Lifetime measurements of the $B_s^0 \to K^+K^-$ system and RICH studies at LHCb

Paul D Sail

University of Glasgow
Department of Physics and Astronomy

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

September 2013

Abstract

Lifetime analyses in heavy flavour systems can provide excellent opportunities to make precision measurements of Standard Model (SM) parameters that probe for new physics. This thesis presents lifetime measurements from the B^0_d and B^0_s mesons to the charmless two-body hadron family of decays, known as $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$, where h can be either a kaon or a pion. A particular interest is given to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ decay channel.

The $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ channel is interesting as it decays into a *CP* even final state, with a small amount of *CP*-violation predicted by the Standard Model. This can be quantified using the decay rate asymmetry parameter $\mathcal{A}_{\Delta\Gamma}$. The decay is also loop dominated, and thus sensitive to New Physics (NP) that could enter in these processes. These effects can be studied by constraining the values of the decay rate difference, $\Delta\Gamma_\mathrm{s}$, and $\mathcal{A}_{\Delta\Gamma}$ from a lifetime measurement of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$.

A method used to measure the effective lifetime of the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ channels, which removes the acceptance bias induced by the event selection, is described. This includes fits to the invariant mass and reconstructed lifetime spectrums, together with the methods employed to determine the per-event acceptance functions from data and treatment of non-parametric backgrounds.

The application of this method to two datasets is then presented. Both datasets are collected using a centre of mass energy $\sqrt{s} = 7\ \mathrm{TeV}$ with the first dataset from the 2010 run comprising a total integrated luminosity of $37\,\mathrm{pb}^{-1}$, and the second from the 2011 run comprising a total integrated luminosity of $1.0\,\mathrm{fb}^{-1}$. The fit methods are verified using various techniques, primarily using a simplified MC simulation and full LHCb Monte Carlo (MC). The fit to the mass spectrum of the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ from the 2011 dataset is used both for the lifetime measurement, and for the normalisation in a search for the $\mathrm{B}^0_{\mathrm{d/s}} \to \mathrm{p}\bar{\mathrm{p}}$ decay. As the details of the lifetime analysis methods are subtly different between datasets, the sources of uncertainty differ between the analyses. The lifetime measurement resulting from the 2010 dataset is found to be

$$\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-} = 1.440 \pm 0.096\ \mathrm{ps\ (stat)}\ \pm\ 0.008\ \mathrm{ps\ (syst)}. \tag{1}$$

The resulting $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes, measured from the 2011 dataset are found to be

$$\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-} = 1.407 \pm 0.016\ \mathrm{ps\ (stat)}\ \pm\ 0.007\ \mathrm{ps\ (syst)}, \tag{2}$$

$$\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-} = 1.524 \pm 0.011\ \mathrm{ps\ (stat)}\ \pm\ 0.004\ \mathrm{ps\ (syst)}, \tag{3}$$

$$\tau_{\mathrm{B}^0_s\to\pi^+\mathrm{K}^-} = 1.597 \pm 0.056 \text{ ps (stat)} \pm 0.012 \text{ ps (syst)}. \quad (4)$$

These measurements are consistent with previous measurements, with the value of $\tau_{\mathrm{B}^0_s\to\mathrm{K}^+\mathrm{K}^-}$ measured from the 2011 dataset having a greater precision than the current world average. From this, we are able to directly measure the quantity $\mathcal{A}_{\Delta\Gamma}$ for the first time, with the value evaluated found to be

$$\mathcal{A}_{\Delta\Gamma} = -0.87 \pm 0.17 \text{ (stat)} \pm 0.13 \text{ (syst)}. \quad (5)$$

Particle IDentification (PID) of the family of $\mathrm{B}\to\mathrm{h}^+\mathrm{h}^-$ decays is crucial for distinguishing the many different final states that are kinematically similar. These final states tend to overlap in mass when reconstructed, providing considerable difficulties when making measurements with them. The Ring Imaging CHerenkov (RICH) detectors provide the experiment with PID capability, which allow high precision measurements from these channels to be made. During data taking periods, it is important to monitor the performance of the many sub-systems that allow the RICH operations to run smoothly. Two studies are presented in this thesis that pertain specifically to RICH 2. The first studies the stability of the spherical and flat mirrors, using the Laser Alignment Monitoring System (LAMS). A direct correlation between observed movement and temperature is found, which is in the order $10-70\ \mu\mathrm{rad/K}$ and agrees with estimates based on the expansion coefficients of materials and the mirror supports. This is an order of magnitude lower than the angular resolution of RICH 2, so does not affect the resolution or performance of the detector. The second study suggests an alternative method to constrain the refractive index of the RICH gas. This method allows for continuous monitoring of the refractive index, with a greater frequency in time than is currently available. The method is purely data driven and is found to be in agreement with the currently employed RICH method to within 1%.

Acknowledgements

There are many people who have helped contribute to this thesis, both in terms of direct contribution to the work and also indirectly by helping me believe that I may actually complete it one day. Hopefully I will be able to mention them all here, but regrettably there may be some I inevitably forget.

To start with I would like to thank my supervisors for giving me the opportunity to join the LHCb group and experience the torture that is a PhD. The 'evolution' of my project through the years has been particularly enjoyable. Initially my supervisors were Paul Soler and Chris Parkes, but most latterly Lars Eklund who took over the unenviable task of being my second supervisor after Chris's 'defection' to Manchester..... we don't bare grudges here Chris, really, we don't:-)

So first of all, a big thanks to my first supervisor, Paul Soler. He was always available for a 'brief' chat, which would tend to turn into a couple of hours quite quickly. I have always come away from our discussions feeling I have learnt something new. Secondly, thanks to Chris Parkes for the brief period you were my second supervisor. I always found our discussions and comments helped me escape my general ignorance on subjects. Finally, thanks to Lars Eklund my second-second supervisor. Your answering of so many of my stupid questions about lifetime fitting, and having to answer the same questions multiple times, I'm surprised you are still sane!

To all the remaining Glasgow LHCb group members past and present who have played a part in this adventure I must say thanks. John, for the infamous 0-0 fifa marathon, you can only play the 1-10-0-0 formation once before it gets boring. Daniel, for your constant attempts to drag me out the office I salute you. Steven, with your attempts to wrest back control of the 'mastermind' chair, I'll leave this to you in my office will. Thanks to Michael and your endless knowledge of the fitter and help getting to grips with fitting with the monster that is G-Fact! Finally, a thanks to everyone else: Laurence, Marco (who I'll mention later), Eduardo, Sylvia, Tomasz, Alison, Christoph and Ian.

I must now move over to my collaborators elsewhere, mainly CERN but I won't be too specific. A special thanks to Vava, Marco and Antonis in particular. To Vava for all things swimming related and his ability to soak up my stupidity and calmly answer my many questions. Marco, who was on hand to answer many lifetime related questions. Finally Antonis, whose ability to repeat and explain many things to me during our many RICH discussions, helped me understand so much that I would previously not understood.

Now, I'll dare to leave the workplace and mention my friends who have attempted (but only occasionally been successful at) keeping me sane over the past few years. Thanks to the guys

in my flat in my first year, particularly Ahmed and Matthias. I think we'll all remember the infamous St Patricks Day of 2010..... and the following day! Thanks to the Engineers, who helped me escape the 'physics bubble' with some very heavy nights out and regular humiliations at poker. Further afield a special thanks to Johan and Kate who helped me realise their is life outside of study and one day I may be a 'real boy' again, and Lee the last dude in the village.

Moving closer to home, I would like to thank Helena for putting up with me these last few years. Without your support, and much moaning about both our theses, I don't think I could have got this far.

Last but not least, I owe a huge thank you and debt of gratitude to my family. I would especially like to give a special thanks to my mum, Christine, who has always supported me though everything I have done. To my sister Nicola and her husband Mark (a spurs fan but I'll let that go) who made me realise how godly I am at wii frisbee golf. To Michael who I've had many a late night discussion with about the mythical 'black box'. Finally my late father Douglas, who is sadly not around to see me complete this journey and whose memory I'd like to dedicate this thesis to.

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.

Paul. D. Sail

Preface

This thesis presents measurements of the beauty sector $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes, where h can be either a kaon or a pion, with the aim to make a high precision measurement of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime. This allows access to the decay rate difference parameter, $\Delta\Gamma_\mathrm{s}$, and the decay rate asymmetry parameter, $\mathcal{A}_{\Delta\Gamma}$, which measures the amount of CP-violation present in the decay. The measurements are made using $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ data, specifically the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ decay modes, collected during the data taking periods of 2010 and 2011 by the LHCb experiment at the LHC at $\sqrt{s} = 7\ \mathrm{TeV}$. The operational performance of the Ring Imaging Cherenkov Detector (RICH) of LHCb is particularly important for distinguishing between the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ final states. Hence, monitoring of the RICH performance with respect to the stability of the mirror alignment and refractive index are also studied.

Chapter 1 presents a discussion of the particle physics theory involved in this thesis, and is derived from various materials that are documented throughout. These primarily include theory and experimental particle physics papers, reference text books and presentations. The chapter begins by providing a brief review of the standard model of particle physics, including the particle content and the standard model Lagrangian. This is followed by a more detailed description of mixing and CP-violation, which is relevant to the physics analysis performed in this thesis. The final emphasis of this chapter is specifically on the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ sector and the physics that can be inferred from measuring $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes.

The LHCb detector is described in Chapter 2. This chapter begins by providing an overview of the LHC complex and the main experiments contained within the accelerator ring. This is then followed by a more detailed description of the LHCb detector, trigger system and reconstruction software. The details given originate primarily from technical documents, performance papers and analysis notes.

Chapter 3 presents performance studies using data from the RICH subdetector. There are two studies performed in this chapter. The first analysis is based on monitoring data from the laser based mirror alignment system, and involves determining the alignment constants for each mirror. This consisted of calculating individual positions of the system components in the correct coordinate system, vectors from component positions to their respective mirrors and various angles to define each plane. Calibrations of each mirror were performed to determine their relative movement coefficients. Possible sources of movement are then

proposed and studied, with the most probable source studied in further depth. The ray tracing code used for this analysis was written and supplied by Antonis Papanestis, with which I made further additions and updates required for the analysis. The second study proposes an alternative method to constrain the refractive index of the gas in RICH 2. This required monitoring data to be downloaded and then analysed in a consistent manner to extract the refractive index. The data was combined with measurements of the molecular gas mass, carried out by Christoph Frei at CERN with a gas chromatograph, in order to extract the refractive index values and to compare to those currently in the database. This work was conducted by me, unless where explicitly stated otherwise.

The method employed to separate the constituent signal and background channels within a given dataset, and extract the effective lifetime of a particle, is presented in Chapter 4. The method is dependent on three discrete components; a fit to the invariant mass, determination of per-event acceptance functions and then a fit to the reconstructed lifetime spectrum. The analysis has been developed over a number of years within the Glasgow LHCb group and collaborators and is well documented. The main areas of work I have contributed have been the development of the mass and lifetime fitter components for measurements of $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes. This chapter is meant as an overview of the method applied to data in this thesis.

The pre-analysis structure of the individual $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analyses using 2010 and 2011 data, is discussed in Chapter 5. An introduction to the experimental status of these measurements is briefly discussed, before moving into the mechanics of how the optimum selection criteria was determined for these data. The mass and lifetime models used to fit the invariant mass and reconstructed lifetime spectrums respectively, as well as the reasoning behind their selection, is then detailed. This is then followed by verification procedures of the constructed fitter using a variety of methods. The majority of these studies were conducted by myself, with the exceptions of the optimisation of the selection for the 2010 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ dataset and the verifications performed with $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ data.

Finally, the results of the fits to both the invariant mass and reconstructed lifetime spectrum are presented in Chapter 6. However, before the lifetime fit results are presented, the results of the fits for the signal yields and systematic studies for the lifetime fits are discussed. These are important to verify that all the dominant effects that could bias the lifetime are incorporated into the fitter. The fits to the reconstructed lifetime for the 2010 and 2011 data are then presented, with a discussion on the physics implications of the measurements. The work incorporated in this chapter has been performed by myself with the exception of

a small number of systematics that include: $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ studies, the production asymmetry, contributions to VELO systematics using $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ data and PID efficiency determinations.

Contents

List of Tables

List of Figures

Chapter 1

The Standard Model of Particle Physics and Flavour Theory

1.1 Theory Introduction

One of the most awe inspiring sights in the world is the night sky viewed from the uppermost ridges of Mauna Kea on the Big Island, Hawaii. The sky is covered with hundreds of thousands of stars, which can be seen as individuals, clusters and constellations. Against this backdrop, the centre of the Milky Way can be seen and is one of the more stunning sights in the natural world. It is from viewing this and similar sights that some of the great questions have been proposed, such as about the meaning of life and the origins of the universe. This thesis will only deal with one of these questions, the latter.

