described in Sec. 5.3, this means that a significant amount of data will be required in order to observe the  $B_s \rightarrow p\bar{p}$  mass peak.

### 5.7.2 Prospects for Decays Involving a $\Lambda$

Other two-body charmless baryonic  $B$  decays which can be reconstructed by LHCb include  $B^+ \rightarrow p\bar{\Lambda}$  and  $B_d \rightarrow \Lambda\bar{\Lambda}$ , in cases where the  $\Lambda$  decays via  $\Lambda \rightarrow p\pi^-$ . As with  $B_d \rightarrow p\bar{p}$ , theoretical calculations of their branching fractions are subject to debate, with predictions differing by up to an order of magnitude (see Table 5.1). The dominant diagram contributing to  $B^+ \rightarrow p\bar{\Lambda}$  is expected to be the gluonic penguin that was shown in Fig. 5.1. This is in contrast to  $B_d \rightarrow p\bar{p}$ , where the tree process is expected to dominate.

LHCb can also look for the  $B_s \rightarrow \Lambda\bar{\Lambda}$  decay. Whereas  $B_s \rightarrow p\bar{p}$  should be rather suppressed compared to  $B_d \rightarrow p\bar{p}$  (see Sec. 5.7.1),  $B_s \rightarrow \Lambda\bar{\Lambda}$  could have a relatively large branching ratio of the same order of that of  $B^+ \rightarrow p\bar{\Lambda}$ , as the penguin diagrams mediating these processes are identical apart from their spectator quarks. The  $u \leftrightarrow s$  transformation relating the spectator quarks (and thus the penguin amplitudes) for these two decays is known as a V-spin transformation. This corresponds to a subgroup of  $SU(3)$ , in analogy to the isospin ( $u \leftrightarrow d$ ) and U-spin ( $d \leftrightarrow s$ ) transformations.

The latest upper limit measurements from Belle, shown on Table 5.2, are close to excluding all theoretical predictions for these decays. Moreover, a discovery of the decay  $B_d \rightarrow \Lambda\bar{\Lambda}$  would crucially constrain the MIT bag model calculation approach in the pole model.

Selections for these decays, which have a different topology to  $B_{d,s} \rightarrow p\bar{p}$ , have yet to be developed within LHCb. The performance of such selections will depend crucially on the efficiency for reconstructing  $\Lambda$  baryons from  $B$  decays. The reconstruction of  $\Lambda$  baryons from  $B$  decays has been shown [138, 139] to be viable for LHCb, however a detailed analysis of the reconstruction efficiency has yet to be made. For those events that are successfully reconstructed, the presence of two protons in the final state will enable excellent suppression of background through PID cuts (compare Sec. 5.3.2). Hence observation of  $B^+ \rightarrow p\bar{\Lambda}$  and  $B_s \rightarrow \Lambda\bar{\Lambda}$  may be feasible, depending on the value of their branching ratios.

## 5.8 Conclusions

The theoretical and experimental status of two-body charmless baryonic  $B$  decays has been reviewed. There is no agreement between different theoretical models on the branching ratios for such decays, and most of the theoretical predictions have already been invalidated experimentally.

The LHCb inclusive selection for  $B \rightarrow h^+h'^-$  decays has been modified to give a selection for the exclusive decays  $B_{d,s} \rightarrow p\bar{p}$ . This selection was found to have a good efficiency for signal events, while suppressing backgrounds from other  $B$  decays sufficiently to give a satisfactory upper limit on the background-to-signal ratio for backgrounds from  $B$  decays. However, the relatively small amount of available simulated minimum bias events prevented a useful constraint being obtained on the amount of minimum bias background passing the selection. Hence the expected signal significances were calculated under the assumption that the number of minimum bias events passing the selection will be negligible. This assumption will be tested using the large amounts of minimum bias events that will be collected once data taking begins.

The potential of LHCb to discover  $B_d \rightarrow p\bar{p}$  using this selection has been explored in detail. It is found that around  $0.25 \text{ fb}^{-1}$  of data (at  $E_{\text{CM}} = 14 \text{ TeV}$ ) will be required for a  $5\sigma$  discovery, assuming that the  $B_d \rightarrow p\bar{p}$  branching fraction is close to the current experimental upper limit. Even if the true branching fraction is a factor of 5 lower than the present upper limit, a discovery can be made with about  $5 \text{ fb}^{-1}$  of data.

Observations of other charmless two-body baryonic  $B$  decays, such as  $B_s \rightarrow p\bar{p}$ ,  $B^+ \rightarrow p\bar{\Lambda}$  and  $B_s \rightarrow \Lambda\bar{\Lambda}$ , may also be possible, depending on their branching ratios. A detailed analysis of the prospects for decays containing a  $\Lambda$  in the final state requires further study, including the development of suitable selections.

# Chapter 6

## Measuring Direct and Mixing-Induced

## $\mathcal{CP}$ Asymmetries in $B_d \rightarrow \pi^+ \pi^-$

This chapter presents a study of the capability of the LHCb detector to measure the time-dependent  $\mathcal{CP}$  asymmetry distribution in the decay  $B_d \rightarrow \pi^+ \pi^-$  and extract the direct and mixing-induced  $\mathcal{CP}$  asymmetries. A toy Monte Carlo simulation, making use of information from the full Monte Carlo simulation of LHCb  $B \rightarrow h^+ h'^-$  events, is used. The statistical sensitivity to the  $\mathcal{CP}$  asymmetries using a small amount of data (compatible with the amount expected to be collected in the initial 2009-2010 run of the LHC) is assessed. Section 6.1 outlines the strategy that will be used to extract the  $\mathcal{CP}$  asymmetries from the toy data. Section 6.2 describes the selection of  $B_d \rightarrow \pi^+ \pi^-$  events, and gives the expected signal and background yields at LHCb. Sections 6.3 and 6.4 discuss the construction of the toy data and fitting functions respectively. The results from the fit to the toy data, including the sensitivity to the  $\mathcal{CP}$  asymmetries, are given in Sec. 6.5. The study is summarised in Sec. 6.6.

### 6.1 Strategy for Fit to Time-Dependent $\mathcal{CP}$ Asymmetry

As discussed in Sec. 1.4, a measurement of the direct and mixing-induced  $\mathcal{CP}$  asymmetries ( $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ ) in  $B_d \rightarrow \pi^+ \pi^-$  can be used to constrain the CKM angles  $\alpha$  (in combination with the decays  $B^+ \rightarrow \pi^+ \pi^0$  and  $B_d \rightarrow \pi^0 \pi^0$ ) and  $\gamma$  (in combination with the decay  $B_s \rightarrow K^+ K^-$ ). From a purely experimental viewpoint, there is also considerable interest in the measurement of  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ , as the two existing measurements from the B-Factories disagree significantly. The toy Monte Carlo study presented in this chapter demonstrates the sensitivity of LHCb to these  $\mathcal{CP}$  asymmetries. The resulting sensitivity to the effective CKM parameter  $\sin 2\alpha_{\text{eff}}$  will also be calculated. However the inclusion of information from other decay modes to extract  $\alpha$  or  $\gamma$  is beyond the scope of this work, and will not be discussed.

In this study, hard cuts are placed on the particle identification (PID) likelihoods for the  $\pi^\pm$  candidates and on the invariant mass of the  $B_d$  candidate, in order to suppress backgrounds arising from other  $B \rightarrow h^+h'^-$  decays. The signal and background yields are estimated using a fit to the invariant mass distribution, and are then fixed in a fit to the proper time distribution that extracts the direct and mixing-induced  $\mathcal{CP}$  asymmetries

This approach is in contrast to the one used in [56], where all of the  $B \rightarrow h^+h'^-$  decays are considered together, and the event yields, mean lifetimes and  $\mathcal{CP}$  asymmetries for all  $B \rightarrow h^+h'^-$  decays are determined in one simultaneous fit to a joint Probability Density Function (PDF) which includes the mass, proper time, tagging and PID information. The different  $B \rightarrow h^+h'^-$  decays are statistically separated from each other during this simultaneous fit, using their mass and PID distributions.

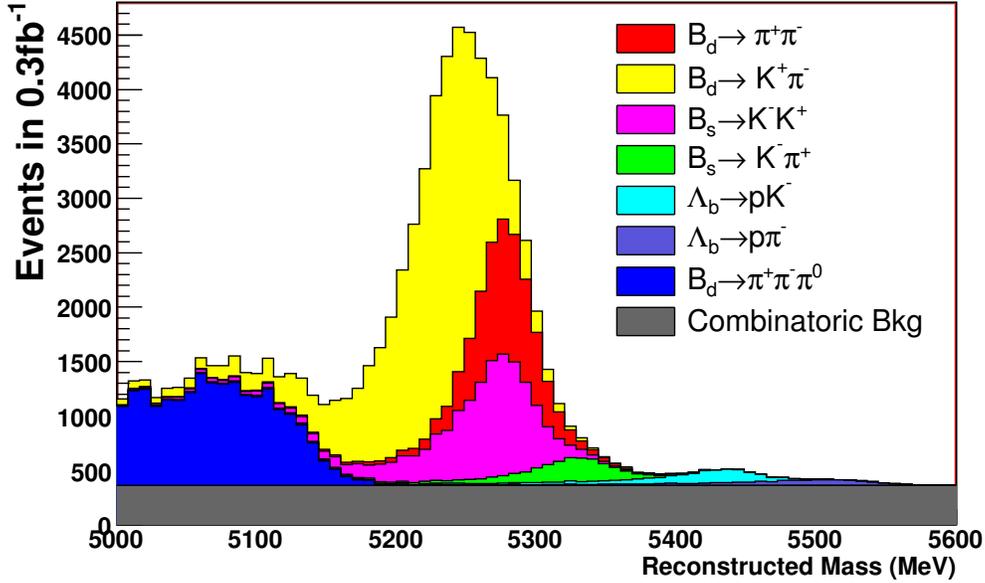
The current study, on the other hand, uses the mass and PID distributions to separate the signal from background *before* the fit. It will be seen that this leaves only two significant backgrounds; the specific background  $B_d \rightarrow K^+\pi^-$  and combinatorial background. These are the only backgrounds that are modelled in the mass fit, and also in the fit to the proper time distribution to extract the signal  $\mathcal{CP}$  asymmetries. The parameters of these backgrounds, and the other parameters of the signal proper time distribution, are fixed in the proper time fit using information from external analyses, which will be discussed. The advantages and disadvantages of the approach used in the current study with respect to the approach from [56] will be pointed out as they arise.

For the toy Monte Carlo study, an integrated luminosity of  $0.3 \text{ fb}^{-1}$  is assumed. This corresponds to the amount of data that is expected to be collected during the 2010 run of the LHC. It will be seen that this sample is sufficient to allow a stable fit for the direct and mixing-induced  $\mathcal{CP}$  asymmetries to be made.

## 6.2 Event Selection and Yields

In this study, the following backgrounds to  $B_d \rightarrow \pi^+\pi^-$  are considered (see Sec. 4.2.3): the two-body specific backgrounds  $B_d \rightarrow K^+\pi^-$ ,  $B_s \rightarrow K^+K^-$ ,  $B_s \rightarrow K^-\pi^+$ ,  $\Lambda_b \rightarrow pK^-$  and  $\Lambda_b \rightarrow p\pi^-$ , partially reconstructed three-body specific backgrounds, and combinatorial background. A number of different three-body decays can act as partially reconstructed backgrounds to  $B_d \rightarrow \pi^+\pi^-$ , however for simplicity the contribution from these backgrounds is modelled using the distributions for the decay  $B_d \rightarrow \pi^+\pi^-\pi^0$ .

The inclusive  $B \rightarrow h^+h'^-$  selection, described in Sec. 4.2.3, is applied to full simulation samples of signal and specific background. All decay modes are reconstructed with a  $\pi\pi$  mass hypothesis for the daughter particles. This shifts the invariant mass distributions for



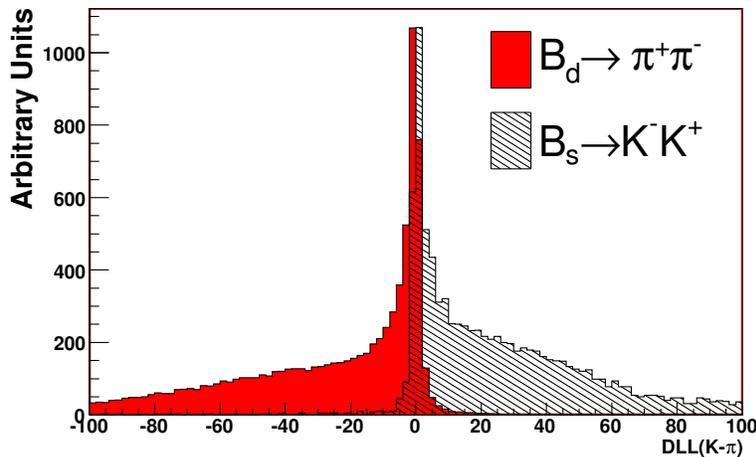
**Figure 6.1:** Cumulative mass distribution for  $B \rightarrow h^+h'^-$  events following the inclusive selection, with the number of simulated events corresponding to  $0.3 \text{ fb}^{-1}$  of data. The contributions from signal and from each background are shown. Yields are taken from [56].

non- $\pi\pi$  final states with respect to their nominal distributions with the correct mass hypothesis. A very wide mass window of  $[5000, 5600]$  MeV is used at this stage. This event selection is run using `v19r14` of the LHCb physics analysis package `DaVinci` [117].

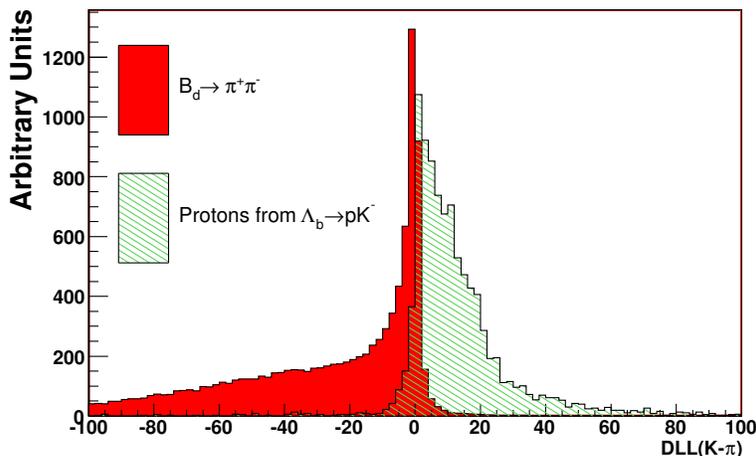
To examine the effect of adding tight mass and PID cuts to the inclusive selection, the mass and PID distributions from the full LHCb Monte Carlo simulation are used. For each specific decay considered, the mass distribution used is that for offline selected events in the full simulation. For the combinatoric background a flat mass distribution is assumed. While the combinatoric background mass distribution in the full simulation does show a shallow negative slope in mass, in the tight mass window that is used in final selection (see below), the distribution can be approximated as being flat.

Figure 6.1 shows the expected cumulative mass distribution after the inclusive selection, with the number of simulated events corresponding to  $0.3 \text{ fb}^{-1}$  of data. The contributions from signal and from each background are shown. The yields shown in Figure 6.1 have not been calculated by the current study; rather the yields found in [56] are used. These yields include the expected trigger efficiency of  $\simeq 36\%$  on offline-selected events. It can be seen that the signal suffers significant contamination, both from combinatoric background and from the specific backgrounds  $B_d \rightarrow K^+\pi^-$  and  $B_s \rightarrow K^+K^-$ .

The specific backgrounds can be drastically reduced by applying cuts on the PID



**Figure 6.2:** Comparison of  $\text{DLL}(K - \pi)$  distributions for selected  $B_d \rightarrow \pi^+\pi^-$  (red) and  $B_s \rightarrow K^+K^-$  (black, hatched) events.



**Figure 6.3:** Comparison of  $\text{DLL}(K - \pi)$  distributions for selected  $B_d \rightarrow \pi^+\pi^-$  (red) and  $\Lambda_b \rightarrow pK^-$  (green, hatched) events.

likelihoods (see Sec. 4.2.3) for the daughter tracks, to form an exclusive selection for  $B_d \rightarrow \pi^+\pi^-$ . The PID likelihood distribution,  $\text{DLL}(K - \pi)$ , for pions is taken from selected events in the full simulation of  $B_d \rightarrow \pi^+\pi^-$ . Similarly the distribution for kaons is taken from  $B_s \rightarrow K^+K^-$ , and that for protons from  $\Lambda_b \rightarrow pK^-$ . These distributions are compared in Figs. 6.2 and 6.3.

Due to small differences in the momentum distributions for different  $B \rightarrow h^+h'^-$  decays, there are in fact small differences between for example the  $\text{DLL}(K - \pi)$  distributions for pions from  $B_d \rightarrow \pi^+\pi^-$  decays and for pions from  $B_d \rightarrow K^+\pi^-$  decays. However these differences are not large enough to significantly affect the efficiencies that are found by the

simplified approach used here, and so are neglected. All daughters for the three-body and combinatoric backgrounds are assumed to be pions. This is a conservative assumption for backgrounds to a  $B_d \rightarrow \pi^+\pi^-$  signal, as the true three-body and combinatoric backgrounds will also contain some kaons and a few protons.