The fundamental constituents of matter have been hypothesised since the time of the ancient Greeks, with the word atom arising from the assumption of the philosopher Democritus. Democritus posed the question: if you break a piece of matter in half, and then break it in half again, how many breaks will you have to make until you can break it no further. These smallest pieces of matter, he surmised, would be called atoms. The theory of the modern atom was not proposed until centuries later in 1808 by J.Dalton [1]. Experimentally, progress was further advanced in the early 20th century with evidence of the composition of the nucleus by E. Rutherford [2], electrons by J.J. Thomson [3] and the discovery of anti-matter. Anti-matter was famously first predicted by the P. Dirac's equation (1928) [4], given by Equation 1.1, before it had been experimentally observed by Carl.D. Anderson (1932) [5].

$$i\hbar\frac{\partial\psi}{\partial t} = (c\alpha\cdot\mathbf{p} + \beta mc^2)\psi \tag{1.1}$$

This key theoretical prediction of a particle before its observation helped kick-start the ideas behind the development of the Standard Model (SM).

The Standard Model, from now on abbreviated as SM, is a theory to describe the fundamental components of matter and the way in which they interact. The theory was developed in the 1960s and finalised in the 1970s once the experimental confirmation of quarks (a fundamental constituent of matter, discussed further in Section 1.2.1) was achieved. The ability of the SM to predict the existence of the c-quark (1974) [6] [7], b-quark (1977) [8], t-quark (1995) [9, 10] and ν_τ-neutrino (2000) [11] has also given credence that it is the correct formalism for nature. The recently published observation [12, 13] of a Higgs-like particle [14] would complete the set of predicted particles of the SM (Figure 1.1).

The SM is currently believed to be a low energy 'effective theory', as there are many aspects of the universe that it cannot fully explain. Two of the key experimental issues that arise from cosmology are the failure of the SM to predict a candidate particle for dark matter, which is believed to exist within the universe, and the baryon asymmetry observed, which differs from the amount expected by the baryogenesis hypothesis [15].

Baryogenesis is the name given to the physical process that produced an imbalance between the number of baryons and anti-baryons at the start of the universe. The concept can be explained by first making the assumption that, in an ideal universe, the initial conditions are such that the density of quarks and their corresponding antiquarks are equal. During this period at the beginning there were two significant processes that occurred, these were proton-antiproton annihilation and proton-antiproton production. The first of these can occur at any temperature, however the second can only occur when the energy of the protons is large enough. This cut off is at $T \geq 2m_p$. Until this point, the production/annihilation processes keep the proton and antiprotons in an equilibrium. Once the temperature falls below $2m_p$, production would cease and the annihilation process would become dominant until all matter and antimatter had annihilated each other. However, we do not see this in current observations so at some point an imbalance must have occurred such that the universe became more populated with matter than antimatter.

The set of necessary conditions needed to produce a baryon (matter/antimatter) asymmetry was proposed by A. Sakharov (1967)[16]. These are:

- the requirement for baryon number violation;
- the violation of both C and CP symmetries;
- the previous 2 conditions must occur outside of thermal equilibrium (when $T < 2m_p$).

With these added into our model, we can begin to understand how the history of matter and antimatter may have proceeded in the universe. To determine the baryon asymmetry we can simply use the equation

$$\eta = \left[\frac{(n_q - n_{\bar{q}})}{n_\gamma}\right]_{BG} \tag{1.2}$$

where n_q and $n_{\bar{q}}$ are the number density of quarks and antiquarks respectively and n_γ is the number density of cosmic background radiation photons. The η is an asymmetry parameter. From this equation and deductions from nucleosynthesis constraints [17], the value of $\eta = (5.5 \pm 0.5) \times 10^{-10}$ is calculated. This differs from the SM predicted value of $\eta \approx 10^{-18}$ originating in quark mixing (discussed further in Section 1.3) by eight orders of magnitude [18]. The Large Hadron Collider beauty (LHCb) experiment at CERN Large Hadron Collider (LHC) is dedicated to the further exploration of the area of CP-violation (CPV), or matter-antimatter asymmetries, proposed in the SM. By studying heavy flavour physics at the high energies used at the LHC, the limits of CP-violation in the quark sector can be probed to determine the accuracy of the theory. This may also lead to new discoveries in the search for additional sources of CP violation, which may account for the baryon asymmetry we observe in our universe.

1.2 The Standard Model

The Standard Model of particle physics is the name given to the current consolidated knowledge of the electromagnetic, strong and weak interactions. The theory attempts to explain all the phenomena of particle physics in terms of the properties and interactions of all the fundamental particles, which come in distinct types (described further in Section 1.2.1). There is one other fundamental force which is not included, this is the gravitational force. This force is neglected due to the interaction between elementary particles being relatively small at currently accessible energies. The SM is arguably the most successful theory ever created and has been able to predict the existence of many particles and phenomena before their experimental observation. A basic overview of the model will now be discussed.

1.2.1 Standard Model Particle Content

Current knowledge of the content of the Standard Model shows the composition to contain a discrete subset of subatomic particles, which constitutes the full tapestry of particles currently observed. These elementary building blocks are the contents of all known particles to date, coupled with the force carrier particles, which determine the interactions between particles. These are organised and displayed in Figure 1.1.

The particles in Figure 1.1 are categorised into groups based on their spin and unique quantum numbers. In addition to the particles displayed, each particle has its own anti-particle

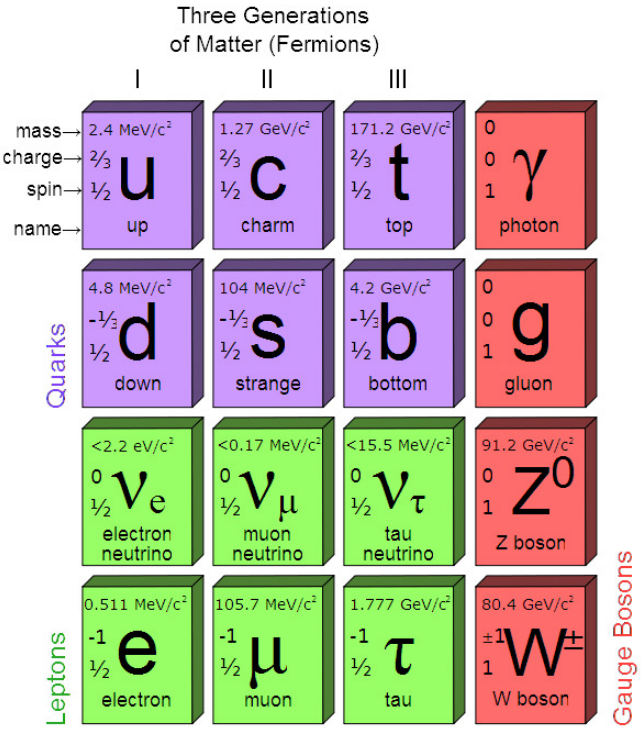


Figure 1.1: Known subatomic particles that make up the Standard Model [19].

which has the same mass but opposite quantum numbers. The 'spin' of a particle can be more simply described as its intrinsic angular momentum, this becomes important as it determines the way in which each particle interacts with each other. Fermi spin statistics [20] describes fermions and bosons in the following way:

- The wavefunction of any system of identical integer spin particles has the same value when the positions of any two particles are interchanged. Particles that exhibit the property of identical wavefunctions under this exchange are known as bosons
- The wavefunction of identical half-integer spin particles changes sign when the positions of these types of particles are interchanged. Particles that display the property of anti-symmetry under exchange are called fermions.

A simple classification of all 'free' particles, (quarks cannot be observed individually due to the strong force), is provided in Figure 1.2. These can be broadly described as:

- Leptons, which are fundamental particles and classified as fermions
- Gauge boson force carrier particles, which are integer spin and therefore classified as bosons.

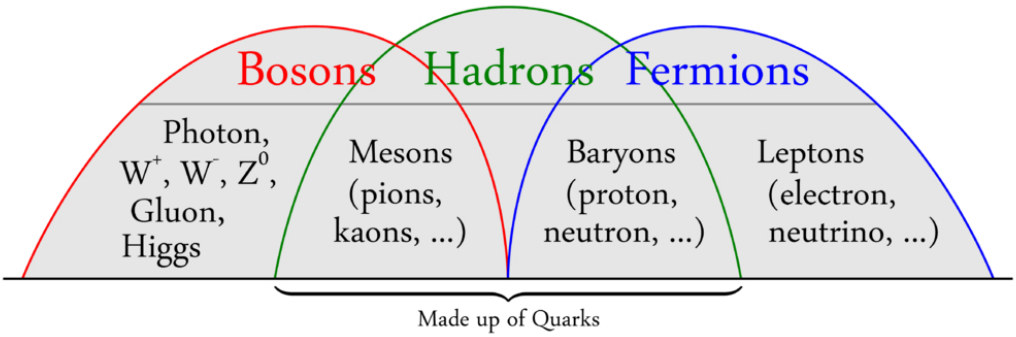


Figure 1.2: Particle classification. Note that mesons are both bosons and hadrons, and baryons are both hadrons and fermions [21].

- Hadrons that are composed of quarks and can have both integer and half-integer spin, so can be classified as a fermion or a boson.

A further discussion of Standard Model Fermions and Bosons is now provided in Sections 1.2.1.1 and 1.2.1.2 respectively.

1.2.1.1 Standard Model Fermions

The elementary half-integer spin fermions are shown in Figure 1.1, and can be further classified in terms of being either quarks or leptons.

There are six known quarks of different flavours: up (u), down (d), strange (s), charm (c), bottom (b) and top (t). The u, c and t quarks have charge $+\frac{2}{3}$e with the d, s and b quarks having charge $-\frac{1}{3}$e, where e is the magnitude of the charge of the electron. The quarks all have corresponding antiparticles which are called antiquarks. The antiquarks have the same mass, mean lifetime and spin of their respective matter counterparts, with the only measurable difference being the opposite charge.

The leptons also come in six flavours. These are the electron (e), muon (μ) and tau (τ) with corresponding neutrinos, the electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_τ). The neutrinos have negligible mass and carry zero charge, whereas the other leptons have large, measurable, masses and charge ± 1. In addition, these fermions are split into three generations with a quark pair and a lepton pair per generation.

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} e^- \\ \bar{\nu_e} \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} \mu^- \\ \bar{\nu_\mu} \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix} \begin{pmatrix} \tau^- \\ \bar{\nu_\tau} \end{pmatrix} \tag{1.3}$$

A fundamental difference between quarks and leptons is that quarks are affected by the strong force, whereas leptons are not. Their flavour quantum numbers are also different,

but although we have observed flavour changes, the overall baryon and lepton number have not been seen to be violated. However, as there is no symmetry guarding them it is believed they can be violated, which would be required for baryogenesis.

1.2.1.2 Standard Model Gauge Bosons

The SM force carrier particles have intrinsic integer spin properties and are thus called gauge bosons. These bosons represent manifestations of three of the four fundamental forces of nature, with the known forces being the strong, the weak, the electromagnetic and gravitational. The hypothetical exchange boson for the gravitational force, the graviton, has yet to be observed and would have spin $S = 2$. However, as this force is seen to have a negligible impact when considering high energy interactions, it will not be discussed in this thesis. The remaining three forces and their related bosons have spin $S = 1$ and will be discussed further, as they all contribute to high energy interactions.

The strong force is mediated by the gluon (g) particle. It is called the strong force as it is the strongest of all the fundamental forces, being $\approx 10^{38}$ stronger than the weakest force (gravity) and $\approx 10^{2}$ stronger than the next strongest (electromagnetic). This force interacts with the quarks due to them possessing a three-fold charge called 'colour'; leptons do not possess this and are distinguished from their fermion partners in this way. It is from this colour charge set of interactions that the field of Quantum Chromo-Dynamics (QCD) originated.

The weak force is mediated by more than one boson, these are the $\mathrm{W}^{\pm}$ and the Z^{0}. These bosons interact with all known fermions and are massive in size compared with the other force carrying particles, with $M_{\mathrm{W}} = 80.385 \pm 0.015\,\mathrm{GeV}/c^2$ [22] and $M_{\mathrm{Z}^0} = 91.1876 \pm 0.0021\,\mathrm{GeV}/c^2$ [22]. The relatively large mass of the $\mathrm{W}^{\pm}$ and Z^{0} also brings two other important aspects. The bosons have very short lifetimes with the $\tau_{\mathrm{W}^{\pm}} = 3.2 \times 10^{-25}\mathrm{s}$ and $\tau_{\mathrm{Z}^0} = 2.6 \times 10^{-25}\mathrm{s}$. More importantly they have a very short range for the interaction, with the coupling constant of the weak force being in the range 10^{-7}-10^{-6} relative to the strong force coupling constant.

The force carrier boson of the electromagnetic interaction is the photon (γ) and couples to electric charge even though the boson is itself neutral. Through experimental studies, the photon is known to be massless and thus in terms of being a force mediator has infinite range in free space. The photon in general does not decay to lighter particles due to its massless nature, unlike the mediators for the strong and weak forces. However, in the case of a photon with high enough energy in an external field, it could decay to a particle-antiparticle pair such as e^+e^-, $\mu^+\mu^-$ or $q\bar{q}$.

1.2.2 Gauge Theories

In order to understand the formalism of the Standard Model, we must first understand the basic premises of Gauge Theories as the SM is built upon these foundations. A gauge theory is a theory which obeys gauge symmetry. Gauge symmetry is a property of a field, where the equations describing that field remain the same after one applies an operation to all particles everywhere in space. The term gauge just means 'measure', such that a field with a gauge symmetry can be remeasured (or re-gauged) from different baselines without changing any properties of the original system.

As discussed in the previous subsection 1.2.1.2 there are three main forces of nature that play an inherent role in high energy physics interactions. Today, these three forces can be described in terms of unitary groups of different dimensions. This combination of gauge groups are known as $SU(3) \times SU(2) \times U(1)$. The Special Unitary ($SU$) groups can be represented by Lie group matrices with dimensions ($n^2 - 1$). The gauge groups $SU(3)$, $SU(2)$, $U(1)$ represent the strong, weak and electromagnetic force respectively. The $U(1)$ acts on the photon, the $SU(2)$ acts on the weak force bosons and the $SU(3)$ acts on the gluon fields of the strong force.

1.2.3 The Mathematical Description of the Standard Model

The description of the interaction between particles is best described by a Lagrangian. The classical form of the Lagrangian takes the form

$$\mathcal{L} = T - V, \tag{1.4}$$

with T being defined as the system kinetic energy and V being the potential energy.

The Lagrangian formalism was introduced by Joseph Louis Lagrange in 1788 in the reformulation of classical mechanics, called Lagrangian mechanics, and is one of the simplest equations used in Physics. If the Lagrangian equation for a system is known, then the equations of motion for that system can be obtained by directly substituting the Lagrangian into the Euler-Lagrange equation, with the one-dimensional example given by

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = 0. \tag{1.5}$$

The single generation Standard Model Lagrangian can be written as

$$\mathcal{L} = \mathcal{L}_{kinetic} + \mathcal{L}_{Yukawa} + \mathcal{L}_{Higgs}, \tag{1.6}$$

with this relationship determining all the interactions between particles and fields within the current framework of high energy physics.
The kinematics of the bosons and fermions are described by the first term, $\mathcal{L}_{kinetic}$, this can be split further into two terms to independently describe the bosons and fermions respectively. The fermions and bosons are treated differently by the $\mathcal{L}$, with the $\mathcal{L}$ of fermions given by

$$\mathcal{L} = \overline{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{1.7}$$

and the boson $\mathcal{L}$ given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{1.8}$$

The fermion and boson $\mathcal{L}$ are both described in terms of vector fields, with the ψ in Equation 1.7 representing the fermion vector field and A is the representation of the boson field in Equation 1.8 where $F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$.
The second term, $\mathcal{L}_{Yukawa}$, is what describes the interactions of the quarks and is the part of the Lagrangian that is of most interest in this thesis as it is where CP-violation originates in the SM.
The final term is $\mathcal{L}_{Higgs}$, which deals with the coupling of the Higgs field to the fermions and bosons to give them their mass.

1.3 Flavour Physics and CP Violation

Flavour Physics can be described as the interactions between the different types, or 'flavours', of quarks and leptons.

> *"The term flavour was first used in the particle physics in the context of the quark model of hadrons. It was coined in 1971 by Murray Gell-Mann and his student at the time Harald Fritzsch, at a Baskin-Robbins ice-cream store in Pasadena. Just as ice-cream has both colour and flavour so do quarks."*
>
> *RMP 81 (2009) 1887*

The work described in this chapter and later chapters will focus primarily on flavour phenomenology that occurs in the quark sector, and in particular the mixing of neutral B mesons. The idea of symmetries is discussed, followed by the introduction of extra generations into the SM before finally moving onto the Cabibbo-Kobayashi-Maskawa matrix ($\mathbf{V_{CKM}}$) which

governs the physics of quark mixing. The nature of neutral meson mixing and the key CP-violation processes that occur through this are discussed, as they have an important contribution in the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime analysis conducted through Chapters 5 and 6.

1.3.1 Symmetries in Nature

Symmetries and conservation laws play an important part in nature and are especially crucial in the building of theories in physics. There are many different examples of continuous symmetries that exhibit conservation laws such as: translational symmetry that gives momentum conservation, rotational symmetry that gives conservation of angular momentum and time translational symmetry that gives energy conservation. These are universal laws of nature, which are valid for all interactions.

There are other, discrete, symmetries that also arise in nature and are only applicable in certain scenarios and interactions. These can be defined as invariant if the Hamiltonian remains unchanged by the transformation. The fundamental discrete symmetries that are of particular interest are: Parity (P), Charge Conjugation (C) and Time reversal (T).

The parity symmetry, in simple terms, reflects a coordinate system through its origin which means converting a set of right-handed coordinates into a set of left-hand coordinates. In terms of vectors, these would change sign if parity was applied such that: $\vec{x} \to -\vec{x}$, although axial vectors would remain unchanged through the transformation: $\vec{L} = \vec{x} \times \vec{p} \to \vec{L}$. If the parity operator, P, operates on the P eigenstate $|\Psi(r,t)\rangle$, then $P|\Psi(r,t)\rangle = q|\Psi(-r,t)\rangle$ and $P^2|\Psi(r,t)\rangle = q^2|\Psi(r,t)\rangle$, where the eigenvalue $q = \pm 1$. Interestingly, parity is not an exact symmetry compared with some of those previously mentioned, i.e. rotational invariance, as it does not hold under all closed systems. This symmetry is noticeably violated through the weak interaction.

The second discrete symmetry of interest is the charge conjugation operation. The charge operator, C, transforms a particle Ψ into its anti-particle $\bar{\Psi}$ such that when it acts on the C eigenstate $|\Psi(r,t)\rangle$, we see $C|\Psi(r,t)\rangle = q|\bar{\Psi}(r,t)\rangle$ and $C^2|\Psi(r,t)\rangle = q^2|\Psi(r,t)\rangle$, where the eigenvalue $q = \pm 1$. The particles that are transformed all keep the same state, so their momentum, positions, etc are all unchanged and only their charge alters. As only the charge, and therefore the magnetic moment of the particles, changes, the electromagnetic interaction remains invariant under this operation. Charge symmetry also remains invariant under the strong interaction, but is again violated in the weak interaction.

The final symmetry is time reversal or T invariance. This is defined as the invariance under the transformation $t \to -t$ and is described by applying the time operator, T, on the T eigenstate $|\Psi(r,t)\rangle$. The T operator changes the direction of motion of a particle for example, the probability of finding the particle at position *r* at time *t* becomes the probability of finding

the particle at position *r* at the time *-t* in the transformed system. Similar to the P and C invariances, the T operation is also invariant under the electromagnetic and strong interactions but is not conserved in the weak interactions.
In the framework of Lagrangian theory, any quantum field theory that is Lorentz invariant is invariant under the combination of the C, P and T transformations [23]. However, CPT is currently the only combination of these three transformations that produces an exact symmetry in the laws of nature.

1.3.2 Additional Generations

The Standard Model in its simplest form only contains one generation of particles, however this differs from what is observed experimentally. It is ideal to start with the addition of a second generation and consider the effects that emerge when this is added. The main phenomenon that occurs is mixing between flavours and generations, this occurs between both quarks and leptons. There are two mixing matrices which govern the mixing phenomenon for these sectors, the CKM matrix [24] for quarks and the PMNS matrix [25] for leptons. Since B-physics is the main topic of this thesis, the PMNS matrix and lepton mixing will not be discussed further.
The first work to link strangeness violation in weak decays to the currently known quarks (u, d and s) was performed by N.Cabibbo in 1963 [26]. His solution was to determine the amount of strangeness violation in weak interactions by postulating that the weak eigenstate *d'* is a rotation of the flavour eigenstates *d* and *s*, such that

$$\begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} u \\ d\cos\theta_c + s\sin\theta_c \end{pmatrix} \tag{1.9}$$

where the Cabibbo angle was experimentally measured to be $\theta_c = 13.04^\circ$ [22]. The Cabibbo angle effectively quantifies the different coupling strengths between the different generations of quarks. The Glashow-Iliopoulis-Maiani, or GIM, mechanism [27] is an extension of this model which included the addition of the then unobserved charm (c) quark as an up-type partner of the *s* quark. A new matrix mixing the *d* and *s* flavour states into the *d'* and *s'* weak eigenstates was postulated to explain some of the inconsistencies within the original Cabibbo model, such as the smallness of the $\mathrm{K}^0 \to \mu^+\mu^-$ decay:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{1.10}$$

The charm quark was later discovered in 1974 as the bound $c\bar{c}$ state, J/ψ [6] [7].

1.3.3 The CKM Matrix

The Cabibbo-Kobayashi-Maskawa matrix [24], $\mathbf{V_{CKM}}$, is an extension of the GIM mechanism which introduces a third generation of quarks, the bottom (b) and top (t) quark. As discussed previously, the Cabibbo mixing angle is related to the relative probability that down and strange quarks decay into up quarks ($|V_{ud}|$ and $|V_{us}|$ respectively).
The matrix $\mathbf{V_{CKM}}$ provides a link between the weak eigenstates ($\mathrm{d}', \mathrm{s}', \mathrm{b}'$) with the flavour eigenstates ($\mathrm{d}, \mathrm{s}, \mathrm{b}$)

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = \mathbf{V_{CKM}} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}, \tag{1.11}$$

where the matrix can consequently be written as

$$\mathbf{V_{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.12}$$

The naming convention of each element represents the transition between quark flavours, with $|V_{ij}|^2$ giving the probability of a transition $i \to j$. Each element of the matrix is allowed to be a complex number, which then attributes eighteen free parameters. Due to the requirement of $\mathbf{V_{CKM}}$ to be a unitary matrix, the number of degrees of freedom is reduced to nine. The application of global phase rotations to the quark fields removes a further five of the phases, reducing the remaining number of free parameters to four. Three of the remaining four parameters consist of mixing angles, which are analogous to the Cabbibo angle in the three and four quark model. The fourth and final free parameter is a complex phase, which if non-zero will cause differences between the interactions of quarks and antiquarks giving rise to CP-violation. To view these parameters more clearly it is possible to choose an explicit parameterisation. The standard parameterisation [28] used by the Particle Data Group (PDG) [22] is given by

$$\mathbf{V_{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.13}$$

where the three mixing angles are θ_{12}, the previously discussed Cabibbo angle θ_C, θ_{23} and θ_{13}. The parameters $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$, while the remaining parameter is the Kobayashi-Maskawa phase, δ, which is currently the dominant source of CPV in the SM.

The $\mathbf{V_{CKM}}$ is unitary. This translates as the complex square matrix, $\mathbf{V_{CKM}}$, times by its transposed conjugate matrix, $\mathbf{V_{CKM}^{T*}}$, should be equal to the Identity matrix. The constraints applied to the on-diagonal components are then

$$\sum_i |V_{ij}|^2 = \sum_j |V_{ij}|^2 = 1, \tag{1.14}$$

with $i \in \{\mathrm{u}, c, \mathrm{t}\}$ and $j \in \{\mathrm{d}, \mathrm{s}, \mathrm{b}\}$. This condition is know as *weak universality* which implies that the sum of all couplings of any of the up-type quarks to all the down-type quarks is the same for all generations. The off diagonal elements in the unitarity requirement are given by

$$\sum_k V_{ki} V_{kj}^* = 0, \tag{1.15}$$

where $k \in \{\mathrm{u}, c, \mathrm{t}\}$ is fixed and $i, j \in \{\mathrm{d}, \mathrm{s}, \mathrm{b}\}$ with $i \neq j$. This gives rise to six different relations, each containing three components which relate to triangles in the complex plane. Three of these relations are

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0, \tag{1.16}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \tag{1.17}$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.18}$$

Each one of the six triangles is unique, however they are special in that all have the same area. This area is related to the CP-violation phase δ, and parameterises the amount of CP-violation in the SM. In the case where the conditions dictate that the SM has parameters that remove CP-violation from the model, the triangle areas would reduce to zero. Each of the three subcomponents of the six relations form the side of a triangle. Of the six triangles, the one most commonly used is the one relating to Equation 1.18, which is know as *the unitarity triangle* (Figure 1.3).

The angles in this triangle represent the level of CP-violation in the different transitions between quark levels, with the angles given by the relationships

$$\alpha = -\arg\left(\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \tag{1.19}$$

$$\beta = -\arg\left(\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right), \tag{1.20}$$

$$\gamma = -\arg\left(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{1.21}$$

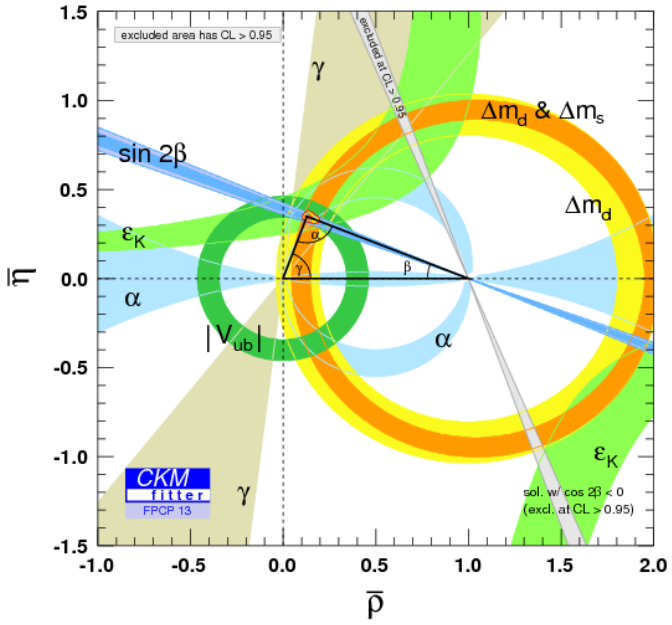


Figure 1.3: The Unitarity Triangle constructed from Equation 1.18 with parameter constraints applied [29].