The chosen PID cut applied to both daughter tracks is

$$\text{DLL}(K - \pi) < 0. \quad (6.1)$$

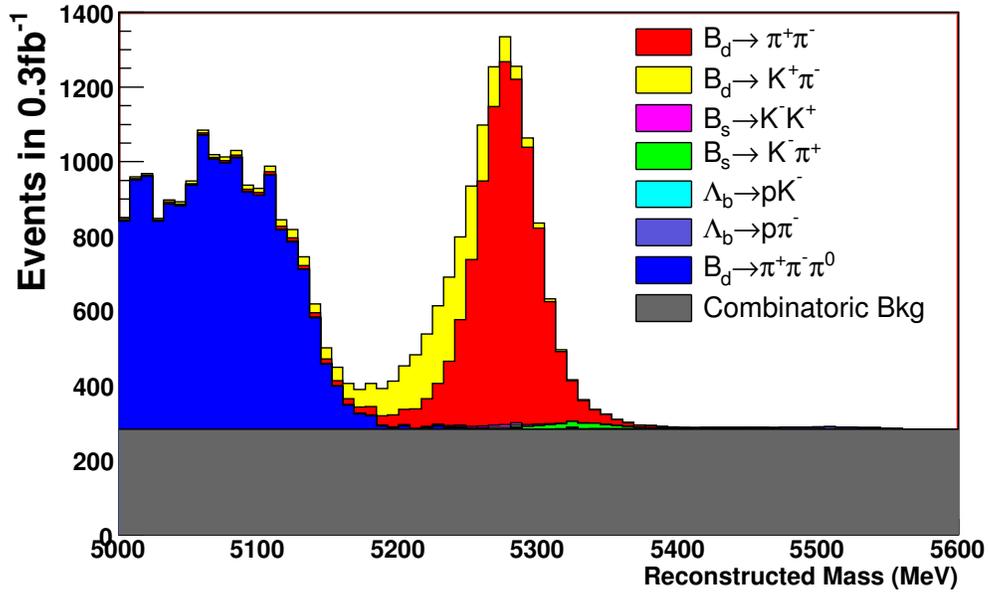
This translates as a requirement that both tracks are, according to the information from the RICH detectors, more likely to be pions than kaons.

It can be seen from Figs. 6.2 and 6.3 that this cut will have a high efficiency for pions while suppressing most of the background from kaons and protons. In fact the efficiency of the PID cut on the distributions used here is  $\simeq 88\%$  for true pions,  $\simeq 8\%$  for true kaons and  $\simeq 13\%$  for true protons. Since each final state particle is assumed to have the same PID distribution independent of which decay it is from (see above), the resulting efficiency for a given decay is simply the product of the efficiency of the cut for each daughter type.

The signal and background yields following these PID cuts are calculated by taking the above before-PID yields, and factoring in the efficiency of the PID cuts for each final state particle as given above. Figure 6.4 shows the expected cumulative mass distribution following the PID cuts, with the number of simulated events corresponding to  $0.3 \text{ fb}^{-1}$  of data. The contributions from signal and from each background are shown. The  $B \rightarrow h^+h'^-$  backgrounds are highly suppressed by the PID cuts, although a significant amount of  $B_d \rightarrow K^+\pi^-$  events remain, due to the high yield of this decay. The combinatoric background and three-body backgrounds are largely unaffected, due to the assumption that they consist only of pions. The presence of some combinatoric background will not adversely affect the fit for the  $\mathcal{CP}$  asymmetries, as it does not carry any  $\mathcal{CP}$  asymmetry itself. This is in contrast to the specific backgrounds, where  $\mathcal{CP}$  asymmetries can affect the proper time distribution and, if they are not taken into account, bias the extraction of the  $\mathcal{CP}$  asymmetries in the signal. This is particularly true for  $B_s \rightarrow K^+K^-$ , where the  $\mathcal{CP}$  asymmetries are currently unmeasured. However, since  $\Delta m_s \gg \Delta m_d$ , the effect of the fast  $B_s$  oscillations should largely integrate out when studying the slower  $B_d$  oscillations, so that any bias on the  $B_d \rightarrow \pi^+\pi^-$   $\mathcal{CP}$  asymmetries should be small.

Finally, a cut is applied on the invariant mass of the  $\pi\pi$  pair. The mass window is tightened from  $[5000, 5600] \text{ MeV}$  to  $[5229, 5329] \text{ MeV}$  (i.e.  $m_{B_d} \pm 50 \text{ MeV}$ ). This cut removes almost all of the three-body specific background, and also removes much of the remaining background from  $B_d \rightarrow K^+\pi^-$  and  $B_s \rightarrow K^-\pi^+$ .

Table 6.1 gives the signal and background yields for  $0.3 \text{ fb}^{-1}$  of data before the PID and mass cuts, after the PID cuts but before the mass cuts, and after both PID and mass cuts.



**Figure 6.4:** Cumulative mass distribution for  $B \rightarrow h^+h'^-$  events following the inclusive selection and PID cuts, with the number of simulated events corresponding to  $0.3 \text{ fb}^{-1}$  of data. The contributions from signal and from each background are shown.

Channel	$0.3 \text{ fb}^{-1}$ Yields		
	Before PID and Mass Cuts	After PID Cuts	After PID and Mass Cuts
$B_d \rightarrow \pi^+\pi^-$	8800	6910	6200
$B_d \rightarrow K^+\pi^-$	32500	2270	1140
$B_s \rightarrow K^+K^-$	10800	70	50
$B_s \rightarrow K^-\pi^+$	2300	170	80
$\Lambda_b \rightarrow pK^-$	1600	20	$\simeq 0$
$\Lambda_b \rightarrow p\pi^-$	1100	120	10
Total two-body background	48300	2650	1280
Three-body backgrounds	15000	11500	40
Combinatoric background	27000	21200	3550
Total background	90300	35400	4870

**Table 6.1:**  $B \rightarrow h^+h'^-$  and background yields with  $0.3 \text{ fb}^{-1}$  of data, before the PID and mass cuts, after the PID cuts but before the mass cuts, and after both PID and mass cuts. Yields for the  $B \rightarrow h^+h'^-$  decays before PID cuts are given to the nearest hundred, and after PID cuts to the nearest ten.

The only significant backgrounds remaining after the PID and mass cuts are the specific background  $B_d \rightarrow K^+\pi^-$  and the combinatoric background. Since the amount of background from  $\Lambda_b \rightarrow ph$  remaining after the PID and mass cuts is so small, it will hereafter be completely neglected. The small contributions from  $B_s \rightarrow K^+K^-$ ,  $B_s \rightarrow K^-\pi^+$  and the three-body backgrounds will be included in the toy data, but not modelled in the PDF used to fit the toy data. This simplifies the fitting functions considerably, but has the potential to introduce a small systematic error if the fit to the decay fractions becomes biased. In Sec. 6.5, the size of this bias will be assessed by comparing the results from fits to toy data including these backgrounds to those from fits to toy data that does not include them.

## 6.3 Construction of Toy Data

As stated in the previous section, the  $\Lambda_b$  backgrounds are neglected due to their tiny contribution to the overall number of events passing the full selection. Distributions of invariant mass and flavour-tagged proper time are generated for the signal, the specific backgrounds  $B_d \rightarrow K^+\pi^-$ ,  $B_s \rightarrow K^+K^-$  and  $B_s \rightarrow K^-\pi^+$ , the three-body backgrounds (represented by  $B_d \rightarrow \pi^+\pi^-\pi^0$ ) and the combinatoric background. The distributions of the PID likelihoods are not included in the toy data, as only the mass distributions are needed to distinguish between the different decays.

### 6.3.1 Mass Distributions for Toy Data

Mass distributions for the signal and specific backgrounds are generated following PDFs that are constructed using the kernel method [116] from the full simulation mass distributions under the  $\pi^+\pi^-$  mass hypothesis (these were shown in Fig. 6.1). For the three-body background, the full simulation distribution used is that for  $B_d \rightarrow \pi^+\pi^-\pi^0$ . For the combinatoric background, the mass distribution is again assumed to be flat.

### 6.3.2 Proper Time Distributions for Toy Data

The proper time distributions for the signal and each specific background are created following the expressions given in Sec. 1.2.2.2. In each decay mode the number of events with each initial  $B$  flavour (and, for the flavour specific modes, each final state) is determined by the value for the direct  $\mathcal{CP}$  asymmetry. The values used for the various  $\mathcal{CP}$  asymmetries are given in Table 6.2. The proper time distributions for the three-body and combinatoric background are assumed to be flavour tagging blind, i.e. the distributions for both flavour tags are identical.

Decay	$\mathcal{CP}$ Asymmetries	Source
$B_d \rightarrow \pi^+\pi^-$	$\mathcal{A}_{\mathcal{CP}}^{\text{dir}} = 0.38, \mathcal{A}_{\mathcal{CP}}^{\text{mix}} = 0.61$	World average of experimental measurements [15]
$B_d \rightarrow K^+\pi^-$	$\mathcal{A}_{\mathcal{CP}} = 0.10$	Average of experimental measurements [59, 140]
$B_s \rightarrow K^+K^-$	$\mathcal{A}_{\mathcal{CP}}^{\text{dir}} = 0.10, \mathcal{A}_{\mathcal{CP}}^{\text{mix}} = 0.25,$ $\mathcal{A}_{\Delta\Gamma} = 0.96$	Theoretical prediction [141]
$B_s \rightarrow K^-\pi^+$	$\mathcal{A}_{\mathcal{CP}} = 0.39$	Experimental measurement [142]

**Table 6.2:** Values used in the toy data for the  $\mathcal{CP}$  asymmetries.

The three-body background will be dominated by  $B^+$  and  $B_d$  decays, due to their high hadronisation fractions,  $f_{B^+} \simeq f_{B_d} \simeq 40\%$ . Here the three-body background proper time distribution is given a 50% contribution with the  $B^+$  lifetime (1.64 ps), and a 50% contribution with the  $B_d$  lifetime (1.53 ps).

The proper time distribution for the combinatoric background in the full simulation is seen to have a dominant component with an effective mean lifetime smaller than  $B$  lifetimes, and a smaller component with a much larger effective mean lifetime. The distribution used here consists of a 90% contribution from “short-lived” events with a mean lifetime of 0.5 ps, and a 10% contribution from “long-lived” events with a mean lifetime of 10 ps.

The tagging efficiencies ( $\epsilon_{\text{tag}}$ ) and mistag rates ( $\omega_{\text{tag}}$ ) are estimated by taking the values found by the detailed study reported in [56]<sup>1</sup>, and removing the contribution from the same side kaon tagger, as this is not applicable to  $B_d$  decays. The tagging strategy used in [56] was optimised for the simultaneous study of  $B_d$  and  $B_s$  decays, and hence is not the optimal strategy when considering a  $B_d$  decay alone.

Table 6.3 compares the tagging efficiencies and mistag rates using the unified tagging approach of [56] with those for the  $B_d$ -only approach used in the current study. The effective tagging power,  $\epsilon_{\text{eff}} \equiv \epsilon_{\text{tag}} \cdot (1 - 2\omega_{\text{tag}})^2$  is also shown. For each strategy, the tagging efficiencies for  $B_d$  decays,  $B_s$  decays, and combinatoric backgrounds, which are expected to be very similar to each other, are assumed to be equal for the sake of simplicity. It can be seen from Table 6.3 that the change of tagging strategy is an advantage of the approach used here: removing the same side kaon tagger improves the effective tagging power for  $B_d$  decays, while degrading it for  $B_s$  decays. This improves the sensitivity to the  $\mathcal{CP}$  asymmetries in the signal, while also reducing the observed  $\mathcal{CP}$  asymmetry in the  $B_s$  backgrounds.

<sup>1</sup>The study in [56] supersedes the study in [118] that was used to give the assumed tagging performance for the proper time resolution studies in chapter 4.

Tagging Strategy	Tagging Efficiency $\epsilon_{\text{tag}}$	Mistag Rate $\omega_{\text{tag}}$		Effective Tagging Power $\epsilon_{\text{eff}}$	
		$B_d$	$B_s$	$B_d$	$B_s$
Unified ( [56])	65%	37.0%	34.7%	4.39%	6.09%
$B_d$ -only (current study)	55%	35.5%	36.1%	4.66%	4.25%

**Table 6.3:** Comparison of tagging efficiencies and mistag rates in different tagging strategies.

The proper time acceptance effects arising from the cuts applied in the trigger and the offline selection are modelled on an event-by-event basis, rather than with an overall acceptance function. Each event is assigned a turn-on point for the proper time acceptance, and the event is rejected if the generated value for the proper time of the  $B$  meson is less than this turn-on point. The turning points are sampled from a Gaussian distribution with a mean of 1.0 ps and a width of 0.1 ps. Sampling the turning points in this way reproduces the proper time acceptance that is observed in the full Monte Carlo simulation (see Fig. 4.10). Both signal and background are expected to have very similar proper time acceptance behaviour, as all events must pass the same impact parameter cuts in order to be accepted.

Each decay is assigned a proper time resolution model following the model described in Sec. 4.2. For each decay, the distributions of the per-event error  $\sigma_{\tau_{\text{rec}}}$  on the proper time are taken from the full Monte Carlo simulation. Since all modes have to pass the same selection cuts, the  $\sigma_{\tau_{\text{rec}}}$  distributions are all broadly similar (the  $\sigma_{\tau_{\text{rec}}}$  distribution for  $B_d \rightarrow \pi^+\pi^-$  was shown in Fig. 4.3). The values used for the two parameters of the proper time resolution model are  $GM = 0.05$  and  $GS = 1.1$ . These choices approximately reproduce the values found from the fits to the full Monte Carlo proper time residual distributions for the  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  decays (see Sec. 4.2.4).

The same values for  $GM$  and  $GS$  are used for all decays. This is justified for the  $B \rightarrow h^+h'^-$  decays by the fact that the values of  $GM$  and  $GS$  for the  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  decays were compatible with each other. For the three-body and combinatoric backgrounds, the actual proper time resolution model may have different parameter values from the  $B \rightarrow h^+h'^-$  decays. However because their proper time distributions do not have oscillations, their dependence on  $GM$  and  $GS$  is very weak, and the same values can be assumed.

The final yields used for signal and background for the  $0.3 \text{ fb}^{-1}$  toy study are given in Table 6.4. The yields are lower than those quoted in Table 6.1 due to the 55% tagging efficiency (untagged events are not generated as they offer no sensitivity to the  $\mathcal{CP}$  asymmetries). Each set of generated toy data contains a total of exactly 6100 events, but the number of events of each decay varies around the mean values given in Table 6.4.

Channel	Number of Events in Toy Data
$B_d \rightarrow \pi^+\pi^-$	3422
$B_d \rightarrow K^+\pi^-$	626
$B_s \rightarrow K^+K^-$	29
$B_s \rightarrow K^-\pi^+$	46
Three-body backgrounds	25
Combinatoric background	1952
Total	6100

**Table 6.4:** Number of signal and background events in a typical toy data sample, corresponding to  $0.3 \text{ fb}^{-1}$  of data.

## 6.4 Construction of Distributions for Fit

As stated at the end of Sec. 6.2, the small contributions from  $B_s \rightarrow K^+K^-$ ,  $B_s \rightarrow K^-\pi^+$  and the three-body backgrounds are not modelled in the fit. This reduces the number of parameters that need to be considered, but may also introduce a small systematic bias on the extracted values for the direct and mixing-induced  $\mathcal{CP}$  asymmetries. The only decays which are modelled in the fit are the  $B_d \rightarrow \pi^+\pi^-$  signal, the specific background  $B_d \rightarrow K^+\pi^-$  and the combinatoric background. In the following, each of these will be referred to as a “decay class”.

To fit the time-dependent  $\mathcal{CP}$  asymmetry, the approach used in [119] to fit the  $B_s$  and  $B_d$  lifetimes using untagged proper time distributions has been developed to include the effects of oscillations and  $\mathcal{CP}$  asymmetries. This development renders the method suitable for fitting flavour-tagged proper time distributions.

The fit is an unbinned log-likelihood fit, carried out within the ROOT data analysis framework [122]. The package MINT [143], written by J. Rademacker, is used to interface to the minimisation package Minuit [123]. MINT provides the user with a number of predefined classes to describe objects such as fit parameters and data events, while maintaining a large amount of flexibility in the construction of the likelihood.

### 6.4.1 Mass Distributions for Fit

The mass distributions used in the fit are the same as those used in the toy data. The mass distributions from the full simulation are expected to give a good description of the data that will be seen, as the deviation from the Gaussian shape for the non- $\pi\pi$  final states is a known function of the momentum asymmetry between the daughters, which is measured for each

event. The effects of final state radiation, which also affects the observed mass distribution, are taken into account in the full simulation as well. The PDFs used for the mass distributions can be tested and improved on data using the fit to the invariant mass distributions.