The angles of this triangle should add up to $180°$ like any other triangle. Any deviation from this would suggest alternative forms of CP-violation or non-unitarity, as yet unaccounted for in the SM. Current experimental knowledge of the unitarity triangle is shown in Figure 1.3.

1.3.4 Neutral Meson Mixing

The phenomenon of neutral meson mixing has been observed to occur in the following meson systems: K^0, D^0, B^0_d and B^0_s. These neutral mesons are unique in being able to evolve into their anti-particle states via the weak interaction. Other neutral mesons, such as $q\bar{q}$ pairs, cannot do this as they are already their own anti-particle. These processes are particularly interesting as they allow us to experimentally measure different types of CP-violation, which is discussed further in Section 1.3.5. The leading order processes that describe how the mixing phenomenon occurs are displayed in Figure 1.4.

Studies of B^0 systems that include mixing allow measurements of these neutral mesons that are superpositions of both B^0 and $\bar{\mathrm{B}}^0$ states. The state of a neutral B^0 meson is thus observed experimentally to be a linear combination of its particle and anti-particle, given by

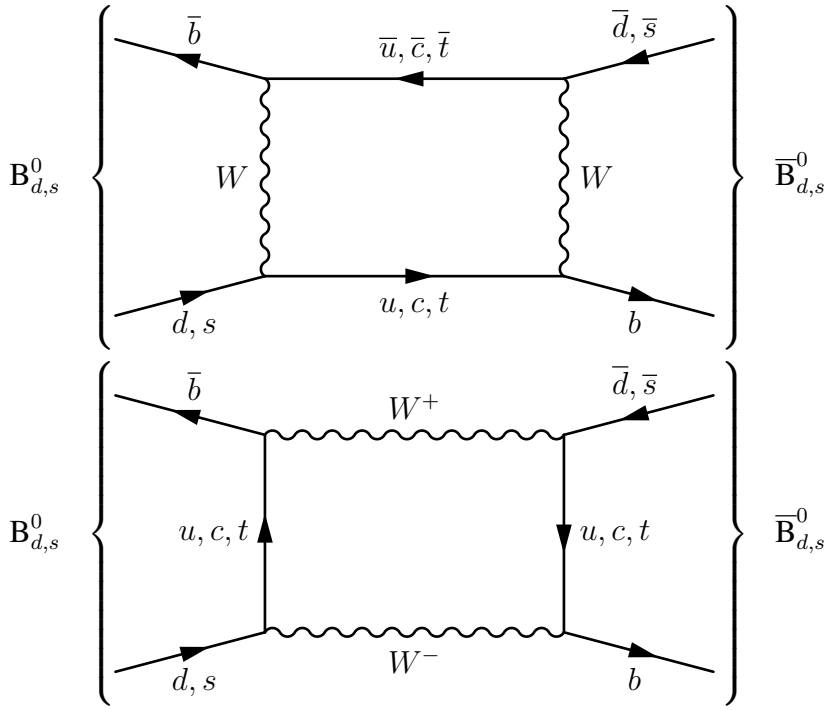


Figure 1.4: Leading order box diagrams involved in B^0-mixing.

$$\begin{aligned} |B_L\rangle &= p|\mathrm{B}^0\rangle + q|\bar{\mathrm{B}}^0\rangle \\ |B_H\rangle &= p|\mathrm{B}^0\rangle - q|\bar{\mathrm{B}}^0\rangle, \end{aligned} \tag{1.22}$$

where the labelling L and H represent the 'light' and 'heavy' states and the complex parameters p and q satisfy the conditions

$$|p|^2 + |q|^2 = 1. \tag{1.23}$$

The instantaneous decay amplitudes of a neutral meson system can be described by the terms A_f and $A_{\bar{f}}$. These terms denote the decay of a particle P into a multi-body final state f or its charge conjugate final state $\bar{f}$.

$$A_f = \langle f|\mathcal{H}|\mathrm{P}\rangle, \;\; A_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\mathrm{P}\rangle \tag{1.24}$$

Further terms, $\bar{A}_f$ and $\bar{A}_{\bar{f}}$, are used to describe the decays of the CP conjugate $\overline{\mathrm{P}}$ into the same final states $\overline{f}$ given by

$$\bar{A}_f = \langle f|\mathcal{H}|\bar{\mathrm{P}}\rangle, \;\; \bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{\mathrm{P}}\rangle, \tag{1.25}$$

where $\mathcal{H}$ is the Hamiltonian governing the weak interactions of the system.
The overview of neutral meson mixing provided in this section has assumed no CP-violation, the inclusion of this will now be discussed.

1.3.5 CP Violation

As discussed previously in Section 1.3.3, CP-violation has origins in the quark sector weak interactions, where components of the quark mixing are elements in the CKM matrix (Equations 1.12 and 1.13). This is currently the only known source of CP-violation in the quark sector. The phenomenon of neutral meson mixing, introduced in Section 1.3.4, allows the observation and potential measurement of three different types of CP-violation. These can be categorised into the following:

- Direct CP-violation. This occurs in both charged and neutral decays when the amplitude between a decay and its CP-conjugate process differ in magnitude.
- CP-violation in mixing. This occurs when the mixing amplitude between one state and its anti-state have different amplitudes.
- CP-violation through interference. This can occur both with or without mixing, for an initial neutral meson state decaying to a final state.

A neutral meson mixing decay system that incorporates two of the three specific CP-violation mechanisms, direct and mixing, is depicted in Figure 1.5. These sources of CP-violation, with the addition of the interference source, will now be elaborated on further.
The condition for direct CP-violation in decay is related to the instantaneous decay amplitude of a particle, P, and its anti-particle, $\bar{\mathrm{P}}$, Equations 1.24 and 1.25, to a final state f. If the ratio of the amplitudes is found to meet the condition

$$\left|\frac{A_f}{\bar{A}_{\bar{f}}}\right| \neq 1, \tag{1.26}$$

the consequence is that the $\mathrm{P} \to f$ decay and its charge conjugate have different rates. This direct CP-violation can be measured using the relationship

$$\mathcal{A}^{\mathrm{CP}} = \frac{N(\mathrm{P} \to f) - N(\bar{\mathrm{P}} \to \bar{f})}{N(\mathrm{P} \to f) + N(\bar{\mathrm{P}} \to \bar{f})}. \tag{1.27}$$

The second mechanism, mixing induced CP-violation, results from the mass eigenstates being different from the CP eigenstates. This originates from the relative phase between the M_{12} and Γ_{12} parameters, which will be discussed in Section 1.4. For now we will use the

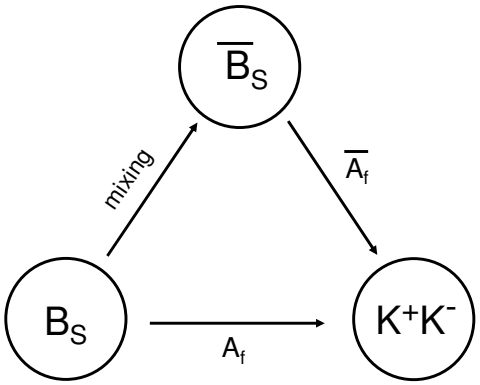


Figure 1.5: Neutral meson mixing diagram showing the different CP-violation effects in the system. This details the special case where the final state f is the same as the conjugate final state $\bar{f}$.

basics of the neutral B system introduced in Section 1.3.4 to generalise the asymmetry. The Schrödinger equation must first be solved for the eigenstates of the Hamiltonian, these take the forms given in Equation 1.22, with the complex parameters satisfying the condition given in Equation 1.23. If the mass eigenstates of a neutral meson are not CP eigenstates, then this causes the rates of $\mathrm{P} \to \bar{\mathrm{P}}$ and $\bar{\mathrm{P}} \to \mathrm{P}$ to be unequal such that

$$|p|^2 + |q|^2 \neq 1. \tag{1.28}$$

This manifests itself as an asymmetry between the mixing of $\mathrm{P} \to \bar{\mathrm{P}}$.
The final source of CP-violation that can be measured via neutral meson mixing, which is not clearly depicted in Figure 1.4, is the interference between the direct and mixing asymmetries. This form of CP-violation is only permissible if the final state f is accessible to both P and $\bar{\mathrm{P}}$, and can occur even if there is no direct or mixing CP-violation in the neutral meson system. This is possible as the processes $\mathrm{P} \to f$ and $\mathrm{P} \to \bar{\mathrm{P}} \to f$ share the same final states, and so can quantum mechanically interfere. This can be described by the complex quantity

$$\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{1.29}$$

CP-violation occurs from this interference when the condition $\mathrm{Im}\ \lambda_\mathrm{f} \neq 0$ is met.

1.4 Two-Body Charmless B Decays and B Meson Lifetimes

Analysis of the class of charmless two body final state decays commonly referred to as $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$, where h can be either a kaon or a pion, can improve current understanding of CP-violation and provide areas where New Physics (NP) could potentially reveal itself [30]. These areas of study include making measurements of the unitarity triangle angles, in particular the angle γ given by Equation 1.21 and displayed in Figure 1.3, measuring CP asymmetries and performing high precision lifetime measurements to gain access to SM parameters. The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays can proceed through many different decay routes, with specific examples provided in Figure 1.6. As the size of the parameter V_{ub} in the CKM matrix is small relative to the other couplings, the strength of the tree processes shown in Figure 1.6 are not dominant. With no particular process having overall dominance, Beyond the Standard Model (BSM) processes that are expected to appear in loops may make sizeable contributions to the decay amplitudes of $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays.

The $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ decay is of particular interest as a lifetime measurement can help constrain the B_s^0 mixing phase Φ_s, decay rate difference $\Delta\Gamma_s$ and decay rate asymmetry $\mathcal{A}_{\Delta\Gamma}$. This is possible due to the difference in lifetime between the two mass eigenstates, which provides sensitivity to CP-violation from the measured 'effective' lifetime. The details of this will be introduced using the following arguments and derivations. These deal in particular with the neutral B_s^0 system, although it is also analogous to the B_d^0 system.

The evolving $B_s \; - \; \bar{B}_s$ system is governed by the Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} |B_s(t)\rangle \\ |\bar{B}_s(t)\rangle \end{pmatrix} = \left(M_s - \frac{i}{2}\Gamma_s\right)\begin{pmatrix} |B_s(t)\rangle \\ |\bar{B}_s(t)\rangle \end{pmatrix}, \tag{1.30}$$

with the mass matrix M_s and the decay matrix Γ_s given by

$$M_s = \begin{pmatrix} M_s & M_{s\,12} \\ M^*_{s\,12} & M_s \end{pmatrix} \tag{1.31}$$

and

$$\Gamma_s = \begin{pmatrix} \Gamma_s & \Gamma_{s\,12} \\ \Gamma^*_{s\,12} & \Gamma_s \end{pmatrix}. \tag{1.32}$$

The mass eigenstates $|B_H\rangle$ and $|B_L\rangle$ with masses M_H, M_L and decay rates Γ_H, Γ_L, are obtained by diagonalising $M_s - \Gamma_s/2$. For the scope of this analysis, the quantity

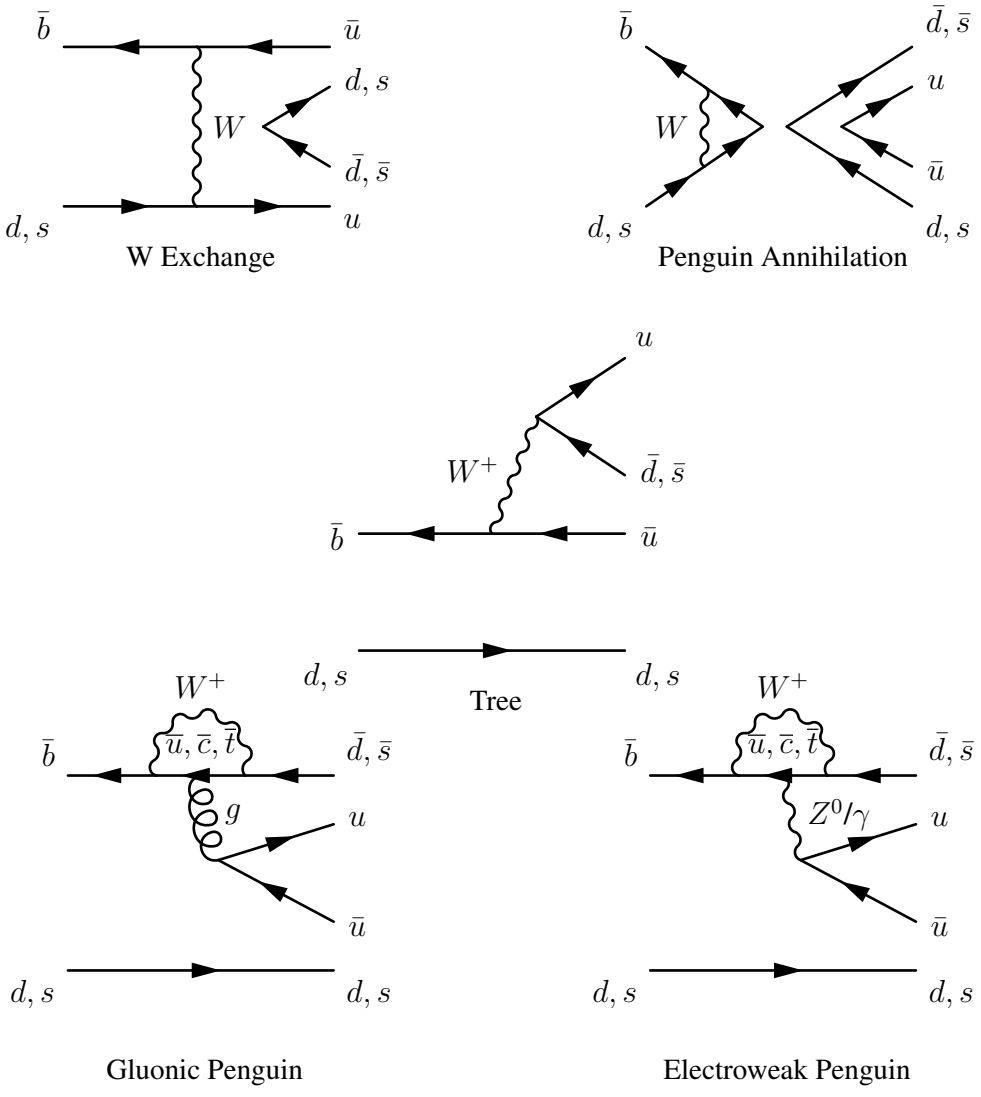


Figure 1.6: Feynman diagrams for possible $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decay route processes.

$$\Delta\Gamma_s = \Gamma_L - \Gamma_H = 2|\Gamma_{12}|\cos\Phi_{M/\Gamma} \tag{1.33}$$

is of particular importance. Within the Standard Model, $\Phi_{M/\Gamma}^{SM} = (4.2 \pm 1.4) \times 10^{-3}$ [31], i.e. $\cos\Phi_{M/\Gamma}^{SM} \approx 1$ to good accuracy. It is important to note that Γ_{12} is dominated by CKM-favoured $b \to c\bar{c}s$ tree level decays and is hence likely to mask new physical phenomena. Any NP affecting $\Delta\Gamma_s$ is expected to enter via the phase $\Phi_{M/\Gamma}$. This can be expressed as

$$\Delta\Gamma_s^{meas} = 2|\Gamma_{12}^{SM}|\cos(\Phi_{M/\Gamma}^{SM} + \Phi^{\Delta}) \tag{1.34}$$

and new CP violating contributions then change $\Delta\Gamma_s^{meas}$ with respect to the Standard Model prediction.

The mass eigenstates $|B_L\rangle$ (light) and $|B_H\rangle$ (heavy) at time $t = 0$ are defined as the linear combination of the flavour eigenstates $|B_s\rangle$ and $|\bar{B}_s\rangle$, as shown in Equation 1.22.

Following the discussion in [32, 33], we introduce the following quantities

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} = \eta_f e^{-i\Phi_M}, \tag{1.35}$$

where $\Phi_M = \arg(M_{s\,12})$, $A_f = \langle f|\mathcal{H}|\mathrm{B}_s^0\rangle$, $\bar{A}_f = \langle f|\mathcal{H}|\bar{\mathrm{B}}_s^0\rangle$ and η_f is the CP parity of $CP|f\rangle = \eta_f|f\rangle$ that is, $\eta_f = \pm 1$. In addition

$$\mathcal{A}_{CP}^{dir} = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \quad \mathcal{A}_{CP}^{mix} = \frac{2\mathrm{Im}(\lambda_\mathrm{f})}{1+|\lambda_f|^2} \quad \text{and} \quad \mathcal{A}_{\Delta\Gamma} = \frac{2\mathrm{Re}(\lambda_\mathrm{f})}{1+|\lambda_\mathrm{f}|^2}, \tag{1.36}$$

which are related via $(\mathcal{A}_{CP}^{dir})^2 + (\mathcal{A}_{CP}^{mix})^2 + (\mathcal{A}_{\Delta\Gamma})^2 = 1$. Additionally the parameter $\mathcal{A}_{\Delta\Gamma}$, also called the decay rate asymmetry and measured in Chapter 6, can be related to more easily interpretable physical quantities by the relationship

$$\mathcal{A}_{\Delta\Gamma} = \frac{R_H - R_L}{R_H + R_L}, \tag{1.37}$$

where R_H is the decay rate of the heavy mass eigenstate and R_L the decay rate of the light mass eigenstate. The general equation for the decay rate of the process $B_s \to f$ for untagged B meson decays (if the final state is accessible by both the B and $\bar{B}$ meson) can be written as

$$\begin{aligned}
\Gamma[f,t] &= \Gamma(B_s(t) \to f) + \Gamma(\bar{B}_s(t) \to f) && (1.38)\\
&= N_f\left[e^{-\Gamma_L t}\left|\langle f|\mathcal{H}|B_L\rangle\right|^2 + e^{-\Gamma_H t}\left|\langle f|\mathcal{H}|B_H\rangle\right|^2\right] && (1.39)\\
&\sim N_f\left|A_f\right|^2\left[1+|\lambda_f|^2\right]e^{-\Gamma t}\left\{\cosh\frac{\Delta\Gamma_s t}{2} + \sinh\frac{\Delta\Gamma_s t}{2}\mathcal{A}_{\Delta\Gamma}\right\}. && (1.40)
\end{aligned}$$

Taking the overall normalisation such that $\sum_f BR[f] = 1$, the overall normalisation N_f can be related to the branching ratio

$$BR[f] = \frac{1}{2}\int_0^\infty dt\Gamma[f,t] \sim \frac{N_f}{2}\left|A_f\right|^2\left[1+\left|\lambda_f\right|^2\right]\frac{\Gamma + \mathcal{A}_{\Delta\Gamma}\Delta\Gamma_s/2}{\Gamma^2 - (\Delta\Gamma_s/2)^2} \tag{1.41}$$

and hence

$$\Gamma[f,t] = 2BR[f]\frac{\Gamma^2-(\Delta\Gamma_s)^2}{\Gamma + \mathcal{A}_{\Delta\Gamma}\Delta\Gamma_s/2}e^{-\Gamma t}\left\{\cosh\frac{\Delta\Gamma_s t}{2} + \sinh\frac{\Delta\Gamma_s t}{2}\mathcal{A}_{\Delta\Gamma}\right\}. \tag{1.42}$$

This can be expressed as

$$\Gamma[f,t] = N_{f'}\frac{|A_f|^2}{2}\left(1+\left|\lambda_f\right|^2\right)\left[\left(1-\mathcal{A}_{\Delta\Gamma}\right)e^{-\Gamma_L t} + \left(1+\mathcal{A}_{\Delta\Gamma}\right)e^{-\Gamma_H t}\right]. \tag{1.43}$$

Currently, the weak phase of the CKM suppressed $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ decay mode is not fully understood. However, using vertex counting it can be estimated that the $b \to u\bar{u}s$ tree diagram is suppressed. It is therefore expected that the dominant contribution to this decay arises from penguin diagrams, with the contribution from Γ_H vanishing. Hence this channel is mainly sensitive to Γ_L.
If contributions from both the light and heavy B meson states are taken into account, the resulting double exponential distribution can be expressed as

$$\Gamma[f,t] = Ae^{-\Gamma_L t} + Be^{-\Gamma_H t} = e^{-\Gamma t}\left[(A+B)\cosh\frac{\Delta\Gamma_s t}{2} + (B-A)\sinh\frac{\Delta\Gamma_s t}{2}\right]. \tag{1.44}$$

If this is fitted with a single exponential $\Gamma[f,t] = \Gamma_f e^{\Gamma_f t}$ using an unbinned maximum likelihood fit, the result for the effective lifetime $\tau_{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-}$, as shown in [34], is found to be of the form

$$\tau^{-1}_{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-} = \Gamma_f = \frac{A/\Gamma_L + B/\Gamma_H}{A/\Gamma_L^2 + B/\Gamma_H^2} \tag{1.45}$$

with

$$A = 1 - \mathcal{A}_{\Delta\Gamma} \quad \text{and} \quad B = 1 + \mathcal{A}_{\Delta\Gamma}. \tag{1.46}$$

Following the discussion in [35], this effective lifetime can be expressed as

$$\frac{\tau_{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-}}{\tau_{B_s}} = \frac{1}{1-y_s^2}\left[\frac{1+\mathcal{A}_{\Delta\Gamma}y_s + y_s^2}{1+\mathcal{A}_{\Delta\Gamma}y_s}\right] = 1 + \mathcal{A}_{\Delta\Gamma}y_s + (2-\mathcal{A}_{\Delta\Gamma}^2)y_s^2 + \mathcal{O}(y_s^3), \tag{1.47}$$

where

$$y_s = \frac{\Delta\Gamma_s}{2\Gamma_s} \tag{1.48}$$

and

$$\tau_{B_s}^{-1} = \Gamma_s = \frac{\Gamma_H^{(s)} + \Gamma_L^{(s)}}{2}. \tag{1.49}$$

Taking the Standard Model (SM) prediction for $\mathcal{A}_{\Delta\Gamma}(B_s \to K^+K^-) = -0.972^{+0.014}_{-0.009}$ from [35] and $\Delta\Gamma_s/\Gamma_s = 0.123 \pm 0.017$ from [36], the ratio can be evaluated as

$$\frac{\tau_{\mathrm{B_s^0 \to K^+K^-}}}{\tau_{B_s}} = 0.940 \pm 0.008 \tag{1.50}$$

Combined with the measurement $\tau_{B_s} = (1.516 \pm 0.011)$ ps [36], the prediction for the effective lifetime in the channel $\mathrm{B_s^0 \to K^+K^-}$ is

$$\tau_{\mathrm{B_s^0 \to K^+K^-}}(SM) = (1.425 \pm 0.016)\ \text{ps}. \tag{1.51}$$

Figure 1.7 from [35] illustrates the dependence of $\tau_{\mathrm{B_s^0 \to K^+K^-}}/\tau_{B_s}$ on the B_s mixing phase and the implication of a measurement with 1% uncertainty. The relationship between the value $\tau_{\mathrm{B_s^0 \to K^+K^-}}/\tau_{B_s}$ and Φ_s depends on Equations 1.34, 1.47 and 1.48, which to first order gives

$$\frac{\tau_{\mathrm{B_s^0 \to K^+K^-}}}{\tau_{B_s}} = 1 + \frac{\mathcal{A}_{\Delta\Gamma}|\Gamma_{12}|\cos\Phi_s}{\Gamma_s} \tag{1.52}$$

The value of Φ_s is varied to determine the effect on $\tau_{\mathrm{B_s^0 \to K^+K^-}}/\tau_{B_s}$, with the 1% error band dominated by the values of $\Delta\Gamma_s^{SM}/\Gamma_s$.

1.5 Summary

This chapter presented a basic overview of the theoretical framework that describes particle physics, with particular emphasis given to the area of flavour physics and the motivations behind the primary measurements performed in this thesis. An outline of the SM is presented in section 1.2, which discusses the particle types and their interactions combined with a short discussion of the mathematical formulation of the SM Lagrangian that facilitate these interactions. A discussion of the key areas of flavour physics that are important to the analyses performed in this thesis, are discussed in section 1.3. These include the generational nature of fermions, specifically quarks and their relationship via the CKM matrix. A description of CP-violation and the phenomenon of neutral meson mixing, which is important to

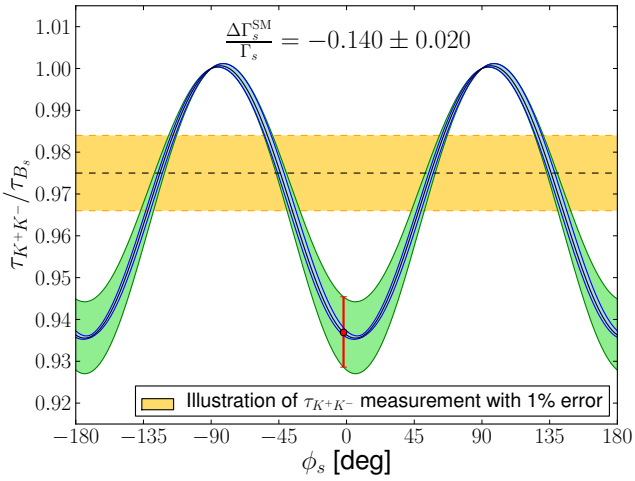


Figure 1.7: Dependence of $\tau_{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-}/\tau_{B_s}$ on the B_s mixing phase. This distribution illustrates the impact of a measurement with 1% uncertainty [35].

measurements made in this thesis, is also provided. Finally, in section 1.4, the $\mathrm{B}\to\mathrm{h}^+\mathrm{h}^-$ family of decays are introduced with particular focus given to lifetime measurements of the $\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-$ decay and its relationship to the parameters $\Delta\Gamma_\mathrm{s}$ and $\mathcal{A}_{\Delta\Gamma}$. A measurement of $\tau_{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-}$ and $\mathcal{A}_{\Delta\Gamma}$ are presented in Chapter 6.