The total PDF  $f(m)$  for the fit to the toy data mass distribution is the sum of the PDFs for each decay class:

$$f(m) = \sum_{\text{decay}} f(m|\text{decay}) \cdot P(\text{decay}), \quad (6.2)$$

where  $f(m|\text{decay})$  is the PDF for each decay class and  $P(\text{decay})$  is the fraction of the total sample coming from each decay. The fit to the invariant mass distribution of the toy data returns values for the decay fractions  $P(\text{decay})$ . The values for  $P(B_d \rightarrow \pi^+\pi^-)$  and  $P(B_d \rightarrow K^+\pi^-)$  are fitted, with  $P(\text{Combinatoric})$  constrained to be  $1 - P(B_d \rightarrow \pi^+\pi^-) - P(B_d \rightarrow K^+\pi^-)$ . As will be seen in the following section, these fractions and the mass PDFs are used to statistically separate the different decay classes in the fit to the proper time distribution. Following the application of tight PID cuts in the selection, it is not necessary to re-use the PID likelihoods in the fit to separate out the decay classes, as the modelled decay classes all have very distinct mass distributions (if for example  $B_s \rightarrow K^+K^-$  were included in the fit, it would be necessary to make use of the PID information as the  $B_s \rightarrow K^+K^-$  and  $B_d \rightarrow \pi^+\pi^-$  mass distributions are very similar under the  $\pi\pi$  mass hypothesis).

## 6.4.2 Proper Time Distributions for Fit PDFs

### 6.4.2.1 Treatment of Proper Time Acceptance

The proper time acceptance is treated on an event-by-event basis, with the acceptance function for each event being a step function that increases from 0 to 1 at some turning point  $t_{\min}$ . The value of  $t_{\min}$  for each event can be measured using a technique known as “swimming”, where the values of the impact parameters are calculated as a function of the decay time of the  $B$ , and the value of the decay time when the impact parameters both become large enough for the event to be accepted is the “turn-on point”<sup>2</sup>, denoted by  $t_{\min}$ . The swimming method was originally developed by the CDF collaboration [144], and has since been adapted for use by LHCb [120].

Once the proper time acceptance is taken into account, the proper time distribution for a given decay class will be a function of the proper time  $t$  of the decay<sup>3</sup>, the tag decision  $q$ , which is  $+1(-1)$  for an initial state tagged as a  $B_d(\bar{B}_d)$ , and  $t_{\min}$ . This can be factorised into

<sup>2</sup>For two-track decays such as  $B \rightarrow h^+h'^-$ , the event-by-event acceptance is a simple step function with one turn-on point and no turn-off point. For decays with more than two tracks, the event-by-event acceptance becomes more complicated.

<sup>3</sup>Here  $t$  is used, rather than  $\tau_{rec}$  as used in chapter 4, to avoid confusion with the mean  $B$  lifetime  $\tau_{B_d}$ .

the proper time distribution for a given value of  $t_{\min}$  and the distribution of  $t_{\min}$  for a given decay class:

$$f(t, q, t_{\min}|\text{decay}) = f(t, q|t_{\min}, \text{decay}) \cdot f(t_{\min}|\text{decay}). \quad (6.3)$$

Since all selected events are assumed to share the same proper time acceptance behaviour,  $f(t_{\min}|\text{decay})$  factorises out to become a common factor in the total likelihood. Hence it will not affect the location of the maximum of the likelihood, and so can be ignored for the purposes of the fit. In the case of different decays displaying different proper time acceptance behaviour on data, the behaviour can be measured with the swimming method, and the different  $f(t_{\min}|\text{decay})$  functions that are measured can be included in the likelihood. Note that in this case the measured  $f(t_{\min}|\text{decay})$  can be used to provide extra discriminating power between the different decay classes.

Since the relevant term in the likelihood is now  $f(t, q|t_{\min}, \text{decay})$ , in which  $t_{\min}$  behaves as a parameter rather than a variable, the only effect on the proper time PDFs of including the event-by-event acceptance function is that the normalisation becomes an integral from  $t = t_{\min}$  to  $t = \infty$ , rather than the usual  $t = 0$  to  $t = \infty$  [119]. The fact that including the proper time acceptance does not further complicate the proper time distribution has the benefit that the normalisation including the proper time acceptance can be done analytically. In many cases where an average acceptance function is used, the resulting integral has to be done numerically (see for example [114]).

#### 6.4.2.2 Conditional Proper Time Distributions

Following the approach of [118] and [56], a joint PDF is used for  $f(t, q|t_{\min}, \text{decay})$ . This PDF simultaneously describes the proper time distribution for events with either flavour tag. This is found by taking the expressions for the decay rate from a pure  $B$  or  $\bar{B}$  initial state (see Sec. 1.2.2.2) and summing them with the appropriate weight to obtain the experimentally observed rate for a given tag (see Eqn. 4.9).

The time PDFs are assumed to have perfect proper time resolution, as the low value of  $\Delta m_d$  means that the effect of including the resolution effects for  $B_d$  decays is negligible (see Sec. 4.1.1). Neglecting the proper time resolution effects allows the PDFs to be greatly simplified, and the integration for the normalisation to remain analytical. When fitting the time-dependent  $\mathcal{CP}$  asymmetry in a  $B_s$  decay such as  $B_s \rightarrow K^+K^-$ , the proper time distributions *will* be significantly affected by the resolution, so the resolution effects will need to be correctly considered in the fit, unlike for the  $B_d$  case here.

With the resolution effects neglected, the rates are

$$f(t, q|t_{\min}, B_d \rightarrow \pi^+\pi^-) = N(t, t_{\min}) \cdot e^{-t/\tau_{B_d}} [1 + q(1 - 2\omega_{\text{tag}})(\mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta m_d t) - \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta m_d t))] \quad (6.4)$$

for the  $B_d \rightarrow \pi^+\pi^-$  signal, and

$$f(t, q|t_{\min}, B_d \rightarrow K^+\pi^-) = N(t, t_{\min}) \cdot e^{-t/\tau_{B_d}} [1 + q(1 - 2\omega_{\text{tag}})\mathcal{A}_{\mathcal{CP}} \cos(\Delta m_d t)] \quad (6.5)$$

for the  $B_d \rightarrow K^+\pi^-$  specific background. In the above equations,  $N(t, t_{\min})$  is the normalisation factor, and  $\tau_{B_d}$  is the mean lifetime of the  $B_d$ , whose current world average experimental value is  $(1.53 \pm 0.01)$  ps [15].

The values of  $\tau_{B_d}$  and the oscillation frequency  $\Delta m$  are fixed to those used in the toy data, as they are already well known experimentally, and will be measured to an even higher precision by LHCb. The charge asymmetry  $\mathcal{A}_{\mathcal{CP}}$  for  $B_d \rightarrow K^+\pi^-$  has already been well-measured by the B-factories and the TeVatron experiments, with the current world average being  $0.10 \pm 0.02$  [15]. In the fit, this asymmetry is fixed to its world average value.

The mistag rate in the above PDFs is fixed to the same value used to generate the toy data, as the mistag rate and direct and mixing-induced  $\mathcal{CP}$  asymmetries cannot be fitted simultaneously. The mistag rate will be determined from data, using the methods discussed in chapter 4. The uncertainty from this determination is likely to be one of the main sources of systematic error in this analysis, however the fact that only the mistag rate for  $B_d$  decays is considered here is an advantage, as this rate can be directly measured from  $B_d \rightarrow K^+\pi^-$ , while the measurement of the mistag rate for  $B_s$  decays is more problematic (see discussion in Sec. 4.1.3). The effect of an incorrect determination of the mistag rate on the extraction of the direct and mixing-induced  $\mathcal{CP}$  asymmetries is discussed in Sec. 6.5.

The normalisation factor  $N(t, t_{\min})$  is calculated by integrating from  $t = t_{\min}$  to  $t = \infty$  and summing over the two discrete values of  $q$ . The summation over  $q$  causes the oscillation terms in the normalisation to cancel, leaving a simple expression:

$$N(t, t_{\min}) = \frac{1}{2 \cdot \tau_{B_d} \cdot e^{-t_{\min}/\tau_{B_d}}}, \quad (6.6)$$

which is the normalisation factor in both Eqn. 6.4 and Eqn. 6.5.

The proper time distribution used for the combinatoric background contribution to the fit PDF is the same as that used for the combinatoric background in the toy data generation. As described in chapter 4, there are various methods that can be employed to measure the combinatoric background proper time distributions on data, either from the upper sideband or by background subtraction [119]. The normalised rate for combinatoric background, which

does not depend on the tag decision, is

$$\begin{aligned} f(t, q|t_{\min}, \text{Combinatoric}) &= f_{\text{short}} \cdot N_{\text{short}}(t, t_{\min}) \cdot e^{-t/\tau_{\text{short}}} + f_{\text{long}} \cdot N_{\text{long}}(t, t_{\min}) \cdot e^{-t/\tau_{\text{long}}} \\ &= f_{\text{short}} \cdot \frac{e^{-t/\tau_{\text{short}}}}{2 \cdot \tau_{\text{short}} \cdot e^{-t_{\min}/\tau_{\text{short}}}} + f_{\text{long}} \cdot \frac{e^{-t/\tau_{\text{long}}}}{2 \cdot \tau_{\text{long}} \cdot e^{-t_{\min}/\tau_{\text{long}}}} \end{aligned} \quad (6.7)$$

where  $f_{\text{short}} = 0.9$  is the fraction of short-lived background with  $\tau_{\text{short}} = 0.5$  ps,  $f_{\text{long}} = 0.1$  is the fraction of long-lived background with  $\tau_{\text{long}} = 10.0$  ps. As with the other decay rates, the normalisation for each component is found by integrating from  $t = t_{\min}$  to  $t = \infty$  and summing over the tag values (the summation in this case simply introduces a factor of 1/2 into the normalisations).

### 6.4.2.3 Total Likelihood for Proper Time Fit

The total likelihood that is used to fit to the proper time distribution is

$$f(t, q, t_{\min}|m) = \sum_{\text{decay}} f(t, q, t_{\min}|\text{decay}) \cdot P(\text{decay}|m), \quad (6.8)$$

where  $P(\text{decay}|m)$ , the probability for an event with a given mass to be of a certain decay class, can be found from the decay fractions  $P(\text{decay})$  and the mass PDFs using Bayes' theorem:

$$P(\text{decay}|m) = \frac{f(m|\text{decay}) \cdot P(\text{decay})}{f(m)}. \quad (6.9)$$

As described above,  $f(t, q|t_{\min}, \text{decay})$  can be used in place of  $f(t, q, t_{\min}|\text{decay})$ , as the proper time acceptance behaviour of all decays classes is assumed to be the same. The total likelihood can then be constructed using the expressions for each  $f(t, q|t_{\min}, \text{decay})$  given in the previous sections, along with the decay fractions  $P(\text{decay})$  from the fit to the invariant mass distribution, and the mass distributions  $f(m|\text{decay})$  for each decay class. This is the likelihood that is used to fit to the proper time distribution of the toy datasets in the following section.

## 6.5 Results from Fit to Toy Data

### 6.5.1 Validation of Proper Time Fitter

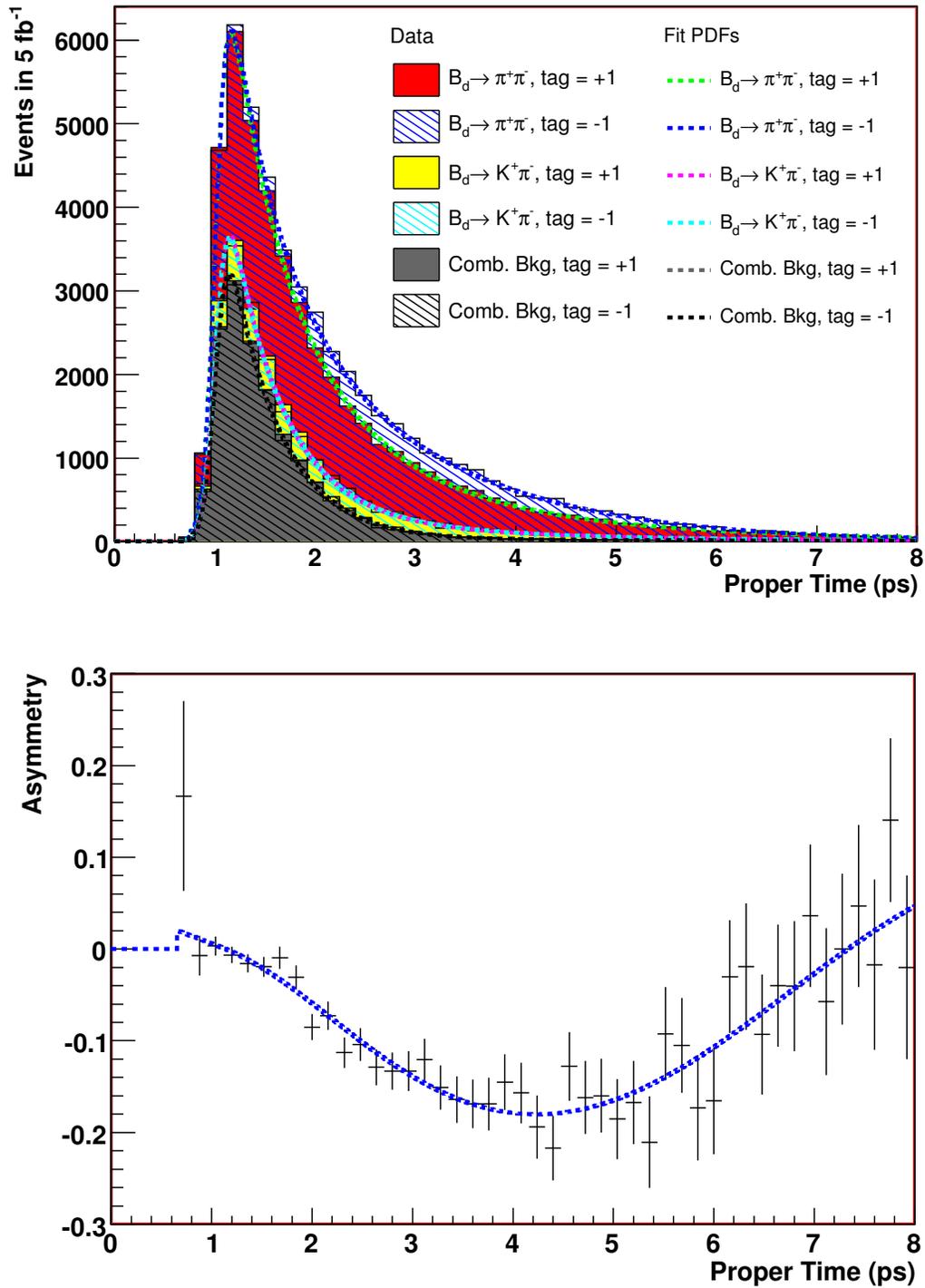
The method used to perform the fit to the proper time distribution can be tested for any inherent bias by performing fits to simulated toy data samples that only contain  $B_d \rightarrow \pi^+\pi^-$ ,  $B_d \rightarrow K^+\pi^-$  and combinatoric background events, i.e. the decay classes that are considered

in the fit PDF. A large sample size is used to minimise the size of the fit errors, so that any bias can be more easily seen. A sample size of 100,000 events was chosen. This corresponds to the number of events for these modes that is expected with around  $5 \text{ fb}^{-1}$  of data, however this sample does not exactly correspond to the dataset that would be observed, as some backgrounds are omitted. The purpose of fitting to this large sample is not to estimate the sensitivity of LHCb to the  $\mathcal{CP}$  asymmetries with  $\simeq 5 \text{ fb}^{-1}$  of data, but rather to validate the method used here for the proper time fit and demonstrate that it is free of inherent bias.

A total of 300 such datasets were generated, with the average fractions of  $B_d \rightarrow \pi^+\pi^-$  and  $B_d \rightarrow K^+\pi^-$  events being 0.561 and 0.103 respectively. These datasets were then fitted with the fit PDF described in Sec. 6.4.2, with all parameters in the PDF, apart from the signal  $\mathcal{CP}$  asymmetries, fixed to their true values (i.e. the values used to generate the toy datasets). The decay fractions  $P(\text{decay})$  are fixed to the average (true) values of  $P(B_d \rightarrow \pi^+\pi^-) = 0.561$  and  $P(B_d \rightarrow K^+\pi^-) = 0.103$ . The signal  $\mathcal{CP}$  asymmetries,  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$ , are seeded in the signal time PDF with values that are close, but not identical, to the values in the toy data: 0.48 for  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and 0.51 for  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  (it was checked that using other seeds for  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  in the interval  $[-1.0, 1.0]$  did not affect the fit results). During the fit the  $\mathcal{CP}$  asymmetries, which are the only floating parameters, are both allowed to float in the interval  $[-10.0, 10.0]$ . Fitted values that violate the physical constraint  $(\mathcal{A}_{\mathcal{CP}}^{\text{dir}2} + \mathcal{A}_{\mathcal{CP}}^{\text{mix}2}) \leq 1$  (see the definitions (1.47)) can occur, as mathematically the fit is measuring the amplitude of an oscillation, which in principle can have any value.

All 300 fits converged successfully and returned a full and accurate covariance matrix. An example fit can be seen in the upper half of Fig. 6.5. The different components of the data are shown as cumulative histograms, with the events of each decay having tag  $q = +1(-1)$  represented by a solid (hatched) histogram (the data distributions with each tag are however *not* displayed cumulatively here). The projection of the PDFs for each decay class with a given tag have been superimposed onto each decay. The fit shown returned  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-) = 0.407 \pm 0.033$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-) = 0.655 \pm 0.024$ . The most relevant feature of Fig. 6.5 with regards to fitting for the  $B_d \rightarrow \pi^+\pi^-$   $\mathcal{CP}$  asymmetries is that the fit PDFs for each tagged subset of  $B_d \rightarrow \pi^+\pi^-$  events follow the data closely, i.e. the green dotted line (tag = +1) follows the solid red histogram, and the blue dotted line (tag = -1) follows the hatched blue histogram. For the other decay classes in the fit, the distributions for each tag are seen to be very similar to each other, as expected. The fit PDFs for each tag in these classes lie almost on top of each other, so that only one of them is visible. The fact that the blue histogram lies above the red histogram, i.e. the time-dependent  $\mathcal{CP}$  asymmetry is negative, for most of the time interval shown is due to  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$  being large and positive (see Eqn. 1.52).

Also shown in Fig. 6.5 is the corresponding proper time distribution of the fractional



**Figure 6.5:** A typical fit to the proper time distribution of a large simulated data sample containing signal and main backgrounds (top), and the corresponding fractional asymmetry distribution (bottom). See text for details.

Parameter	Input	Fitted Values Distribution		Pull Distribution		Fit Errors
		Mean	Width	Mean	Width	Mean
$\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$	0.38	$0.379 \pm 0.002$	$0.034 \pm 0.002$	$-0.01 \pm 0.06$	$1.00 \pm 0.05$	0.033
$\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$	0.61	$0.610 \pm 0.002$	$0.024 \pm 0.001$	$0.00 \pm 0.07$	$1.03 \pm 0.06$	0.024

**Table 6.5:** Summary of results from proper time fits to large simulated data samples containing signal and main backgrounds.

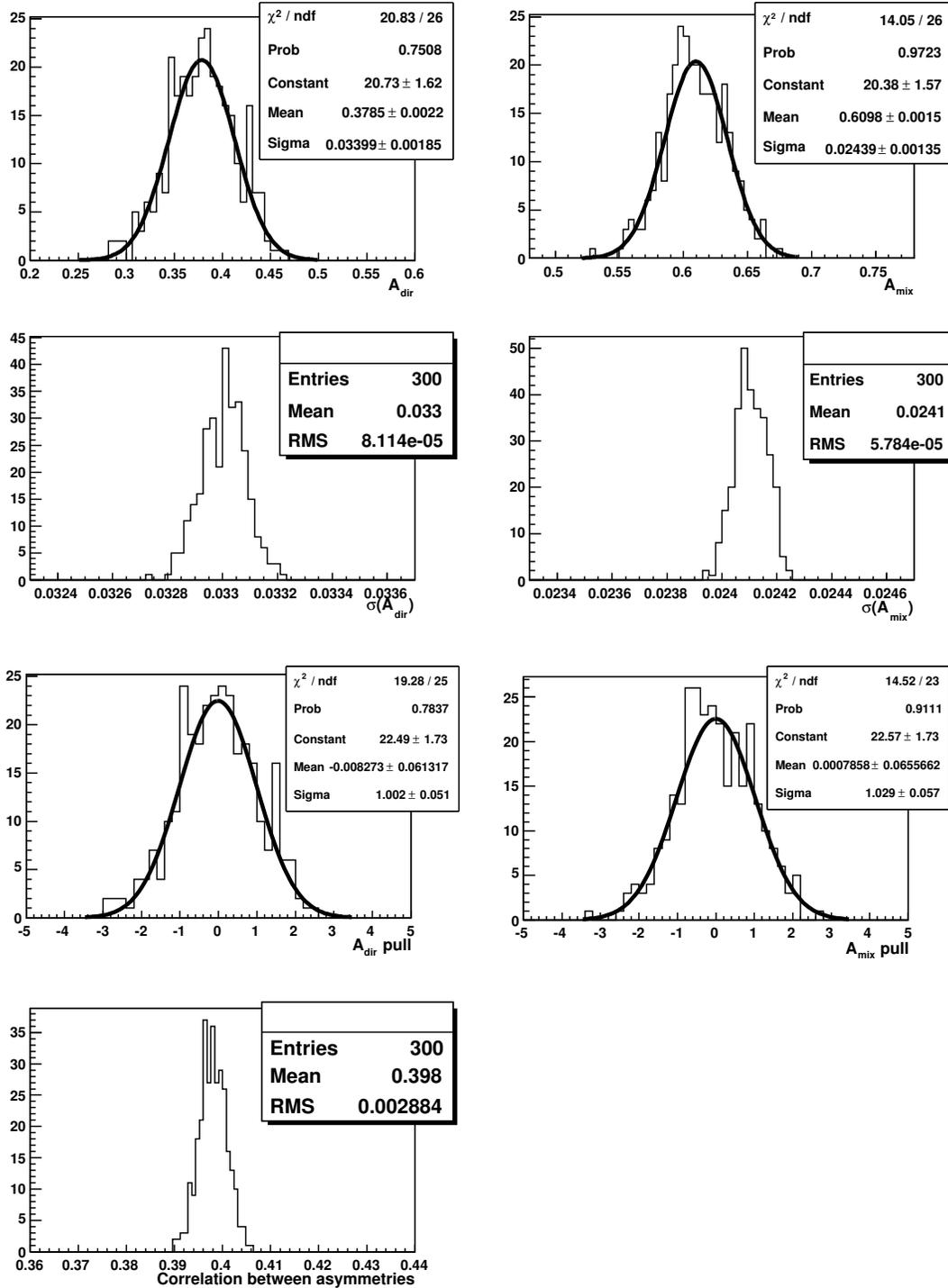
asymmetry (i.e. the difference in the number of events with  $q = 1$  and  $q = -1$ , divided by the sum), with the fractional asymmetry of the fit PDF overlaid in blue. The asymmetry distribution is shown only for illustration, it is *not* the distribution that is fitted. The fit is made to the proper time distribution (see Eqn. 6.8).

The distributions of the fit results for  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$  from the 300 toy datasets are shown in Fig. 6.6. The distributions of the fitted value, fit error and pull (i.e. residual divided by fit error) for each parameter are shown. A single Gaussian fit has been performed on the distributions of fitted values and pulls. The distribution of the (signed) correlation between the two  $\mathcal{CP}$  asymmetries in the fit is also shown. The results are summarised in Table 6.5. It can be seen from Fig. 6.6 and Table 6.5 that, for each fit parameter,

- The mean of the distribution of the fitted values is equal to the input value for the parameter,
- The mean of the pull distribution is compatible with 0, and the width with 1,
- The width of the distribution of the fitted values matches the mean of the distribution of the fit errors.

Hence it can be concluded that the method has successfully recovered the input values for both  $\mathcal{CP}$  asymmetries, with correctly estimated errors. From the uncertainty on the mean of the fitted values distribution, and also on the mean and width of the pull distribution, an upper limit of  $\simeq 0.1 \cdot \sigma(\mathcal{A}_{\mathcal{CP}})$  can be placed on the size of any inherent bias in the method. This corresponds to an absolute bias of  $\simeq 0.003$ , which is several times smaller than the statistical error achievable at LHCb during the initial phase of LHC operation (a naïve extrapolation of the fit errors observed here to  $10 \text{ fb}^{-1}$  would yield  $\sigma(\mathcal{A}_{\mathcal{CP}}) \simeq 0.02$ ).

Another feature of the fit results is that the error on  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  is considerably smaller than that on  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ . This arises through the combination of three features of the signal proper time distribution (see Eqn. 6.4): the turn-on point, the exponential decay and the sinusoidal terms. Since the typical turn-on point value is 1 ps, the vast majority of the signal events that pass the selection and the trigger have proper time values between 1 ps and 4 ps. At  $t \simeq 3.1$  ps,



**Figure 6.6:** Distributions of fit results for  $\mathcal{CP}$  asymmetries from proper time fits to large simulated data samples containing signal and main backgrounds. The first, second and third rows show respectively the distributions of fitted values, fitted errors, and pulls for  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$ . The fourth row shows the distribution of the correlation between the two  $\mathcal{CP}$  asymmetries in the fit.

$\Delta m_d t \simeq \pi/2$  so that  $\sin(\Delta m_d t) \simeq 1$  and  $\cos(\Delta m_d t) \simeq 0$ . So in the region where most of the data lies, the magnitude of the term multiplying  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  is larger than the magnitude of the term multiplying  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ . This can be seen in the asymmetry plot in Fig. 6.5: the points with  $t \lesssim 4$  ps, which have higher statistics, have smaller error bars and are mainly below the x-axis, reflecting the dominance of the  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  term in this time region. Hence the fit determines  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  more accurately than  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ . If the proper time acceptance behaviour of data differs from that used in this study, then the ratio of the errors on  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  may change.

The distribution of the (signed) correlation between the two  $\mathcal{CP}$  asymmetries in the fit, shown in Fig. 6.6, is narrowly peaked around +0.40. A non-negligible correlation is expected, as the sum of the sine and cosine terms in the decay rate (see Eqn. 6.4) can be considered as a single sinusoidal function whose amplitude and phase both depend on (different combinations of)  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ .

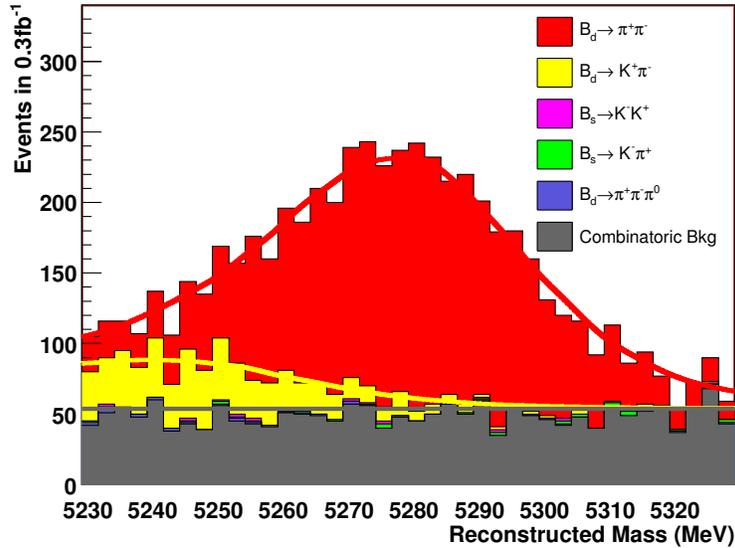
With the method having been seen to carry no (or very little) inherent bias, any biases found from fits to “realistic” data that include the minor backgrounds can be ascribed to the presence of those backgrounds and not to the method itself.

## 6.5.2 Mass and Proper Time Fits to Realistic Data

The performance of the mass and proper time fits on smaller data samples, corresponding to the number of triggered, selected, and tagged events expected in  $0.3 \text{ fb}^{-1}$  of data, was studied. Apart from the  $B_d \rightarrow \pi^+\pi^-$  signal and the dominant backgrounds ( $B_d \rightarrow K^+\pi^-$  and combinatoric background), these data samples also include the “minor” backgrounds  $B_d \rightarrow \pi^+\pi^-\pi^0$  (representing three-body backgrounds),  $B_s \rightarrow K^+K^-$  and  $B_s \rightarrow K^-\pi^+$ . However the minor backgrounds are not modelled in the fit. The results of the previous section mean that any bias that is found on the extracted values of the  $\mathcal{CP}$  asymmetries will be due to the effect of these minor backgrounds, and not the fit method itself.

A total of 1000 toy datasets were generated as described in Sec. 6.3, with the average number of each events set as given in Table 6.4 (i.e. 6100 total events, of which an average of 3422 are signal events).

As described in Sec. 6.4, the fractions of each decay class,  $P(\text{decay})$ , are first measured by fitting the invariant mass distribution. The time fits are performed with the decay fractions fixed not to the true values but to the mean values found by the invariant mass fit. This takes any biasing effect of the minor backgrounds on the mass fit into account.



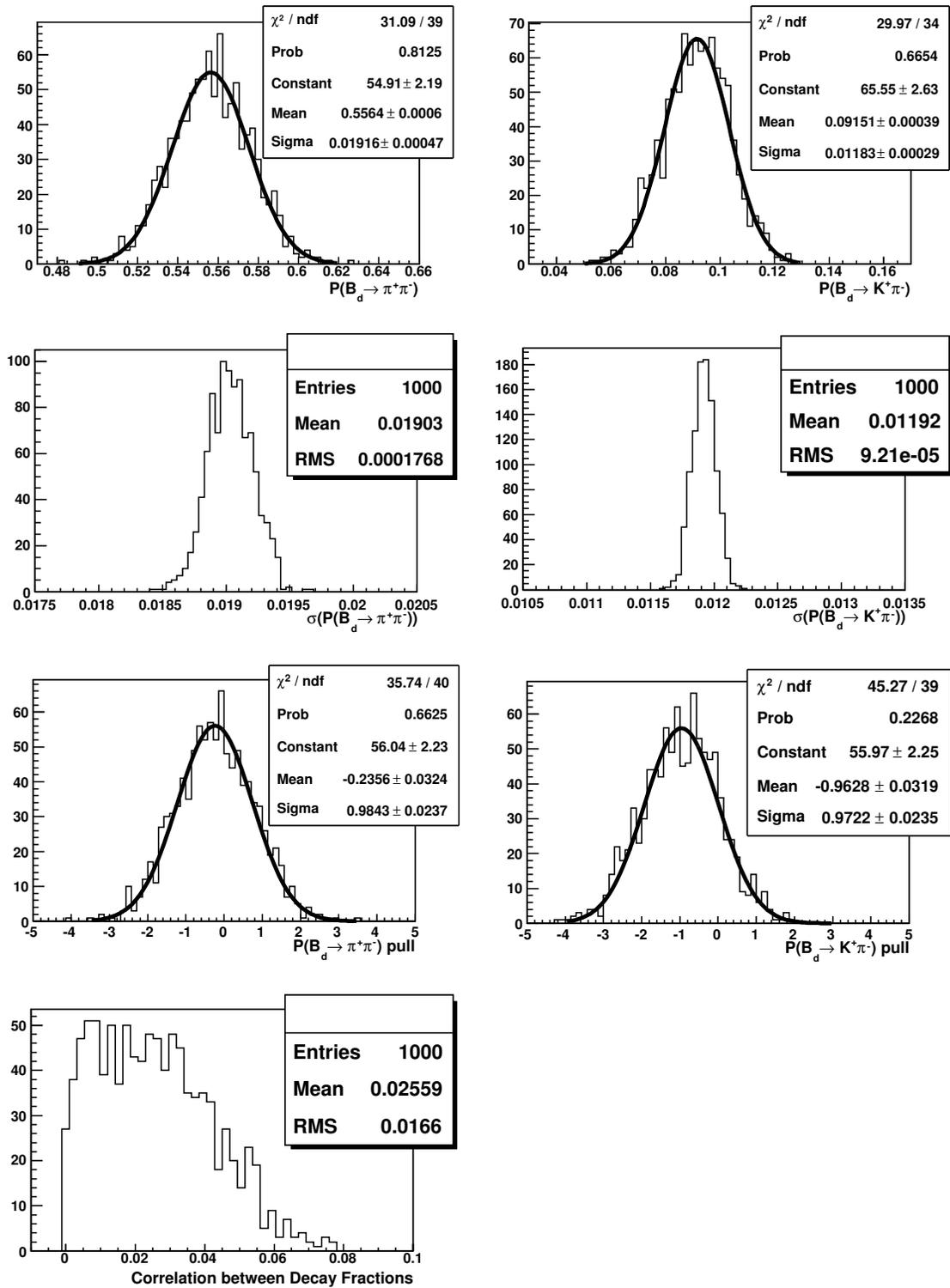
**Figure 6.7:** A typical fit to the mass distribution of a simulated  $0.3 \text{ fb}^{-1}$  data sample containing signal and all considered backgrounds. See the text for details.

### 6.5.2.1 Mass Fits

The mass distributions of the toy datasets described above were fitted with the total mass PDF  $f(m)$  as discussed in Sec. 6.4.1, with only the signal and the two dominant backgrounds being included in the total PDF. The mass distribution,  $f(m|\text{decay})$ , for each modelled decay class matches that used in the toy data generation. The two independent decay fractions,  $P(B_d \rightarrow \pi^+\pi^-)$  and  $P(B_d \rightarrow K^+\pi^-)$ , are each seeded in the fit with the value 0.33, and floated in the interval  $[0,1]$ . As the  $f(m|\text{decay})$  used in the fit match the generated distributions, the only possible source of bias in the extracted values of  $P(B_d \rightarrow \pi^+\pi^-)$  and  $P(B_d \rightarrow K^+\pi^-)$  is the presence of the minor backgrounds.

All 1000 fits converged successfully and returned a full and accurate covariance matrix. An example fit is shown in Fig. 6.7. The six different components of the data are shown as cumulative histograms. The three different components of the fit PDF are shown as smooth lines. It can be seen that the minor backgrounds have little effect on the total distribution. The fitted values for this dataset are  $P(B_d \rightarrow \pi^+\pi^-) = 0.556 \pm 0.019$  and  $P(B_d \rightarrow K^+\pi^-) = 0.092 \pm 0.012$ .