Chapter 2

The LHCb Experiment and Detector

2.1 The Large Hadron Collider at CERN

The Large Hadron Collider (LHC) [37] (Figure 2.1) is a particle accelerator located on the Franco-Swiss border at the CERN research facility [38] near Geneva, Switzerland. The accelerator uses the same tunnel as the previous LEP [39] collider, and is driven by the requirement to further probe the SM of particle physics at the electroweak symmetry breaking scale and beyond. This also includes testing new theories and determining if the recently observed new particle at the LHC [12] [13] is the hypothesised Higgs Boson [14].
The LHC is located underground at a mean depth of $100\,\mathrm{m}$, which due to geological considerations has a slope of 1.4% [41], and is housed in a tunnel of circumference 26.7 km [37] ($\approx$ 17miles). The actual depth varies from $50\,\mathrm{m}$ to $175\,\mathrm{m}$, with this depending on the tunnels position relative to Lake Geneva or the Jura mountain range respectively. The accelerator itself is a proton-proton collider that has the potential to accelerate bunches of 1.1×10^{11} particles up to a Centre of Mass (CM) energy of 14 TeV, with the world's highest luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. It has also been designed to collide heavy (Pb) ions up to an energy of $2.8\,\mathrm{TeV}$ per nucleon reaching a peak luminosity of $10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Along the collider ring are located six main detector experiments, these are: ATLAS [42], CMS [43], LHCb [44], ALICE [45], TOTEM [46] and LHCf [47].

ATLAS and CMS

ATLAS and CMS are two multipurpose detectors, which are designed to run at the peak luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. These experiments are primarily engaged in Higgs and BSM searches. Additional topics of interest include: top physics, B physics, heavy ions and

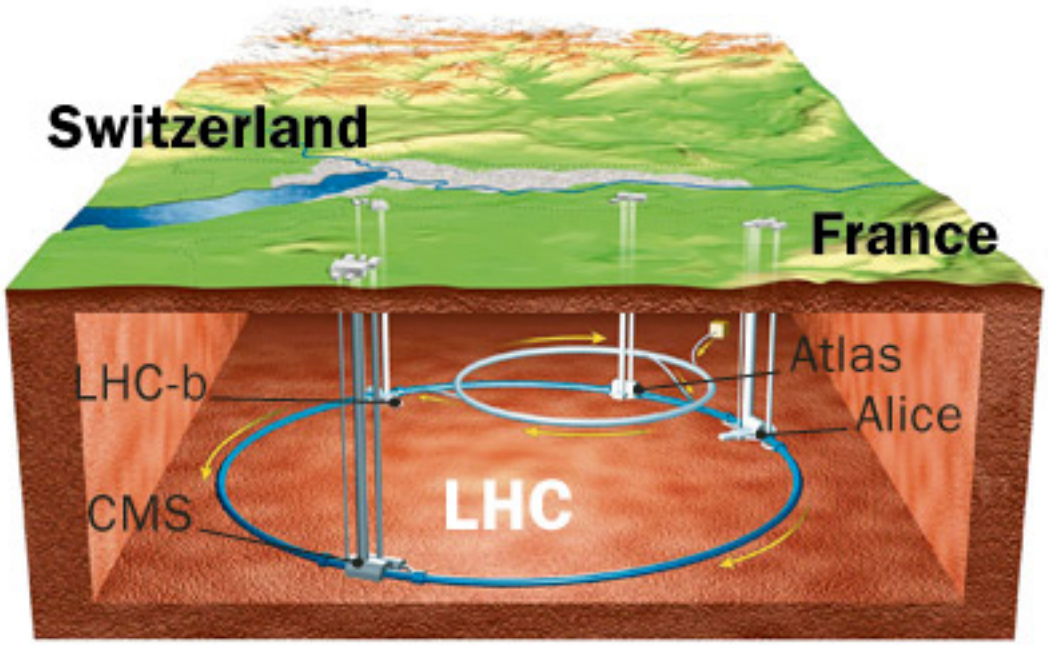


Figure 2.1: The LHC accelerator and experiments at CERN [40].

searching for evidence of dark matter and extra dimensions.

ALICE

ALICE is a dedicated heavy ion experiment studying the quark-gluon plasma that existed shortly after the big bang, by colliding heavy lead ions. It aims to reach, and use, a peak luminosity of $10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ with 156 bunches.

LHCb

LHCb is primarily designed to focus on heavy flavour physics, in particular the study of CP-violation and rare decays, using a peak luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. A more extensive description will be discussed in the remainder of this chapter.

TOTEM

TOTEM is one of the two smaller experiments that is part of the LHC. Although technically integrated with CMS, it can also be operated as a standalone experiment. Its purpose is to measure the total pp cross-section with a luminosity-independent method, and study elastic and diffractive scattering at the LHC.

LHCf

LHCf is an experiment dedicated to the measurement of neutral particles emitted in the very forward region of LHC collisions. The physics goal is to provide data for calibrating the hadron interaction models that are used in the study of Extremely High-Energy Cosmic-Rays.

2.2 LHCb Experiment

The LHCb detector [44] is one of the four largest experiments (alongside ALICE [45], ATLAS [42] and CMS [43]) which have been designed to run at the LHC [44] during its operational lifetime. The LHCb experiment is designed specifically to look at heavy flavour physics, and more specifically B-physics. The main physics goals are:

- Study CP-violation in charm and beauty decays
- Precision measurements of Standard Model observables
- Study of rare decays.

All of these areas will contribute to the search for NP beyond the SM.
At the time of writing, all heavy flavour physics studies conducted by other accelerators around the world (the TEVATRON, BABAR, BELLE, CLEO, etc) have achieved results that largely prove to be consistent with the SM predictions and the CKM mechanism.
Due to the large $\mathrm{b\bar{b}}$ cross section, $\sigma(pp \to b\bar{b}X)$ = $(282 \pm 20 \pm 49\,\mu\mathrm{b})$ at $7\,\mathrm{TeV}$ [48], the LHCb is designed to be the most abundant source of B-hadrons in the world producing a full spectrum of B^0_d, B^0_s, Λ_b's and charm hadrons. At the LHCb design luminosity [49] of $2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, $10^{12}\,\mathrm{b\bar{b}}$ pairs are produced in $10^7\,\mathrm{s}$ which equates to one year of data taking. The impact of the luminosity cannot be underestimated, as running at a lower luminosity

with respect to the other LHC experiments means that produced events are dominated by single pp interactions per bunch crossing. As of 2011, the typical number of visible interactions per bunch crossing is 1.7. This type of event is significantly easier to analyse compared with multiple interaction events, as well as reducing radiation damage and keeping the detector occupancy at a low level. The LHCb also has the technical capability to regulate its luminosity. The regulation is performed by displacing the beams, i.e. not colliding them head on, around the interaction point. The focusing (β^*) is constant throughout each fill such that this provides the optimal experimental luminosity. In 2012, the design luminosity of the LHCb was actually exceeded as it ran at $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which is twice the design luminosity.
As discussed in section 1.1, the amount of matter in the universe requires a far larger level of CP-violation than the SM currently predicts. It is therefore hypothesised that new mechanisms of CP-violation are needed to account for what we observe, with evidence of this potentially being observable through heavy flavour physics analysis. Extended BSM models of new physics include extra factors that could contribute to changes in CP violating phases, and enhance rare decay branching fractions. These can be examined by doing high statistics studies of common decays, which give access to the CP violating phases, and rare decays of charm and beauty hadrons.
In order for the LHCb to fully utilise its capacity to produce B-hadrons effectively, a variety of both hardware and software solutions must be employed. This includes reducing the large amount of background that is produced in a high occupancy hadron collider environment, by the introduction of both hardware and software triggers (discussed further in 2.10). Excellent vertex and momentum resolution are also required for high precision proper-time measurements, as well as identification of final state hadrons in order to cleanly reconstruct B and D hadron decay modes. These will all be discussed in more detail in the following sections.

2.3 The LHCb Detector

The LHCb detector is a forward facing spectrometer and is designed to cover a small region of the full solid angle around the beam line ($\approx 2\%$ [37]). Unlike the other experiments at the LHC the collision point is not located directly in the centre of the detector, but on one side. The configuration of the detector itself is justified by the production directions of the $\mathrm{b\bar{b}}$ pairs. The kinematics involved in the pp interactions in the bunch crossings cause $(26.56 \pm 7.02)\%$ [48] of all $\mathrm{b\bar{b}}$-hadrons produced to be in the same forward or backward cone, see Figure 2.2. These B hadrons are highly forward (or backward) boosted such that they are almost collinear with the beams, and within the forward acceptance region of the spectrometer.

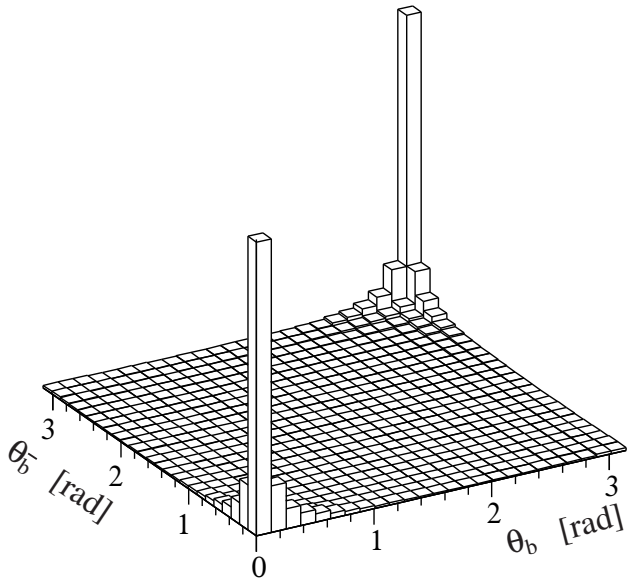


Figure 2.2: $b\bar{b}$ production angles, relative to the beam, from proton-proton collisions at 14 TeV [50].

Figure 2.3 shows a schematic drawing of the LHCb detector. A right handed coordinate system is chosen, with the z-axis mapped to the direction of the beam and the y-axis being along the perpendicular vertical.

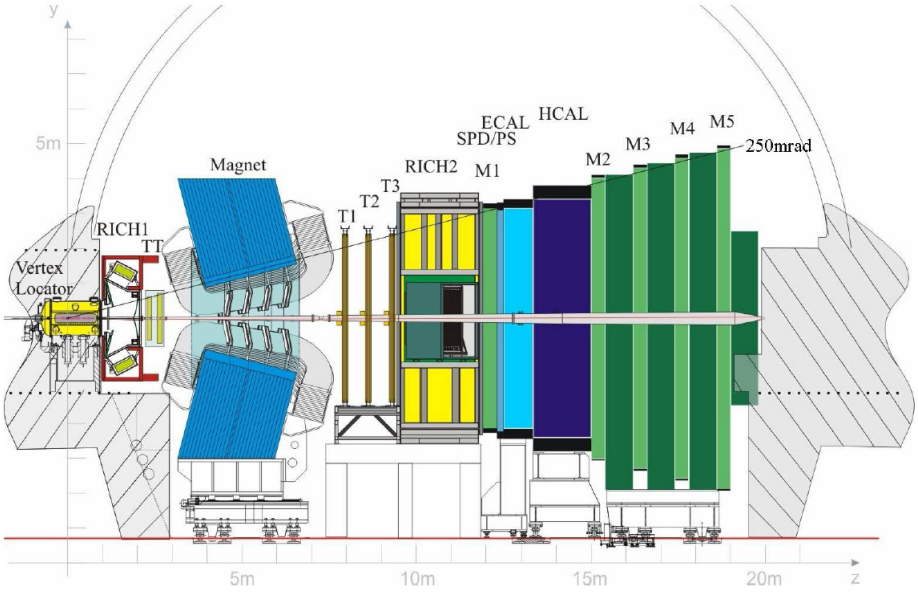


Figure 2.3: LHCb detector geometry [44].

The detector itself has some generic detector components, much as all high energy particle detectors generally have. These comprise the tracking detectors described in Section 2.5, a

magnet in Section 2.6, the electromagnetic and hadronic calorimeters in Sections 2.8.2 and 2.8.3 respectively, and muon chambers in Section 2.9. However, it also has more specific sub-detector components such as the vertex locator in Section 2.4 and the Cherenkov detectors in Section 2.7. It is these that give the LHCb the capability to make high precision measurements. Each subdetector will be discussed in the following sections.