The distributions of the fit results for  $P(B_d \rightarrow \pi^+\pi^-)$  and  $P(B_d \rightarrow K^+\pi^-)$  from the 1000 toy datasets are shown in Fig. 6.8. The distributions of the fitted value, fit error and pull for each parameter are shown. A single Gaussian fit has been performed on the distributions of fitted values and pulls. The distribution of the (signed) correlation between the two decay fractions in the fit is also shown. The results are summarised in Table 6.6.



**Figure 6.8:** Distributions of fit results for decay fractions from mass fits to simulated  $0.3 \text{ fb}^{-1}$  data sample containing signal and all considered backgrounds. The first, second and third rows show respectively the distributions of fitted values, fitted errors, and pulls for the fractions of  $B_d \rightarrow \pi^+\pi^-$  and  $B_d \rightarrow K^+\pi^-$ . The fourth row shows the distribution of the correlation between the two decay fractions in the fit.

Parameter	Input	Fitted Values Distribution		Pull Distribution		Fit Errors
		Mean	Width	Mean	Width	Mean
$P(B_d \rightarrow \pi^+\pi^-)$	0.561	$0.556 \pm 0.001$	$0.019 \pm 0.001$	$-0.24 \pm 0.03$	$0.98 \pm 0.02$	0.019
$P(B_d \rightarrow K^+\pi^-)$	0.103	$0.092 \pm 0.001$	$0.012 \pm 0.001$	$-0.96 \pm 0.03$	$0.97 \pm 0.02$	0.012

**Table 6.6:** Summary of results from mass fits to simulated  $0.3 \text{ fb}^{-1}$  data samples containing signal and all considered backgrounds.

Figure 6.8 and Table 6.6 show that the presence of the minor backgrounds causes a small negative bias in the extracted values for the decay fractions. The bias is  $\simeq 0.005$  for  $P(B_d \rightarrow \pi^+\pi^-)$  and  $\simeq 0.011$  for  $P(B_d \rightarrow K^+\pi^-)$ . For  $P(B_d \rightarrow \pi^+\pi^-)$  the systematic error is several times smaller than the statistical error with this data sample, but for  $P(B_d \rightarrow K^+\pi^-)$  they are roughly equal, and as more data is collected and the statistical error decreases, the systematic error will become dominant. Future studies should therefore investigate the inclusion of the minor backgrounds in the fit PDF, to remove this bias.

The effect of the bias on the extraction of the  $\mathcal{CP}$  asymmetries in the proper time fit will be explored in the next section. This bias has shifted the means of the pull distributions away from 0, but the widths of the pull distributions are still compatible with 1, showing that the mass fit is determining the errors correctly.

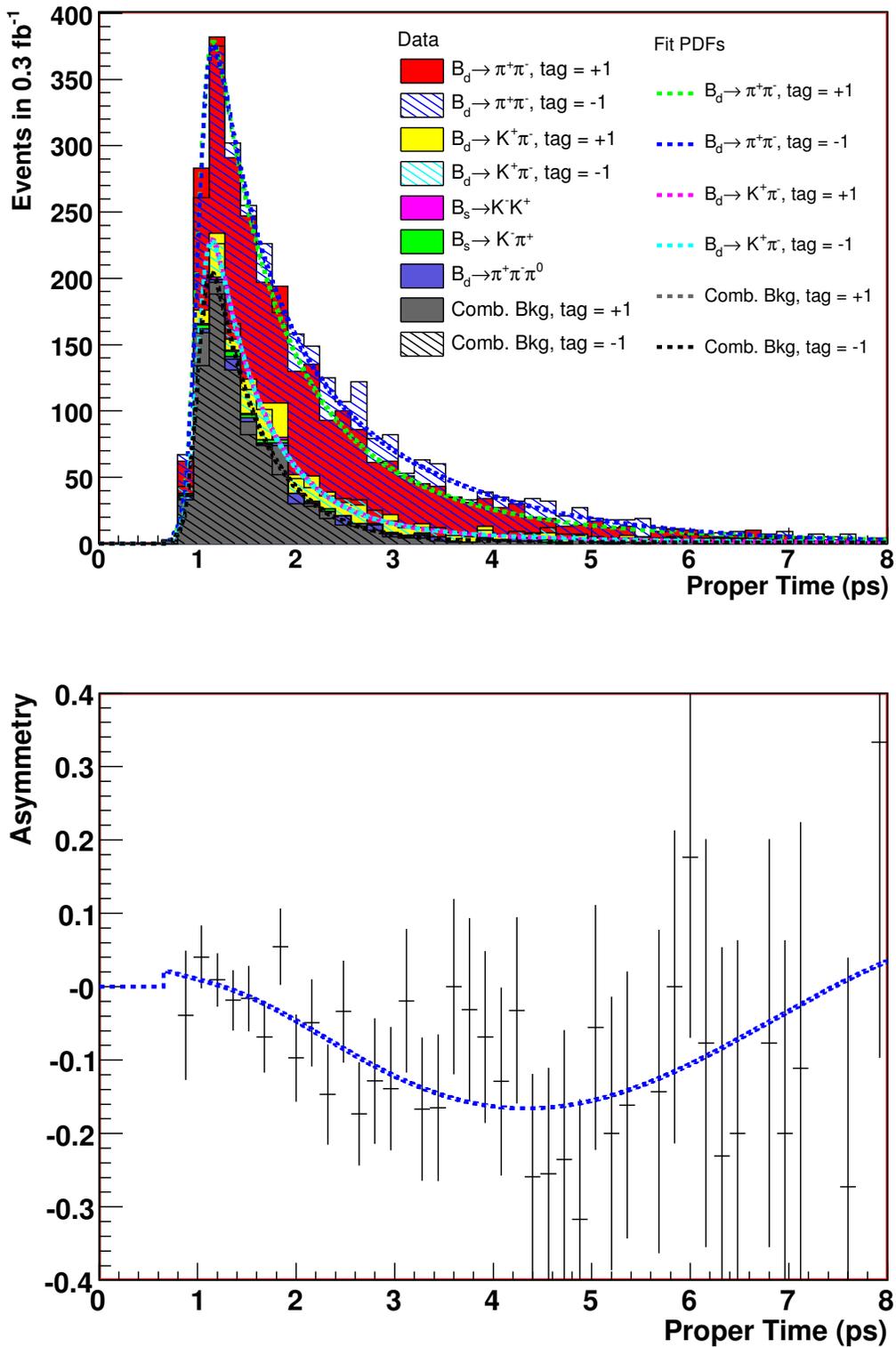
Finally, the correlation between the two decay fractions in the fit is seen to be very low. This is expected as the shapes of the mass distributions for the three main decay classes are very distinct from each other (see Fig. 6.7).

### 6.5.2.2 Proper Time Fits

The fit PDF used is the same as that used for the fits described in Sec. 6.5.1, except that the decay fractions are fixed not to their true values but to the mean values found by the mass fit in the previous section, i.e.  $P(B_d \rightarrow \pi^+\pi^-) = 0.556$  and  $P(B_d \rightarrow K^+\pi^-) = 0.092$ .

As with the mass fit, all 1000 fits converged successfully and returned a full and accurate covariance matrix. An example fit is shown in Fig. 6.9. As expected, the proper time distributions are very similar to those seen in Fig. 6.5, except that the minor backgrounds now appear in the data. The distributions for the minor backgrounds are not displayed separately by tag to allow them to be more clearly identified. It can be seen that the amount of combinatoric background is slightly overestimated, as indicated by the small negative biases on  $P(B_d \rightarrow \pi^+\pi^-)$  and  $P(B_d \rightarrow K^+\pi^-)$  from the mass fit. The fit shown returned  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-) = 0.406 \pm 0.133$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-) = 0.572 \pm 0.096$ .

The corresponding proper time distribution of the fractional asymmetry is also shown in Fig. 6.9. The distribution is similar to that shown in Fig. 6.5, except that the error bars are



**Figure 6.9:** A typical fit to the proper time distribution of a simulated  $0.3 \text{ fb}^{-1}$  data sample containing signal and all considered backgrounds (top), and the corresponding fractional asymmetry distribution (bottom). See text for details.

Parameter	Input Value	Fitted Values Distribution		Pull Distribution		Fit Errors
		Mean	Width	Mean	Width	Mean
$\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$	0.38	$0.369 \pm 0.005$	$0.135 \pm 0.003$	$-0.07 \pm 0.03$	$1.01 \pm 0.03$	0.134
$\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$	0.61	$0.598 \pm 0.003$	$0.093 \pm 0.002$	$-0.12 \pm 0.03$	$0.96 \pm 0.02$	0.097

**Table 6.7:** Summary of results from proper time fits to simulated  $0.3 \text{ fb}^{-1}$  data samples containing signal and all considered backgrounds.

far larger due to the lower statistics involved here.

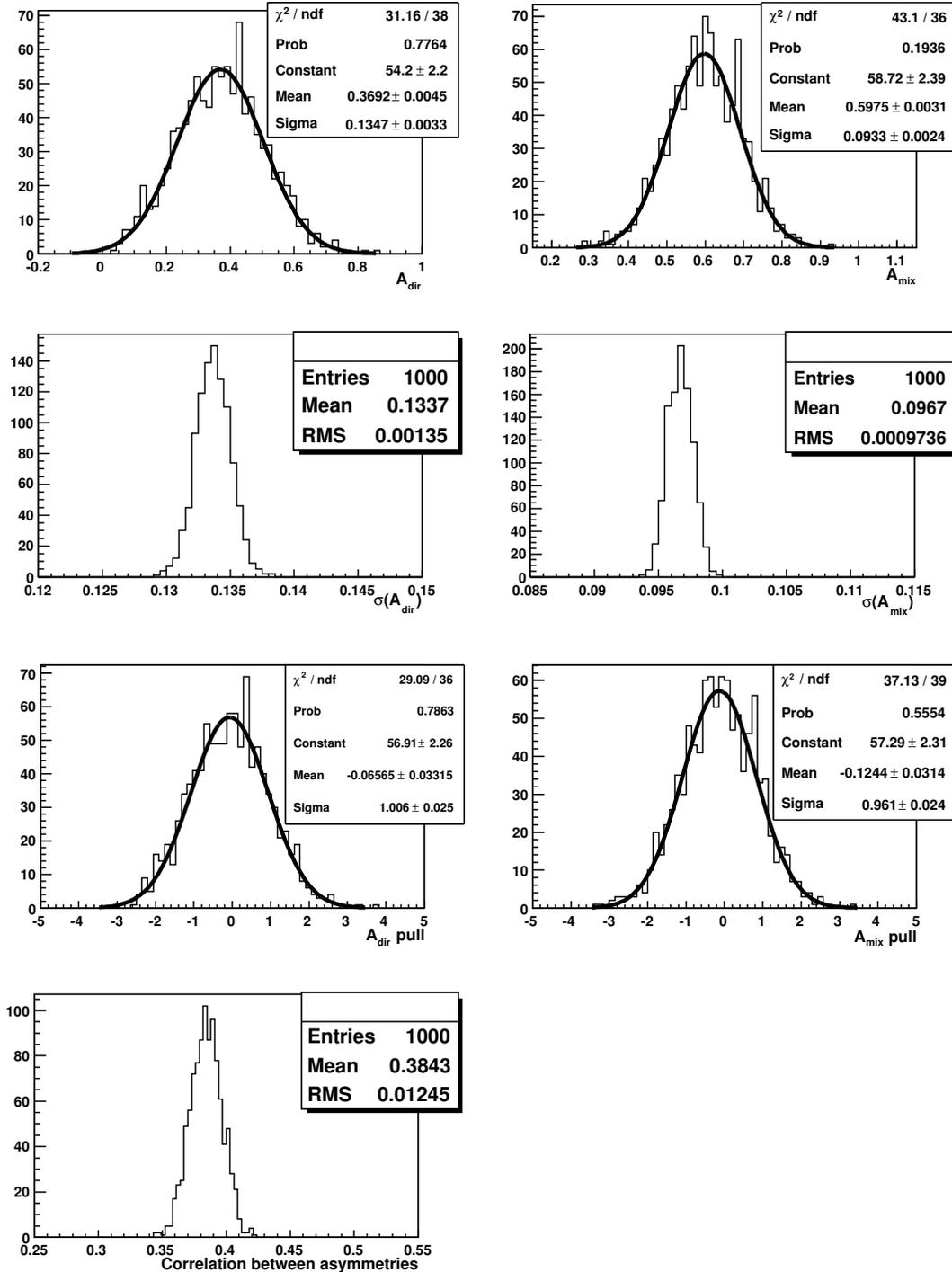
The distributions of the fit results for  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$  from the 1000 toy datasets are shown in Fig. 6.10. The distributions of the fitted value, fit error and pull for each parameter are shown. A single Gaussian fit has been performed on the distributions of fitted values and pulls. The distribution of the (signed) correlation between the two  $\mathcal{CP}$  asymmetries in the fit is also shown. The results are summarised in Table 6.7. It can be seen that the bias introduced by the minor backgrounds is small and negative; around  $-0.01$  for both  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ . This bias is negligible compared with the statistical errors from this small sample, which are  $\simeq 0.13$  for  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\simeq 0.10$  for  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ . Due to this small bias, the pull distributions do not have mean 0, but their widths are still very close to 1, indicating that the fit is calculating the errors correctly. As was seen for the results in Sec. 6.5.1, and for the same reasons, the error on  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  here is again considerably smaller than that on  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ . The correlation distribution has a mean value that is very similar to that for the fit to the large data samples (see Fig. 6.6), but the spread is considerably larger due to the lower statistics involved.

Before a complete estimate of the sensitivity to the  $\mathcal{CP}$  asymmetries with the  $0.3 \text{ fb}^{-1}$  data sample can be made, further sources of systematic error, beyond that associated with the minor backgrounds, need to be considered.

### 6.5.3 Sources of Systematic Error

The presence of minor backgrounds that are not modelled in the fit has been seen to introduce a small systematic error on the extraction of the  $\mathcal{CP}$  asymmetries. Another possible source of systematic error is an incorrect determination of one of the parameters that is held fixed in the proper time fit. The physics parameters that are fixed in the proper time fit —  $\tau_{B_d}$ ,  $\Delta m_d$  and  $\mathcal{A}_{\mathcal{CP}}(B_d \rightarrow K^+\pi^-)$  — are tightly constrained by current experimental measurements, whose precision will be improved by LHCb. Hence none of the physics parameters are expected to be a source of significant systematic uncertainty.

The other fixed parameters in the proper time fit are experimental parameters — the



**Figure 6.10:** Distributions of fit results for  $\mathcal{CP}$  asymmetries from proper time fits to simulated  $0.3 \text{ fb}^{-1}$  data samples containing signal and all considered backgrounds. The first, second and third rows show respectively the distributions of fitted values, fitted errors, and pulls for  $A_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $A_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$ . The fourth row shows the distribution of the correlation between the two  $\mathcal{CP}$  asymmetries in the fit.

Parameter	Fit $\omega_{\text{tag}}$	Fitted Values Distribution		Pull Distribution		Fit Errors
		Mean	Width	Mean	Width	Mean
$\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$	35.0%	$0.365 \pm 0.006$	$0.120 \pm 0.004$	$-0.15 \pm 0.05$	$0.96 \pm 0.04$	0.129
$\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$	35.0%	$0.587 \pm 0.004$	$0.087 \pm 0.003$	$-0.29 \pm 0.05$	$0.96 \pm 0.04$	0.093
$\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$	36.0%	$0.372 \pm 0.007$	$0.142 \pm 0.005$	$-0.03 \pm 0.05$	$1.04 \pm 0.04$	0.139
$\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$	36.0%	$0.613 \pm 0.005$	$0.099 \pm 0.004$	$0.02 \pm 0.05$	$0.99 \pm 0.04$	0.100

**Table 6.8:** Summary of results from proper time fits to simulated  $0.3 \text{ fb}^{-1}$  data samples containing signal and all considered backgrounds, with incorrect estimation of  $\omega_{\text{tag}}$  in the fit PDF.

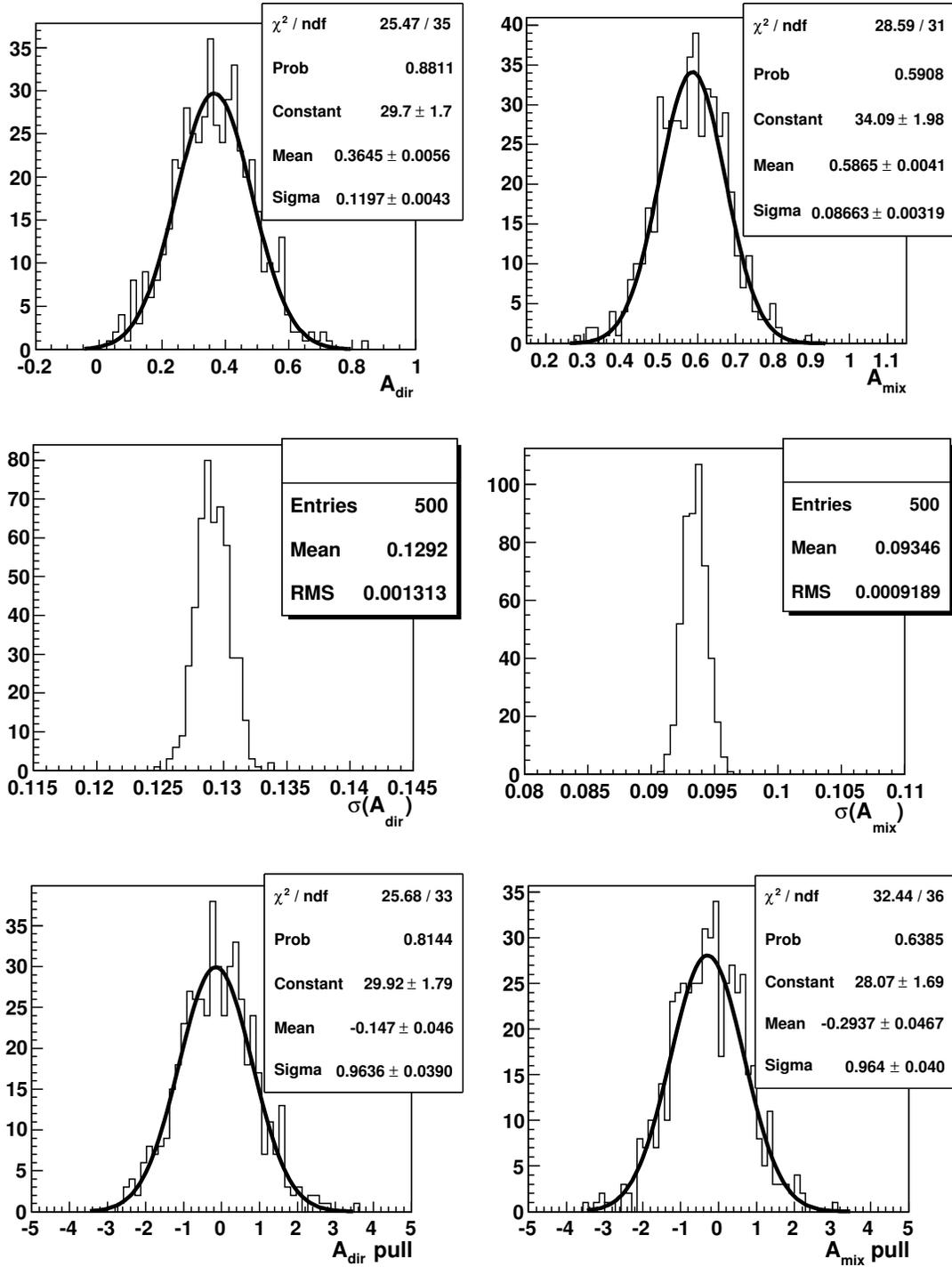
mistag rate for  $B_d$ , and the parameters relating to the proper time distribution for the combinatoric background. These will be measured on data using the methods mentioned in Sec. 6.4.2, however these determinations will have some uncertainty associated with them. As the coefficients of the oscillations in the decay rate are the product of  $(1 - 2\omega_{\text{tag}})$  and a  $\mathcal{CP}$  asymmetry (see Eqn. 6.4), an incorrect determination of the mistag rate is likely to have the most direct impact on the extracted values of the  $\mathcal{CP}$  asymmetries. The size of this effect can be estimated by carrying out fits to the datasets used in Sec. 6.5.2, but with the value of  $\omega_{\text{tag}}$  in the fit PDF fixed to a different value from that used to generate the toy data, which is 35.5%. Since the mistag rate for  $B_d$  can be directly determined using  $B_d \rightarrow K^+\pi^-$  (see Sec. 4.1.3.1), its uncertainty should be relatively small. Here the effect of a 0.5% change in the value of  $\omega_{\text{tag}}$  used in the fit will be estimated.

The 1000 datasets used in Sec. 6.5.2 were divided into two sets of 500. One set was fitted with a fit PDF having  $\omega_{\text{tag}} = 35.0\%$ , and the other with a fit PDF having  $\omega_{\text{tag}} = 36.0\%$ . The other parameters (including the decay fractions) in the fit were fixed to the same values used for the fits described in Sec. 6.5.2.2. All 1000 fits converged successfully and returned a full and accurate covariance matrix.

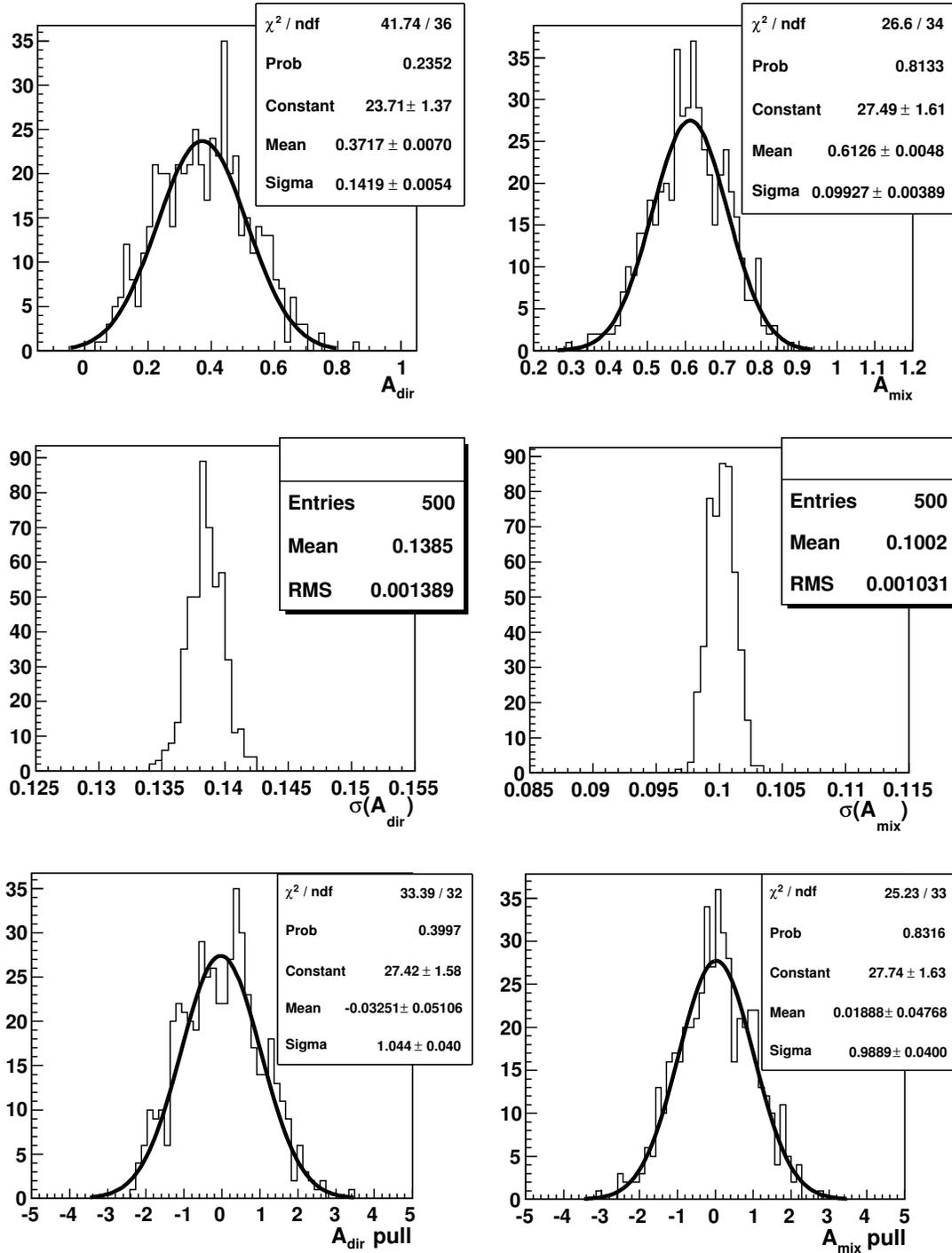
The distributions of the results from the fits to these samples are shown in Figs. 6.11 and 6.12 (the distributions of the correlation between the asymmetries are not shown as they are almost identical to the distribution in Fig. 6.10). A summary is given in Table 6.8.

Comparing these results with those from the previous section, it can be seen that underestimating (overestimating) the mistag rate decreases (increases) the measured values of the  $\mathcal{CP}$  asymmetries by around 0.01. This is the expected behaviour, because e.g. an underestimation of the mistag causes more dilution of the oscillations to be observed than is expected, which is compensated for by lower values of the  $\mathcal{CP}$  asymmetries (see Eqn. 6.4).

In the overestimation case, there is a fortunate cancellation with the systematic arising from the minor backgrounds, and the final results come out unbiased. However the fit errors



**Figure 6.11:** Distributions of fit results for  $\mathcal{CP}$  asymmetries from proper time fits to simulated  $0.3 \text{ fb}^{-1}$  data samples containing signal and all considered backgrounds, with  $\omega_{\text{tag}}$  underestimated by 0.5% in the fit PDF. The first, second and third rows show respectively the distributions of fitted values, fitted errors, and pulls for  $A_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $A_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$ .



**Figure 6.12:** Distributions of fit results for  $\mathcal{CP}$  asymmetries from proper time fits to simulated  $0.3 \text{ fb}^{-1}$  data samples containing signal and all considered backgrounds, with  $\omega_{\text{tag}}$  overestimated by 0.5% in the fit PDF. The first, second and third rows show respectively the distributions of fitted values, fitted errors, and pulls for  $A_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  and  $A_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$ .

are seen to increase slightly when the mistag is overestimated.

Due to the cancellation noted above, the means of the pull distributions in the overestimation case are compatible with zero. In the underestimation case they are significantly negative. In both cases, the widths of the pull distributions are compatible with 1, indicating that the fit is still correctly estimating the errors.

#### 6.5.4 Sensitivity to $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$ and $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$

The complete sensitivity to the  $\mathcal{CP}$  asymmetries in  $B_d \rightarrow \pi^+\pi^-$  with  $0.3 \text{ fb}^{-1}$  of data can now be estimated using the results from the previous sections:

$$\begin{aligned}\sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{dir}}) &= 0.135(\text{stat}) \oplus 0.011(\text{bkgs}) \oplus 0.004(\omega_{\text{tag}}) = 0.135(\text{stat}) \oplus 0.012(\text{syst}), \\ \sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{mix}}) &= 0.093(\text{stat}) \oplus 0.012(\text{bkgs}) \oplus 0.013(\omega_{\text{tag}}) = 0.093(\text{stat}) \oplus 0.018(\text{syst}),\end{aligned}$$

where the biases arising from the minor backgrounds and the mistag determination have been combined in quadrature to give an estimate for the total systematic error.

Although the total errors as calculated here are slightly larger than that those on the individual measurements from Babar and Belle (see Eqn. 1.68), a measurement with such an error, expected at LHCb after only the initial 2009–2010 data-taking run, can still help to resolve the disagreement between the Babar and Belle measurements of  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ .

As the data sample grows beyond  $0.3 \text{ fb}^{-1}$ , the statistical error will fall while the systematic error will remain the same, assuming that the analysis continues not to model the minor backgrounds and the estimation of the mistag rate does not improve. Under these assumptions, the total errors with one year of data at nominal luminosity ( $2 \text{ fb}^{-1}$ ) would be:

$$\begin{aligned}\sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{dir}}) &= 0.052(\text{stat}) \oplus 0.012(\text{syst}), \\ \sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{mix}}) &= 0.036(\text{stat}) \oplus 0.018(\text{syst}),\end{aligned}$$

These values are similar to the  $2 \text{ fb}^{-1}$  sensitivities reported in [56], obtained using the simultaneous fit strategy (see Sec. 6.1). With this amount of data, the disagreement between the B-Factories will be resolved beyond doubt.

Arguably the most important physics measurement that can be made using  $B \rightarrow h^+h'^-$  decays is the measurement of the CKM angle  $\gamma$  utilising the direct and mixing-induced  $\mathcal{CP}$  asymmetries in  $B_s \rightarrow K^+K^-$  (see Sec. 1.4.4). Assuming that the sensitivity attained for the  $\mathcal{CP}$  asymmetries in  $B_s \rightarrow K^+K^-$  is similar to that found for those in  $B_d \rightarrow \pi^+\pi^-$  by the current study, a value for  $\gamma$  can be calculated with an error of around  $7^\circ$ , using  $2 \text{ fb}^{-1}$  of LHCb data [56].

### 6.5.5 Sensitivity to $\sin 2\alpha_{\text{eff}}$

In the absence of penguin amplitudes, a measurement of  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)$  would directly give a value for  $\sin 2\alpha$ . However the presence of penguin amplitudes causes  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(\pi^+\pi^-)$  to be nonzero, and the measured quantity becomes the effective physics parameter  $\sin 2\alpha_{\text{eff}}$  (see Sec. 1.4.3). From the definition

$$\sin 2\alpha_{\text{eff}} \equiv \frac{\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(\pi^+\pi^-)}{\sqrt{1 - \mathcal{A}_{\mathcal{CP}}^{\text{dir}2}(\pi^+\pi^-)}}, \quad (6.10)$$

the uncertainties  $\sigma_{\text{dir}}$  and  $\sigma_{\text{mix}}$  on  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  given in the previous section can be translated into an uncertainty  $\sigma_{\text{sin eff}}$  on  $\sin 2\alpha_{\text{eff}}$ :

$$(\sigma_{\text{sin eff}})^2 = \frac{\mathcal{A}_{\mathcal{CP}}^{\text{dir}2} \cdot \mathcal{A}_{\mathcal{CP}}^{\text{mix}2}}{(1 - \mathcal{A}_{\mathcal{CP}}^{\text{dir}2})^3} \cdot \sigma_{\text{dir}}^2 + \frac{1}{1 - \mathcal{A}_{\mathcal{CP}}^{\text{dir}2}} \cdot \sigma_{\text{mix}}^2 + 2 \cdot \frac{\mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cdot \mathcal{A}_{\mathcal{CP}}^{\text{mix}}}{(1 - \mathcal{A}_{\mathcal{CP}}^{\text{dir}2})^2} \cdot \rho_{\text{dir,mix}} \cdot \sigma_{\text{dir}} \cdot \sigma_{\text{mix}}, \quad (6.11)$$

where  $\rho_{\text{dir,mix}}$  is the correlation coefficient between  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$  (see Fig. 6.10).

The input values of  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}} = 0.38$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}} = 0.61$  give an input value of  $\sin 2\alpha_{\text{eff}} = 0.659$ . Using the values found in Sec. 6.5.2.2 for the mean values of  $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$  and  $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ , the statistical errors  $\sigma_{\text{dir(mix)}}$  and  $\rho_{\text{dir,mix}}$ , the resulting measurement of  $\sin 2\alpha_{\text{eff}}$  with  $0.3 \text{ fb}^{-1}$  of data is:

$$\sin 2\alpha_{\text{eff}} = 0.643 \pm 0.119.$$

This sets the scale of the expected error on  $\sin 2\alpha$  using the isospin method, and is the minimum possible error in the ideal case of no further uncertainty being introduced by the inclusion of measurements from the decays  $B^+ \rightarrow \pi^+\pi^0$  and  $B_d \rightarrow \pi^0\pi^0$ .

## 6.6 Conclusions

A study has been carried out to estimate the sensitivity of LHCb to the direct and mixing-induced  $\mathcal{CP}$  asymmetries for the decay  $B_d \rightarrow \pi^+\pi^-$ , via a study of the time-dependent  $\mathcal{CP}$  asymmetry distribution. Cuts on particle identification likelihoods and invariant mass are used to reduce the background level from other  $B \rightarrow h^+h'^-$  decays. The proper time fit method was validated using large data samples containing only the decays that are modelled in the fit PDF. No evidence of any biases was found.

Smaller sets of toy data corresponding to the expected yields with  $0.3 \text{ fb}^{-1}$  of data, were then generated. These datasets included the minor backgrounds of  $B_s \rightarrow K^+K^-$ ,  $B_s \rightarrow K^-\pi^+$  and three-body backgrounds, but were fitted taking only the  $B_d \rightarrow \pi^+\pi^-$  signal and the dominant backgrounds —  $B_d \rightarrow K^+\pi^-$  and combinatoric background — into account. The fit was performed in two stages: firstly the relative fractions of the signal and

background decay classes are found from the invariant mass distribution. These were found to be biased by a small absolute amount of 0.005–0.01. This bias was taken into account when the decay fractions were used as fixed parameters fit to the proper time distribution. The resulting measured values for the  $\mathcal{CP}$  asymmetries displayed a small bias of  $\simeq 0.01$ , which, due to the results from the fitter validation, can be ascribed to the influence of the minor backgrounds.

An incorrect determination of the mistag rate  $\omega_{\text{tag}}$  was considered as the most likely source of additional error. It was found that an incorrect determination of  $\omega_{\text{tag}}$  by 0.5% contributes a further systematic error of  $\simeq 0.01$ .