2.4 VErtex LOcator (VELO)

The VELO subdetector is one of the main components of the LHCb tracking system together with the Silicon Trackers (ST) described in section 2.5. These combine to provide excellent spatial resolution and track reconstruction. Specifically, the VELO reconstructs the Primary Vertex (PV) from the pp collisions, and the secondary vertices from long lived particles such as B and D mesons. The VELO was constructed to be able to resolve and identify displaced secondary vertices, which are a distinctive feature of b and c-hadron decays. It is able to accomplish this by being located within $7\,\mathrm{mm}$ of the interaction point. This short distance to the beam is important as it minimises the extrapolation distance to the vertices, thus improving the spatial resolution. The B-mesons themselves have a lifetime of approximately $1.6\,\mathrm{ps}$, such that they travel typically $1\,\mathrm{cm}$ before decaying. This configuration of the detector means they are more easily detected, with the mother and daughter decay vertices able to be reconstructed with a higher efficiency.

2.4.1 VELO Design

The design of the VELO detector has been carefully considered due to its difficult working environment. As the sensors are mounted perpendicular to the beam, they receive a highly non-uniform radiation dose so must therefore be constructed of a radiation tolerant material. The maximum radiation exposure of the sensors at nominal conditions per year is comparable to a neutron fluence of $1.3 \times 10^{14}\ \mathrm{n_{eq}}/\,\mathrm{cm}^2$, where $\mathrm{n_{eq}}$ is $1\,\mathrm{MeV}$ equivalent fluence. The material choice was n-implants in n-bulk technology, with strip isolation given by the use of a p-spray. This material was chosen as it evolves over time until it undergoes type inversion due to the radiation exposure. This process actually extends the lifetime of the sensors, with close monitoring of the voltage required to maintain constant performance. Prototype sensors made of this material were developed and qualified up to a fluence corresponding to three years nominal conditions. The VELO is also designed to move such that the sensors, which will be described, are retracted away from the beam during injection. The sensors are then closed around the beam, using tracks to locate the luminous region, once collisions

begin.

The VELO is composed of 42 individual semicircular silicon modules, Figure 2.4, with each incorporating a measure of the r and ϕ coordinates required for track reconstruction. The minimum pitch of the sensor modules is $38\,\mu\mathrm{m}$ at the innermost radius of the sensors with the size increasing linearly up to $101.6\,\mu\mathrm{m}$ at the outer radius of $41.9\,\mathrm{mm}$, this is illustrated in Figure 2.5

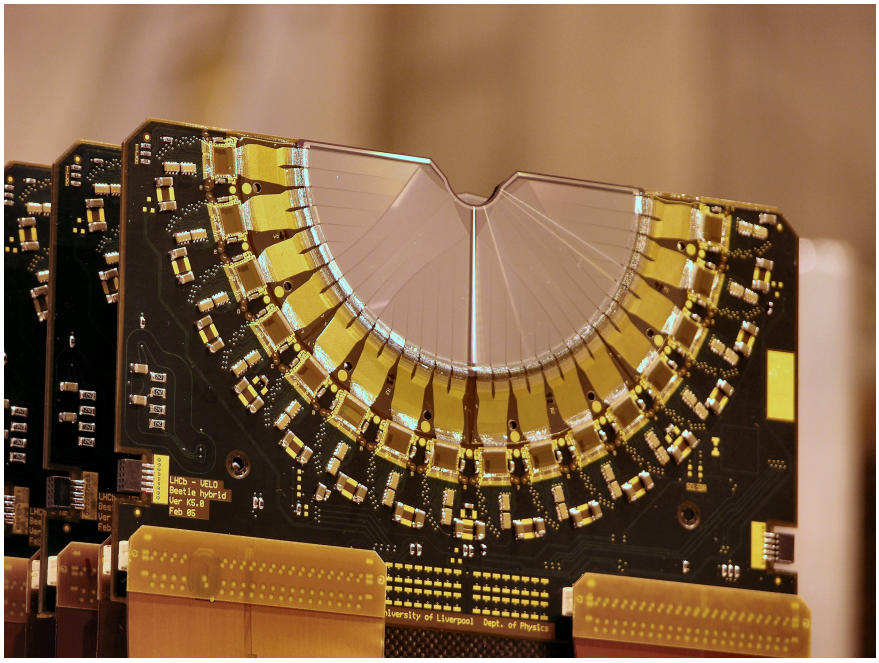


Figure 2.4: One of the 48 VELO sensors [51].

The sensors themselves are a composite of R- and ϕ- sensors, see Figure 2.5, and are attached in pairs of R and ϕ to the modules holding the readout electronics. The technology used for both are identical, with both sets of sensors being $300\,\mu\mathrm{m}$ thick and the readout method being the same. The pitch of both increases linearly from the inner radius outwards given by the following relations:

$$R - sensor : 40 + (101.6 - 40) \times \frac{r - 8190}{41949 - 8190} \tag{2.1}$$

$$\phi - sensor : 37.7 + (79.5 - 37.7) \times \frac{r - 8190}{17250 - 8170} \; (r < 17250) \tag{2.2}$$

$$\phi - sensor : 39.8 + (96.9 - 39.8) \times \frac{r - 17250}{42000 - 17250} \; (r > 17250) \tag{2.3}$$

These modules are situated in an evacuated vessel, with the sensors separated from the beam

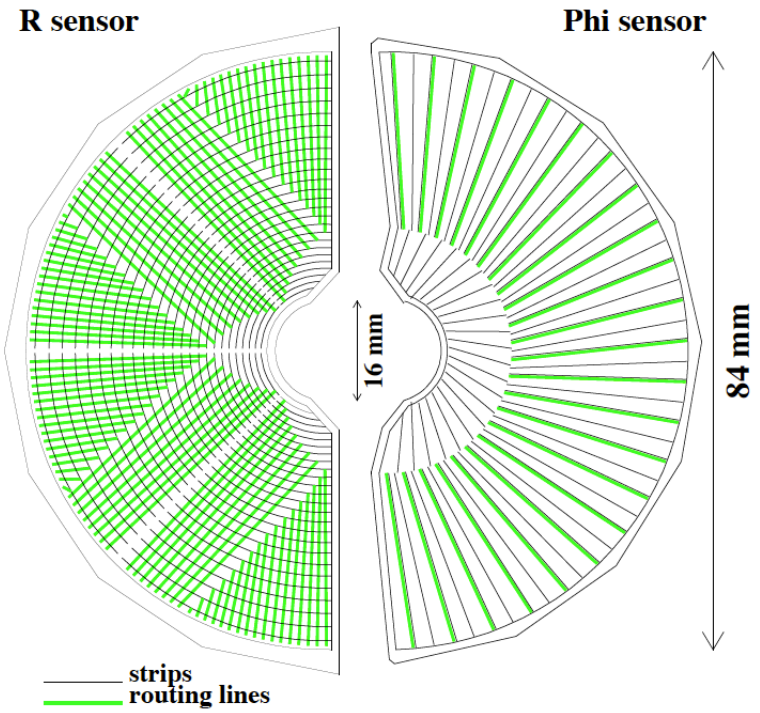


Figure 2.5: $r\phi$ geometry of the VELO sensors [52].

pipe by two thin walled corrugated aluminium sheets, Figure 2.6, called *RF-foils*. Aluminium was chosen due to its low Z, such that the particles produced in the collisions have minimal material to traverse and less energy loss will occur. Of the 42 sensors positioned along the beam line, 21 of these are located on either side of the beam axis, with a pair of modules opposite one another being known as a station. Geometrically, the VELO is required to cover the same angular acceptance as the downstream detectors, thus detecting particles in the pseudorapidity region $1.6 < \eta < 4.9$. To cover the full azimuthal acceptance, the two detector halves are required to overlap. The sensor halves are offset along the z-axis by $1.5\,\mathrm{cm}$, which allows the two halves to overlap when the VELO is in the closed position. Once the VELO is closed, a track is only reconstructed when a particle has crossed at least 3 VELO stations. The stations are densely packed around the IP to reduce the average extrapolation distance from the first measured hit to the vertex.
The detector design was limited by a number of constraints. The requirements for the VELO performance were:

- to aim for an initial signal to noise (S/N) ratio of > 14;
- to have an overall channel efficiency of at least 99%;
- to maintain a spatial resolution of $\approx 4\,\mu\mathrm{m}$ for $100\,\mathrm{mrad}$ tracks that is not degraded by

radiation exposure.

Full details of the VELO design are provided in [44].

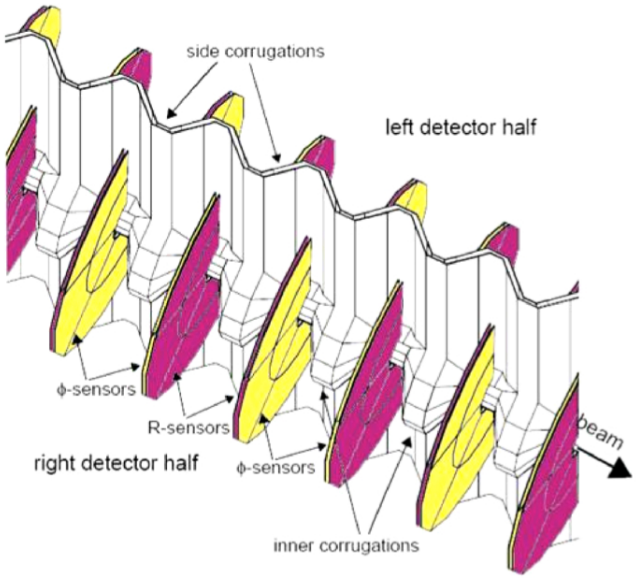


Figure 2.6: Inside the RF Foil [44].

2.4.2 VELO Performance

The performance of the VELO is very important to all LHCb physics analyses. The VELO arrangement was carried out to minimise the material within its acceptance, while providing good geometrical coverage. As mentioned previously, all events that are reconstructed must traverse at least three VELO modules. Individual resolution of the sensors has been performed using test beams. These have shown the resolutions to be a strong function of the projected angle and sensor pitch. The VELO performance during the delivered luminosity of 1.2 fb^{-1} of data over the first two years of LHCb operation has shown a number of interesting features [53]. The best single hit resolution achieved from studies of the 2011 data is $4\,\mu\mathrm{m}$ for a $40\,\mu\mathrm{m}$ pitch, with the Impact Parameter (IP) resolution found to be better than $15\,\mu\mathrm{m}$ for high p_T tracks, Figure 2.7.

The IP resolution feeds directly into the resolution of the PVs, which is found to be below $15\,\mu\mathrm{m}$ for both the x and y coordinates of the VELO (where z is the along the beam axis) when using greater than 25 tracks in a fit. Other performance characteristics of note include; The signal to noise of the sensors being $\sim\ 20\ :\ 1$, with the best hit resolution observed at the optimal track angle. The track finding efficiency of $\sim 98\%$ or above. The decay time resolution observed for the channel $\mathrm{B}_\mathrm{s}^0 \to \mathrm{J}/\psi\,\phi$ of $50\,\mathrm{fs}$, which is due to the impact parameter resolution of $35\,\mu\mathrm{m}$ for particles of transverse momentum greater than $1\,\mathrm{GeV}$.

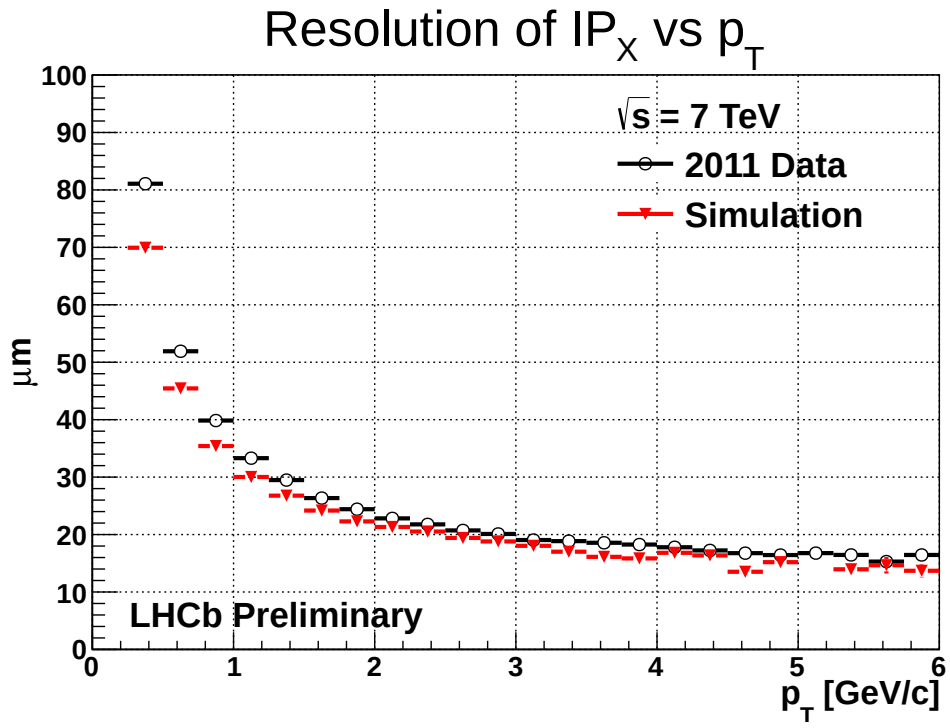


Figure 2.7: Measurement of VELO impact parameter resolution with increasing $\mathrm{p_T}$, from simulation and 2011 data [54].