The expected total error on the  $\mathcal{CP}$  asymmetries using  $0.3 \text{ fb}^{-1}$  of data is found to be:

$$\begin{aligned}\sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{dir}}) &= 0.135(\text{stat}) \oplus 0.012(\text{syst}), \\ \sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{mix}}) &= 0.093(\text{stat}) \oplus 0.018(\text{syst}).\end{aligned}$$

These are comparable to the uncertainties on the current experimental measurements. Clearly, LHCb can expect to supersede these measurements as more data is collected. The corresponding measurement of the effective physics parameter  $\sin 2\alpha_{\text{eff}}$  would be:

$$\sin 2\alpha_{\text{eff}} = 0.643 \pm 0.119,$$

for an input value of  $\sin 2\alpha_{\text{eff}} = 0.659$ . The study described here could be continued by considering further sources of systematic error, such as an incorrect determination of the proper time distribution for the combinatorial background or of the proper time acceptance turning points. Furthermore the known systematic error arising from the minor backgrounds could be tackled by including these backgrounds in the fit PDF. The addition of  $B_s$  decays to the fit PDF would require incorporating a proper time resolution model into the fit PDFs, and a strategy for taking the (currently poorly measured)  $\mathcal{CP}$  asymmetries in these channels into account would need to be developed. The natural longer-term development of the fitter described here is to enable a fit to the  $\mathcal{CP}$  asymmetries for  $B_s \rightarrow K^+K^-$ , which can then be combined with the  $B_d \rightarrow \pi^+\pi^-$  asymmetries to extract a value for  $\gamma$ . However this will require careful consideration of the possible contamination of each selected sample with the other decay of interest.

# Chapter 7

## Conclusions and Outlook

### 7.1 Summary

This thesis began by presenting a review of the Standard Model (SM) of particle physics, with particular focus on the SM description of quark mixing, the CKM mechanism, and how the phenomenon of  $\mathcal{CP}$  violation is included in this mechanism. Several ways in which studies of hadronic decays of  $B$  hadrons can constrain the CKM mechanism, and possibly reveal new effects beyond the SM, were discussed.

The LHCb experiment, which is about to begin physics data taking at the Large Hadron Collider (LHC), will make precision measurement of  $B$  decays, including hadronic decays. In chapter 2, a brief description was given of the LHC accelerator complex and its attending experiments, and the LHCb detector was described in depth.

The Ring Imaging Cherenkov (RICH) counters of the experiment provide excellent particle identification (PID) for charged hadrons over a wide momentum range. In chapter 2 the RICH counters were described in more detail than the other subdetectors, as the characterisation of the entire ensemble of the photon detectors for the RICHes, the Hybrid Photon Detectors (HPDs), was the subject of chapter 3.

All 557 HPDs produced for LHCb underwent thorough quality assurance testing and characterisation at dedicated test facilities. Their overall performance was found to be outstanding, with only 2.2% of HPDs judged to be unusable for the RICHes. The LHCb requirements and the contractual specifications are met and often exceeded in key areas, for example the dead pixel number, ion feedback probability and dark-count rate. The quantum efficiency of the HPD photocathode is found to be significantly higher than expectations, leading to increased photoelectron yield in the RICHes.

The measurement of the single photoelectron detection efficiency  $\eta$  of the HPD anode was described in detail. The efficiency was found to be  $\eta = (87.9 \pm 1.4)\%$  when a 25 ns

digital readout window was used. The efficiency was independent of the light input level. This value confirms that the production HPDs exceed the contractual requirement in this area, and is in agreement with previous measurements of  $\eta$ . It will be used to improve the simulation of the RICH detectors. The RICH detectors are now fully populated with the HPDs and are ready to participate in the study of the first collisions within LHCb.

The PID information from the RICHes is crucial for the study of hadronic  $B$  decays, as different final states can be distinguished from one another only if the final state particles are correctly identified. One interesting class of hadronic  $B$  decays is the two-body charmless decays, referred to within LHCb as  $B \rightarrow h^+h'^-$  decays.

The detector performance with respect to the proper time resolution for the decay  $B_s \rightarrow K^-K^+$  has to be measured in order to correctly extract the  $\mathcal{CP}$  asymmetries for that decay, which are needed for the measurement of  $\gamma$  using  $B \rightarrow h^+h'^-$  decays. The study presented in chapter 4 proposed a method which, without the use of information from Monte Carlo simulations, was found to be able to correctly extract the parameters of the proper time resolution model for  $B \rightarrow h^+h'^-$  decays. It was demonstrated that this can be done by measuring the proper time resolution for the flavour-specific decay  $B_s \rightarrow K^-\pi^+$ , via a fit to its flavour-tagged reconstructed proper time distribution.

A proper time resolution model was developed that accurately describes the proper time residual distributions in simulated data for four  $B_{d,s} \rightarrow \{\pi, K\}^+\{\pi, K\}^-$  channels. The model makes use of the per-event error on the reconstructed proper time. The two parameters of the model,  $GM$  and  $GS$ , are the effective scaling factors between the per-event error and the mean and width, respectively, of the resolution distribution. The values of  $GM$  and  $GS$  were found to be compatible with each other for all four decays.

The parameters obtained from these fits to the full simulation were used to guide a toy Monte Carlo study, which estimated how well the parameters of the model can be measured using data. The statistical uncertainty on the parameters of the model with  $2 \text{ fb}^{-1}$  of data is found to be  $\sigma_{GM(\text{stat})} = 0.21$  and  $\sigma_{GS(\text{stat})} = 0.30$ . The dominant systematic uncertainty is expected to be that associated with an incorrect determination of the mistag rate. This uncertainty affects  $GS$  but not  $GM$ , and has been estimated to be  $\sigma_{GS(\text{syst})} = 0.13$ , which is similar in size to the statistical uncertainty on  $GS$  after  $10 \text{ fb}^{-1}$  of data. With one nominal year of LHCb data, the fit to the resolution model parameters can be expected to constrain the resolution model for  $B_s \rightarrow K^-K^+$  sufficiently to provide useful input for the study of its time-dependent  $\mathcal{CP}$  asymmetry distribution, and hence for the measurement of  $\gamma$ .

While some  $B \rightarrow h^+h'^-$  decays, such as  $B_s \rightarrow K^-K^+$ , are experimentally well established, there are a number of such decays that remain unobserved. In the case of two-body charmless baryonic  $B$  decays, this is despite a range of theoretical calculations which predict

that these decays should already be experimentally accessible.

In chapter 5, a study was presented which examined the prospects of LHCb to discover new baryonic  $B$  decay modes, with particular focus on the experimentally most promising mode,  $B_d \rightarrow p\bar{p}$ . Measurements of the branching ratios for two-body charmless baryonic  $B$  decays can help to constrain, and discriminate between, different theoretical models that make certain approximations or assumptions when calculating decay amplitudes.

The LHCb inclusive selection for  $B \rightarrow h^+h'^-$  decays was adapted to produce an exclusive selection for the modes  $B_{d,s} \rightarrow p\bar{p}$ . By estimating the number of background events that would pass this selection, a  $5\sigma$  discovery of  $B_d \rightarrow p\bar{p}$  was found to be possible with only  $0.25 \text{ fb}^{-1}$  of nominal LHCb data, if the true  $B_d \rightarrow p\bar{p}$  branching fraction is close to the current experimental upper limit. Even if the true branching fraction is a factor of 5 lower than the present upper limit, a discovery can be made with about  $5 \text{ fb}^{-1}$  of data.

A detailed experimental study is required before any definitive statement can be made about the prospects of LHCb to observe other two-body charmless baryonic  $B$  decays. However, those channels involving a  $\Lambda$  baryon in the final state may also be worthy of study at LHCb.

Finally, chapter 6 described a study of the prospects of LHCb to measure the direct and mixing-induced  $\mathcal{CP}$  asymmetries for the decay  $B_d \rightarrow \pi^+\pi^-$ , via a study of the time-dependent  $\mathcal{CP}$  asymmetry distribution. A strategy was adopted for the event selection and fit which aimed to minimise systematic effects by reducing problematic, poorly constrained backgrounds to small levels.

The method used to fit to the  $B_d \rightarrow \pi^+\pi^-$  proper time distribution was examined for inherent bias using fits to large simulated data samples that contained only the decays that are modelled in the fit PDF. A tight limit was placed on any such bias.

Fits were then carried out to smaller sets of simulated data, to evaluate the sensitivity to the  $\mathcal{CP}$  asymmetries with  $0.3 \text{ fb}^{-1}$  of data, which is approximately the amount of data that is expected to be collected in the 2009–2010 data taking run of LHCb. These smaller data sets included backgrounds which are expected to be present in the sample of selected events for  $B_d \rightarrow \pi^+\pi^-$ , but which are not modelled in the fit PDF.

A two-stage fit was performed: firstly the relative fractions of the signal and the modelled backgrounds were obtained using a fit to the invariant mass distribution. There was a small bias of 0.005–0.01 on these fractions due to the unmodelled backgrounds. This shift was considered in the handling of the fit to the proper time distribution.

The measured values for the  $\mathcal{CP}$  asymmetries from this proper time fit carried a small bias of  $\simeq 0.01$ . This can be ascribed to the influence of the minor backgrounds, as the fit method was found to be free of intrinsic bias. An incorrect determination of the mistag rate

$\omega_{\text{tag}}$  was considered as an additional source of possible systematic error. It was found that an incorrect determination of  $\omega_{\text{tag}}$  by 0.5% introduces a bias of  $\simeq 0.01$  to each asymmetry.

The expected total sensitivity to the  $\mathcal{CP}$  asymmetries with  $0.3 \text{ fb}^{-1}$  of data was found to be:

$$\begin{aligned}\sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{dir}}) &= 0.135(\text{stat}) \oplus 0.012(\text{syst}), \\ \sigma_{\text{tot}}(\mathcal{A}_{\mathcal{CP}}^{\text{mix}}) &= 0.093(\text{stat}) \oplus 0.018(\text{syst}).\end{aligned}$$

This total expected error is comparable to the uncertainties on the current experimental measurements. As more data is collected and the statistical error falls, LHCb expects to come to dominate the world average values for these  $\mathcal{CP}$  asymmetries.

Using the values for the direct and mixing-induced  $\mathcal{CP}$  asymmetries expected with  $0.3 \text{ fb}^{-1}$  of data, the measurement of the effective physics parameter  $\sin 2\alpha_{\text{eff}}$  is found to be:

$$\sin 2\alpha_{\text{eff}} = 0.643 \pm 0.119,$$

for an input value of  $\sin 2\alpha_{\text{eff}} = 0.659$ . Apart from  $\sin 2\alpha_{\text{eff}}$ , a range of other interesting physics measurements can be made using  $B \rightarrow h^+h'^-$  decays. Notably, a combination of the measurements of the  $\mathcal{CP}$  asymmetries for  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^-K^+$  can be used to extract a value for the CKM angle  $\gamma$ .

## 7.2 Outlook

With the LHC due to begin physics-quality data taking in early 2010, the LHCb experiment is ready to analyse the first collisions to begin to calibrate and understand the detector. Calibration of the Particle Identification (PID) performance of the experiment, which is crucial for the study of hadronic  $B$  decays, will involve a detailed understanding of the Hybrid Photon Detectors (HPDs). This in turn depends on the knowledge gained from the comprehensive testing programme that all HPDs underwent at the Photon Detector Test Facilities (PDTFs) in Scotland. One HPD parameter that is required to be known in order to fully understand the PID performance is the single photoelectron detection efficiency of the HPD anode, which was measured in chapter 3 of this thesis.

Another aspect of the detector that needs to be well understood is the proper time resolution behaviour. Poor understanding of the proper time resolution has the potential to introduce systematic errors in time-dependent studies of  $B$  hadron decays. In chapter 4 a method was proposed to measure this behaviour for  $B \rightarrow h^+h'^-$  decays.

Apart from detector understanding, there are several potentially world-leading physics measurements that can be made with early data. One example would be the observation of

the decay  $B_d \rightarrow p\bar{p}$ , or indeed any two-body charmless baryonic  $B$  decay. The potential of LHCb to discover  $B_d \rightarrow p\bar{p}$  was explored in chapter 5.

As more data is collected, LHCb will soon be able to make an accurate measurement of  $\gamma$  using the decays  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^+K^-$ . The method of chapter 6 could be developed to allow the time-dependent  $\mathcal{CP}$  asymmetry distribution of  $B_s \rightarrow K^+K^-$ , as well that for  $B_d \rightarrow \pi^+\pi^-$ , to be studied. The extension of the  $\mathcal{CP}$  asymmetry fitter to include  $B_s \rightarrow K^+K^-$  would require the proper time resolution model to be correctly treated in the fit PDFs. The method laid out in chapter 4 can be used to measure the proper time resolution from data, to provide the necessary input to the fitter. This would then allow for a measurement of  $\gamma$  to be made using this method.

Since the decays  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^+K^-$  involve penguin amplitudes, the value of  $\gamma$  measured using them may be affected by New Physics. A comparison of  $\gamma$  as measured in these decays with that measured from decays that do not involve penguin amplitudes may lead to the discovery of New Physics effects, and help to open the door to a new era of discovery in High Energy Physics.

# Bibliography

- [1] F. Halzen and A. D. Martin. *Quarks & Leptons*. John Wiley and Sons, 1984.
- [2] A. D. Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *J. Exp. Theor. Phys. Lett.*, 5:24–27, 1967.
- [3] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.
- [4] S. Davidson. The Standard Model. In *Proceedings of the School for Experimental High Energy Physics Students, 3–15 September 2006, Abingdon, Oxfordshire, UK*, pages 113–206, 2007. RAL-TR-2007-004.
- [5] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys. B*, 22:579–588, 1961.
- [6] A. Salam and J. C. Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13:168–171, 1964.
- [7] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.
- [8] M. Kaku. *Quantum Field Theory*. Oxford University Press, 1993.
- [9] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.
- [10] Z. Maki, M. Nakagawa, and S. Sakata. Remarks on the Unified Model of Elementary Particles. *Prog. Theor. Phys.*, 28:870–880, 1962.
- [11] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:532–533, 1963.
- [12] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D*, 2:1285–1292, 1970.

- 
- [13] M. Kobayashi and T. Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.
- [14] L. Wolfenstein. Parameterisation of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.
- [15] C. Amsler et al. Review of Particle Physics. *Phys. Lett. B*, 667(1), 2008.
- [16] E. Noether. Invariante Variationsprobleme. *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys.*, pages 235–257, 1918.
- [17] C. S. Wu. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, 1957.
- [18] L. D. Landau. On the conservation laws in weak interactions. *Nucl. Phys.* 3, 3(127), 1957.
- [19] J. H. Cristenson et al. Evidence for the  $2\pi$  decay of the  $K_2^0$  meson. *Phys. Rev. Lett.*, 13:138–140, 1964.
- [20] A. Angelopoulos et al. First direct observation of time-reversal non-invariance in the neutral kaon system. *Phys. Lett. B*, 444(43), 1998.
- [21] G. D. Rochester and C. C. Butler. Evidence for the Existence of New Unstable Elementary Particles. *Nature*, 160:855–857, 1947.
- [22] J. E. Augustin et al. Discovery of a Narrow Resonance in  $e^+ - e^-$  Annihilation. *Phys. Rev. Lett.*, 33:1406–1408, 1974.
- [23] J. J. Aubert et al. Experimental Observation of a Heavy Particle  $J$ . *Phys. Rev. Lett.*, 33:1402–1406, 1974.
- [24] S. W. Herb et al. Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, 1977.
- [25] D. Andrews et al (CLEO Collab.). The CLEO Detector. *Nucl. Instr. Meth.*, 211:47–71, 1983.
- [26] D. Besson et al (CLEO Collab.). Observation of New Structure in the  $e^+ - e^-$  Annihilation Cross-Section Above  $B$  anti- $B$  Threshold. *Phys. Rev. Lett.*, 54(382), 1985.
- [27] Y. Kubota et al (CLEO Collab.). The CLEO II Detector. *Nucl. Instr. Meth.*, 320:66–113, 1992.
-

- 
- [28] B. Aubert et al (Babar Collab.). The Babar Detector. *Nucl. Instr. Meth. A*, 479:1–116, 2002.
- [29] A. Abashian et al (Belle Collab.). The Belle Detector. *Nucl. Instr. Meth. A*, 479:117–232, 2002.
- [30] H. Albrecht et al (ARGUS Collab.). Observation of  $B^0 - \bar{B}^0$  mixing. *Phys. Lett. B*, 192:245–252, 1987.
- [31] H. Albrecht et al (ARGUS Collab.). Argus: A universal detector at DORIS II. *Nucl. Instr. Meth. A*, 275:1–48, 1989.
- [32] V. M. Abazov et al (D0 Collab.). D0 Run IIB upgrade technical design report. FERMILAB-PUB-02-327-E, December 2002.
- [33] R. Blair et al (CDF Collab.). The CDF II Detector: Technical Design Report. FERMILAB-PUB-96-390-E, November 1996.
- [34] V. M. Abazov et al (D0 Collab.). Direct Limits on the  $B_s^0$  Oscillation Frequency. *Phys. Rev. Lett.*, 97(021802), 2006.
- [35] A. Abulencia et al (CDF Collab.). Observation of  $B_s^0 - \bar{B}_s^0$  Oscillations. *Phys. Rev. Lett.*, 97(242003), 2006.
- [36] I. I. Bigi and A. I. Sanda. *CP Violation*. Cambridge University Press, 2000.
- [37] V. Fanti et al (NA48 Collab.). A new measurement of direct CP violation in two pion decays of the neutral kaon. *Phys. Lett. B*, 465:335–348, 1999.
- [38] A. Alavi-Harati et al (KTeV Collab.). Observation of Direct CP Violation in  $K_{S,L} \rightarrow \pi\pi$  Decays. *Phys. Rev. Lett.*, 83:22–27, 1999.
- [39] B. Aubert et al (Babar Collab.). Direct CP Violating asymmetry in  $B_d \rightarrow K^+\pi^-$  decays. *Phys. Rev. Lett.*, 93(131801), 2004.
- [40] Y. Chao et al (Belle Collab.). Evidence for direct CP violation in  $B_d \rightarrow K^+\pi^-$  decays. *Phys. Rev. Lett.*, 93(191802), 2004.
- [41] B. Aubert et al (Babar Collab.). Search for T, CP and CPT violation in  $B^0 - \bar{B}^0$  Mixing with inclusive dilepton events. *Phys. Rev. Lett.*, 96(251802), 2006.
- [42] E. Nakano et al (Belle Collab.). Charge Asymmetry of Same-Sign Dileptons in  $B^0 - \bar{B}^0$  Mixing. *Phys. Rev. D*, 73(112002), 2006.
-

- 
- [43] O. Schneider.  $B^0 - \bar{B}^0$  Mixing Review. <http://pdg.lbl.gov/2009/reviews/rpp2009-rev-b-bar-mixing.pdf>.
- [44] LHCb Technical Proposal. CERN-LHCC-98-004.
- [45] D. Binosi and L. Theußl. JaxoDraw: A graphical user interface for drawing Feynman diagrams. *Comput. Phys. Commun.*, 161:76–86, 2004. Package available from <http://jaxodraw.sourceforge.net/>.
- [46] P. Ball et al. B Decays at the LHC. CERN-TH-2000-101, March 2000.
- [47] B. Aubert et al (Babar Collab.). Observation of CP violation in the  $B^0$  meson system. *Phys. Rev. Lett.*, 87(091801), 2001.
- [48] K. Abe et al (Belle Collab.). Observation of large CP violation in the neutral  $B$  meson system. *Phys. Rev. Lett.*, 87(091802), 2001.
- [49] T. Aaltonen et al (CDF Collab.). First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in  $B_s \rightarrow J/\psi\phi$  Decays. *Phys. Rev. Lett.*, 100(161802), 2008.
- [50] V. M. Abazov et al (D0 Collab.). Measurement of  $B_s^0$  Mixing Parameters from the Flavor-Tagged Decay  $B_s^0 \rightarrow J/\psi\phi$ . *Phys. Rev. Lett.*, 101(241801), 2008.
- [51] J. Albrecht et al. Measurement of mixing-induced CP violation in  $B_s \rightarrow J/\psi\phi$ . Chapter 4 of [55].
- [52] The Heavy Flavor Averaging Group. <http://www.slac.stanford.edu/xorg/hfag/>, July 2009.
- [53] CKMFitter: Home. <http://ckmfitter.in2p3.fr/>, July 2009.
- [54] M. Battle et al (CLEO Collab.). Observation of  $B$  decay to two charmless mesons. *Phys. Rev. Lett.*, 71:3922–3926, 1993.
- [55] The LHCb Collaboration. Roadmap for selected key measurements of LHCb. CERN-LHCb-2009-029, November 2009. [arXiv:hep-ex/09124179].
- [56] A. Bates et al. Charmless charged two-body B decays. Chapter 3 of [55].
- [57] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner. Decays of  $B$  mesons to two pseudoscalars in broken SU(3) symmetry. *Phys. Rev. D*, 52:6356–6373, 1995.
-

- 
- [58] C. Newby. *The production and evaluation of photodetectors for the LHCb Rich system and a study of the sensitivity to very rare hadronic B decays*. PhD thesis, University of Oxford, May 2005. CERN-THESIS-2007-018.
- [59] B. Aubert et al (Babar Collab.). Measurement of CP Asymmetries and Branching Fractions in  $B^0 \rightarrow \pi^+\pi^-$ ,  $B^0 \rightarrow K^+\pi^-$ ,  $B^0 \rightarrow \pi^0\pi^0$ ,  $B^0 \rightarrow K^0\pi^0$  and Isospin Analysis of  $B \rightarrow \pi\pi$  Decays. In *Proceedings of the 34th International Conference on High Energy Physics, ICHEP08, 29 July–5 August 2008, Philadelphia, Pennsylvania, USA, 2008*. [arXiv:hep-ex/08074226].
- [60] H. Ishino et al (Belle Collab.). Observation of Direct CP violation in  $B^0 \rightarrow \pi^+\pi^-$  Decays and Model-Independent Constraints on the Quark-Mixing Angle  $\phi_2$ . *Phys. Rev. Lett.*, 98(211801), 2007.
- [61] R. Fleischer. New strategies to extract  $\beta$  and  $\gamma$  from  $B_d \rightarrow \pi^+\pi^-$  and  $B_s \rightarrow K^+K^-$ . *Phys. Lett. B*, 459:306–320, 1999.
- [62] J. Charles. Taming the penguin contributions in the  $B_d^0(t) \rightarrow \pi^+\pi^-$  CP asymmetry: Observables and minimal theoretical input. *Phys. Rev. D*, 59(054007), 1999.
- [63] M. Gronau and D. London. Isospin analysis of CP asymmetries in  $B$  decays. *Phys. Rev. Lett.*, 65:3381–3384, 1990.
- [64] O. Deschamps et al. The CKM angle  $\alpha$  at LHCb. CERN-LHCb-2007-046, May 2007.
- [65] M. Adinolfi et al. The tree-level determination of  $\gamma$ . Chapter 2 of [55].
- [66] The Large Hadron Collider. <http://lhc.web.cern.ch/lhc/>, January 2009.
- [67] CERN: The European Organization for Nuclear Research. <http://www.cern.ch>, January 2009.
- [68] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3(S08003), 2008.
- [69] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3(S08004), 2008.
- [70] A. Augusto Alves Jr et al. The LHCb Detector at the LHC. *JINST*, 3(S08005), 2008.
- [71] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3(S08002), 2008.
-

- [72] LHCb Technical Design Report: Reoptimized Detector Design and Performance. CERN-LHCC-2003-030.
- [73] LHCb Technical Design Report: Vertex Locator. CERN-LHCC-2001-011.
- [74] LHCb Technical Design Report: RICH. CERN-LHCC-2000-037.
- [75] M. Calvi, O. Leroy, and M. Musy. Flavour Tagging Algorithms and Performances in LHCb. CERN-LHCb-2007-058, May 2007.
- [76] J. D. Jackson. *Classical Electrodynamics*. John Wiley and Sons, 2nd edition, 1975.
- [77] N. Brook et al. LHCb RICH 1 Engineering Design Review Report. CERN-LHCb-2004-121, October 2005.
- [78] M. Adinolfi et al. LHCb RICH 2 engineering design review report. CERN-LHCb-2002-009, March 2002.
- [79] M. Alemi et al. First operation of a hybrid photon detector prototype with electrostatic cross-focussing and integrated silicon pixel readout. *Nucl. Instr. Meth. A*, 449:48–59, 2000.
- [80] M. Patel, M. Losasso, and T. Gys. Magnetic Shielding Studies of the RICH Photon Detectors. CERN-LHCb-2005-055, January 2006.
- [81] A. MacGregor. *The Laser Mirror Alignment System for the LHCb RICH Detectors*. PhD thesis, University of Glasgow, April 2006.
- [82] K. Wyllie. Level-0 electronics for the LHCb rich. CERN-LHCb-2000-075, May 2001.
- [83] J. Podolansky and R. Armenteros. *Phil. Mag.*, 45(13), 1954.
- [84] LHCb Technical Design Report: Magnet. CERN-LHCC-2000-007.
- [85] A. Keune. Parameterization of the LHCb Magnetic Field Map. CERN-LHCb-2009-002, January 2009.
- [86] LHCb Technical Design Report: Inner Tracker. CERN-LHCC-2002-029.
- [87] LHCb Technical Design Report: Outer Tracker. CERN-LHCC-2001-024.
- [88] LHCb Technical Design Report: Calorimeters. CERN-LHCC-2000-036.
- [89] LHCb Technical Design Report: Muon System. CERN-LHCC-2001-010.

- [90] LHCb Technical Design Report: Trigger System. CERN-LHCC-2003-031.
- [91] LHCb Technical Design Report: Online System. CERN-LHCC-2001-040.
- [92] G. Haefeli et al. TELL1: Specification for a common readout board for LHCb. CERN-LHCb-2003-007, February 2003.
- [93] R. Jacobsson, B. Jost, and Z. Guzik. Readout supervisor design specifications. CERN-LHCb-2001-012, March 2001.
- [94] ETM Professional Control. <http://www.etm.at/>, May 2009.
- [95] L. Carson, A. Pickford, and P. Soler. Measurement of the Photoelectron Detection Efficiency of the HPD Anode. CERN-LHCb-2009-038, November 2009.
- [96] L. Somerville. *Performance of the LHCb RICH Photon Detectors and Tagging Systematics for CP Violation Studies*. PhD thesis, University of Oxford, May 2006. CERN-THESIS-2006-020.
- [97] LHCbPIX1 Page. <http://kwyllie.home.cern.ch/kwyllie/LHCbPIX1.htm>, May 2009.
- [98] The LHCb Collaboration. Contract with Delft Electronic Products - Photonis, Supply of Hybrid Photon Detectors for the LHCb RICH Counters. CERN Contract F549/PH/LHCb, January 2005.
- [99] R. Lambert. *LHCb Hybrid Photon Detectors and Sensitivity to Flavour Specific Asymmetry in Neutral B-Meson Mixing*. PhD thesis, University of Edinburgh, January 2009. CERN-THESIS-2009-001.
- [100] K. Wyllie. Pixel HPD Test System Hardware Description. [http://kwyllie.home.cern.ch/kwyllie/LHCbPIX1doc/Test\\_hardware\\_doc\\_1.0.pdf](http://kwyllie.home.cern.ch/kwyllie/LHCbPIX1doc/Test_hardware_doc_1.0.pdf).
- [101] CERN S-link Homepage. <http://hsi.web.cern.ch/HSI/s-link>.
- [102] National Instruments LabVIEW Page. <http://www.ni.com/labview/>.
- [103] T. Ypsilantis and J. Seguinot. A historical survey of ring imaging Cherenkov counters. *Nucl. Instr. Meth. A*, 343:1–29, 1994.
- [104] E. H. Darlington. Backscattering of 10–100 keV electrons from thick targets. *J.Phys. D*, 8:85–93, 1975.

- 
- [105] ORTEC TRUMP PCI. <http://www.ortec-online.com/trump.htm>.
- [106] M. Moritz. Die LHCb RICH Pixel HPDs. In *Deutsche Physikalische Gesellschaft Frühjahrstagung des Fachverbandes Teilchenphysik 2004, 29 March–1 April 2004, Mainz, Germany, 2004*.
- [107] T. Tabarelli de Fatis. Light Spectra Sum Rule. *Nucl. Instr. Meth. A*, 385:366–370, 1997.
- [108] GSL — GNU Scientific Library. <http://www.gnu.org/software/gsl/>.
- [109] M. Moritz et al. Performance Study of New Pixel Hybrid Photon Detector Prototypes for the LHCb RICH Counters. *IEEE Trans. Nucl. Sci.*, 51:1060–1066, 2004.
- [110] M. Adinolfi et al. Performance of the LHCb RICH photodetectors in a charged particle beam. *Nucl. Instr. Meth. A*, 574:39–49, 2007.
- [111] M. Adinolfi et al. Performance of the LHCb RICH photo-detectors and readout in a system test using charged particles from a 25 ns-structured beam. *Nucl. Instr. Meth. A*, 603:287–293, 2009.
- [112] P. Vankov and G. Raven. Proper-Time Resolution Modeling. CERN-LHCb-2007-055, July 2007.
- [113] L. Carson. Measurement of the Proper Time Resolution for  $B \rightarrow h^+h'^-$  Channels from Data. CERN-LHCb-2008-060, November 2008.
- [114] J. Borel. *The analog readout of the LHCb vertex detector and study of the measurement of the  $B_s$  oscillation frequency*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, June 2008. CERN-THESIS-2009-035.
- [115] G. Raven. Selection of  $B_s \rightarrow J/\psi\phi$  and  $B_u \rightarrow J/\psi K^+$ . CERN-LHCb-2003-118, December 2003.
- [116] K. Cranmer. Kernel Estimation in High Energy Physics. *Comput. Phys. Commun.*, 136:198–207, 2001.
- [117] The DaVinci Project. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/>.
- [118] A. Carbone et al. Charmless charged two-body B decays at LHCb. CERN-LHCb-2007-059, September 2007.
-

- 
- [119] M. Gersabeck. *Alignment of the LHCb Vertex Locator and Lifetime Measurements of Two-Body Hadronic Final States*. PhD thesis, University of Glasgow, September 2009.
- [120] V. Gligorov and J. Rademacker. Monte Carlo Independent Lifetime Fitting at LHCb in Lifetime Biased Channels. CERN-LHCb-2007-053, June 2007.
- [121] The RooFit Toolkit for Data Modeling. <http://roofit.sourceforge.net/>.
- [122] ROOT : A Data Analysis Framework. <http://root.cern.ch/drupal/>.
- [123] MINUIT : Function Minimization and Error Analysis. <http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html>.
- [124] L. Carson and E. Rodrigues. Two-Body Charmless Baryonic  $B$  Decays at LHCb. CERN-LHCb-2009-006, June 2009.
- [125] Y. T. Tsai et al (Belle Collab.). Search for  $B^0 \rightarrow p\bar{p}$ ,  $B^0 \rightarrow \Lambda\bar{\Lambda}$  and  $B^+ \rightarrow p\bar{\Lambda}$  at Belle. *Phys. Rev. D*, 75(111101), 2007.
- [126] B. Aubert et al (Babar Collab.). Search for the decay  $B^0 \rightarrow p\bar{p}$ . *Phys. Rev. D*, 69(091503), 2004.
- [127] T. E. Coan et al (CLEO Collab.). Search for exclusive rare baryonic decays of  $B$  mesons. *Phys. Rev. D*, 59(111101), 1999.
- [128] H. Y. Cheng and K. C. Yang. Charmless exclusive baryonic  $B$  decays. *Phys. Rev. D*, 66(014020), 2002.
- [129] V. Chernyak and I. Zhitnitsky. B-meson exclusive decays into baryons. *Nucl. Phys. B*, 345:137–172, 1990.
- [130] P. Ball and H. G. Dosch. Branching ratios of exclusive decays of bottom mesons into baryon-antibaryon pairs. *Z. Phys. C*, 51:445–453, 1991.
- [131] M. Jarfi et al. Relevance of baryon-antibaryon decays of  $B_d^0$ ,  $\bar{B}_d^0$  in tests of CP violation. *Phys. Lett. B*, 237:513–520, 1990.
- [132] M. Jarfi et al. Pole model of  $B$ -meson decays into baryon-antibaryon pairs. *Phys. Rev. D*, 43:1599–1632, 1991.
- [133] G. Feldman and R. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, 1998.
-

- 
- [134] The CDF Collaboration. Measurement of CP-violating asymmetries and of branching fractions of  $\Lambda_b^0 \rightarrow p\pi^-$  and  $\Lambda_b^0 \rightarrow pK^-$  decays. CDF Note 9092, January 2008.
- [135] Pythia Homepage. <http://home.thep.lu.se/~torbjorn/Pythia.html>.
- [136] B. Popovici and S. Stoica. Calibration of LHCb RICH Detectors with  $\Lambda \rightarrow p\pi^-$  Decay using Data. In *Physics at LHC-2008, 29 September–4 October 2008, Split, Croatia*, 2008. CERN-Poster-2008-047.
- [137] V. Vagnoni. Private communication.
- [138] F. Legger. Reconstruction of the decays  $\Lambda_b^0 \rightarrow \Lambda(1115)\gamma$  and  $\Lambda_b^0 \rightarrow \Lambda(1670)\gamma$  at LHCb. CERN-LHCb-2006-012, December 2006.
- [139] F. Legger. Polarized radiative  $\Lambda_b^0$  decays at LHCb. CERN-LHCb-2006-013, December 2006.
- [140] S. W. Lin. Hadronic B Decays at Belle and Babar. In *Proceedings of the 41st Rencontres de Moriond on QCD and High Energy Hadronic Interactions, 17–24 March 2007, La Thuile, Italy*, 2007. [arXiv:hep-ex/07053125].
- [141] R. Fleischer.  $B_{s,d} \rightarrow \pi\pi, \pi K, KK$ : Status and Prospects. *Eur. Phys. J. C*, 52(267–281), 2007.
- [142] M. Kreps. Branching Fractions and Direct CP Asymmetries of Charmless B Decay Modes at CDF. In *Proceedings of the 15th International Conference on Supersymmetry and the Unification of Fundamental Interactions, SUSY07, 26 July–1 August 2007, Karlsruhe, Germany*, 2007. [arXiv:hep-ex/07094555].
- [143] The MINT Package. Source code available from [http://www.phy.bris.ac.uk/people/rademacker\\_j/index.html](http://www.phy.bris.ac.uk/people/rademacker_j/index.html).
- [144] F. Azfar et al. A Monte Carlo Independent Method for Lifetime Fits in Data biased by the Hadronic Trigger. CDF-6756, November 2003.