2.5 The Trackers

The Silicon Trackers are part of the LHCb tracking system coupled with the VELO (section 2.4) and consist of four planar tracking stations. The Tracker Turicensis (TT) is located upstream of the dipole magnet with the T1-T3 stations downstream of the magnet.
As with the VELO, the TT also uses silicon strip detectors, these detectors are also partially used in the T1-T3 downstream trackers. The Inner Tracker (IT), which covers the region closest to the beam pipe, is part of the T1-T3 downstream trackers and uses these silicon strip detectors. The remaining downstream trackers covering the outer region use straw tubes. Both the TT and IT were developed commonly as the Silicon Tracker (ST) .

2.5.1 Silicon Tracker

The ST, which comprises both the TT and IT, are both silicon strip detectors as mentioned previously, with a strip pitch of approximately $200\ \mu\mathrm{m}$. Each of the four tracker sub-detectors has been built using four detection layers in an (x-u-v-x) arrangement, displayed in Figure 2.8. The strips are organised with vertical strips in both the first and last layers, while in the second and third layers there is an offset rotation of -5° and $+5^\circ$ respectively. The single

hit spatial resolution of these detectors is about $50\,\mu\mathrm{m}$. As with all the LHCb sub-detectors, radiation damage must also be taken into account. For 10 years of nominal luminosity, 1 MeV neutron equivalent fluences are expected not to exceed $5 \times 10^{14}\,\mathrm{cm}^{-2}$ in the innermost regions of the TT and $9 \times 10^{12}\,\mathrm{cm}^{-2}$ in the IT.

The TT is an upstream planar tracking station, with dimensions of $150\,\mathrm{cm}$ in width and $130\,\mathrm{cm}$ in height. This tracker is constructed of 143360 readout strips of length up to $38\,\mathrm{cm}$, the total active area is about $8.4\,\mathrm{m}^2$. The entire sub detector is housed in a large, light tight, thermal and electrically insulated detector volume that is kept at a constant temperature of $+5°C$. This volume is also continually flushed through with nitrogen to prevent condensation on the cold surfaces within it.

The IT, being downstream and part of the T1-T3 tracking stations, covers a $120\,\mathrm{cm}$ wide and $40\,\mathrm{cm}$ high cross shaped region in the centre of these tracking stations, Figure 2.9. This particular tracking station is assembled with 129024 readout strips of length either $11\,\mathrm{cm}$ or $22\,\mathrm{cm}$, with a total active area of $4.0\,\mathrm{m}^2$. To avoid any acceptance gaps, adjacent modules in the detection layer are staggered by $4\,\mathrm{mm}$ in the z direction and $3\,\mathrm{mm}$ in the x direction so as to overlap.

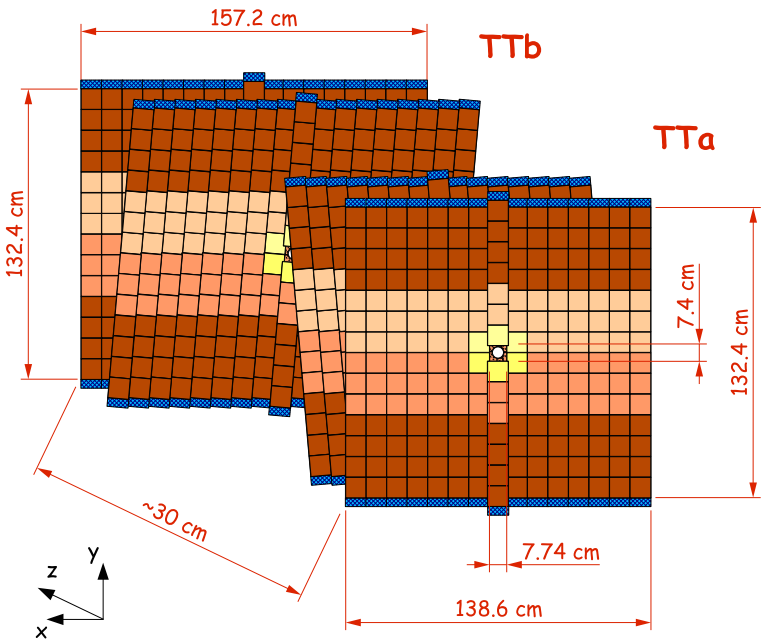


Figure 2.8: Tracker Turicensis Schematic [55].

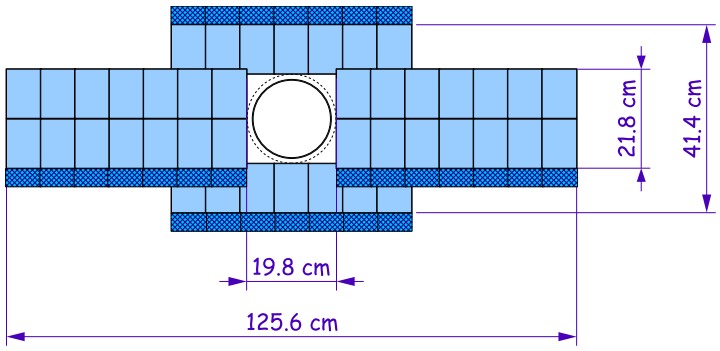


Figure 2.9: Inner Tracker Schematic [55].

2.5.2 Outer Tracker

The remaining part of the T1-T3 tracking stations comprise the Outer Tracker OT. The LHCb OT is a drift time detector, which tracks charged particles allowing for the measurement of their momentum over a large acceptance. For the precise determination of the invariant mass of reconstructed b-hadrons, excellent momentum resolution is required; a mass resolution of $10\,\mathrm{MeV}/c^2$ of the $\mathrm{B}_s^0 \rightarrow \mathrm{D}_s^- \pi^+$ decay equates to a required momentum resolution of $\delta p/p \approx 0.4\%$.

The OT is designed as an array of gas-tight, individual straw tube modules (displayed in Figure 2.10). Each of these modules contains 2 layers of drift tube which are staggered from each other. To introduce a drift-coordinate resolution of 200 $\mu\mathrm{m}$, a gas mixture comprising 70% Argon and 30% CO_2 is used. This composition of gas also has the additional properties that guarantee a fast drift time of below $50\,\mathrm{ns}$.

2.6 Dipole Magnet

In order to measure the momentum of charged particles, the LHCb utilises a large dipole magnet, where the schematic (Figure 2.11) is displayed on the same scale as the detector (Figure 2.3). The forward acceptance of $\pm 250\,\mathrm{mrad}$ in the vertical and $\pm 300\,\mathrm{mrad}$ in the horizontal is used to make the measurement.

A warm magnet design was used instead of a super-conducting magnet originally proposed in the Technical Proposal [50], primarily due to investment costs and construction times. The design for the magnet, with an integrated magnetic field of $4\,\mathrm{Tm}$ for tracks of length $10\,\mathrm{m}$, has to accommodate the differing requirements of the sub-detectors in the field. The RICH requires the magnetic field strength to be less than $2\,\mathrm{mT}$ within its envelope, while the field everywhere else and in particular within the region between the VELO and trackers, must

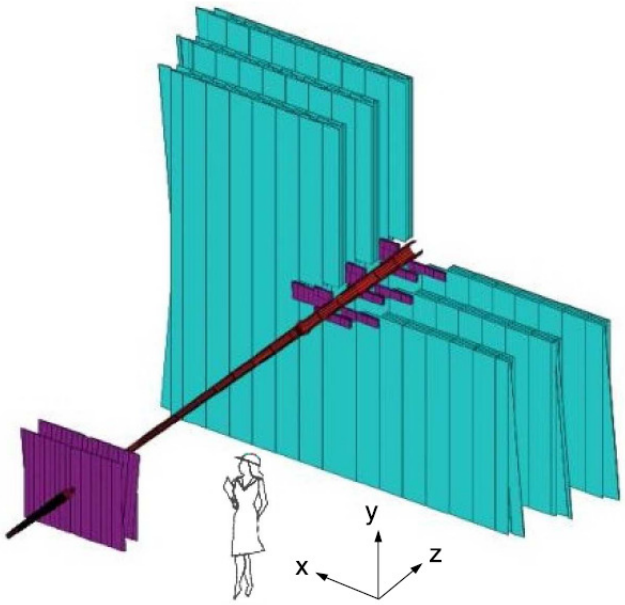


Figure 2.10: Illustration of the tracker configuration. The TT and OT are shown in purple, with the OT in blue and beam-pipe in red [44].

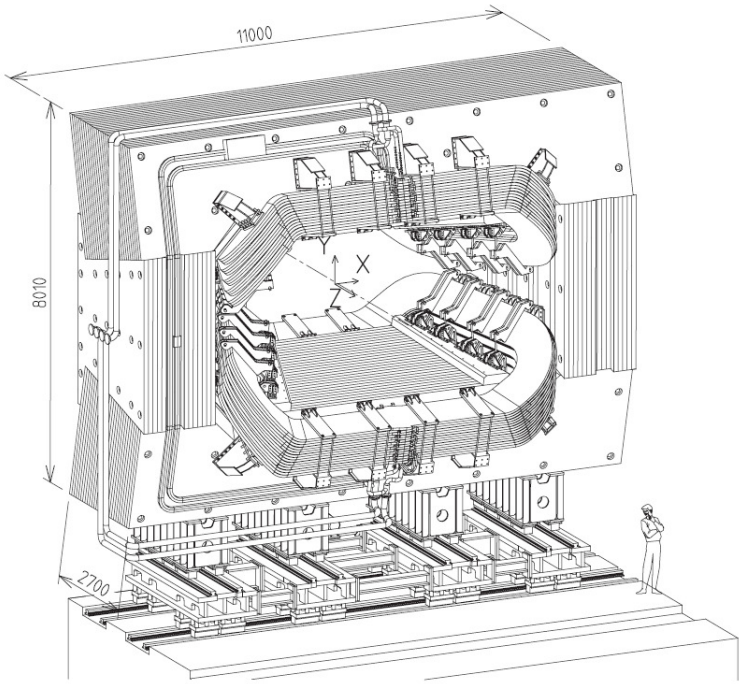


Figure 2.11: Dipole Magnet Schematic [44].

be as strong as possible. The two coils situated in the magnet yoke are saddle shaped and have been fitted symmetrically about each other. Each of the coils are made of 15 'pancakes' which are then further broken up into 5 triplets. These triplets are produced from a pure

Al-99.7 hollow conductor which has a central cooling channel to aid in thermal conduction. Cast aluminium clamps are then used to hold together the triplets that make up the full coil and hold it in its correct position. The pancakes that comprise the coils are left free to move and slide around on their supports, as thermal stress is one of the main constraints within the magnet. To keep this movement stable, only one of the coils extremities is left fixed on the symmetric axis. The total weight of the yoke is 1500 tons with the two coils having a combined weight of 54 tons.

One of the primary objectives for LHCb is to measure CP asymmetries. It is very important to be able to control systematic uncertainties that are inherent to the detector, with the primary method of correcting for this being to periodically reverse the magnetic field. However, hysteresis effects within the magnet may affect the reproducibility of the magnetic field. Another important measurement for the magnet is the measurement of the momentum of a charged particle. In order to achieve the required momentum resolution, the magnetic field integral has to be measured to a relative precision of about 10^{-4}, with the position of the B-field peak being located within an accuracy of a few mm. To map the three components of the magnetic field, a semi-automatic device constructed of multiple Hall probes was used that could be positioned in both the horizontal and vertical direction to cover the areas of interest within the field. The device itself was composed of 60 sensor cards, each containing three Hall probes which were mounted orthogonally on a cube together with a temperature sensor and readout electronics.
A complete measurement of the magnetic field within the tracking regions and also inside the magnetic shielding for the RICH photon detectors was carried out. The precision of the field measurement made within the tracking region was 4×10^{-4}, with the absolute field being reproducible for both magnet polarities to better than this value.

2.7 Ring Imaging CHerenkov (RICH) Detectors

One of the fundamental attributes required by LHCb is the ability to identify the final state particles. Lepton final states are determined by other sub-detectors, electrons by the ECAL and muons with the muon chambers, with light charged hadronic final states (i.e. pions, kaons and protons) being determined primarily by the RICH detectors. These detectors utilise the Cherenkov effect to determine the velocity of a particle within a medium.
The Cherenkov [56] effect occurs when a charged particle traverses a dielectric medium at a speed that exceeds the local speed of light in that medium. This can be described by

$$v > \frac{c}{n}, \tag{2.4}$$

where v is the velocity of the particle, c is the speed of light in a vacuum and n is the refractive index of the medium that the particle is traversing. As the charged particles traverse the medium, they cause the surrounding medium to polarise. However, once the charged particles have moved past the region of polarisation, the constituents of the medium rapidly depolarise emitting a cone of electromagnetic (EM) radiation, Figure 2.12.

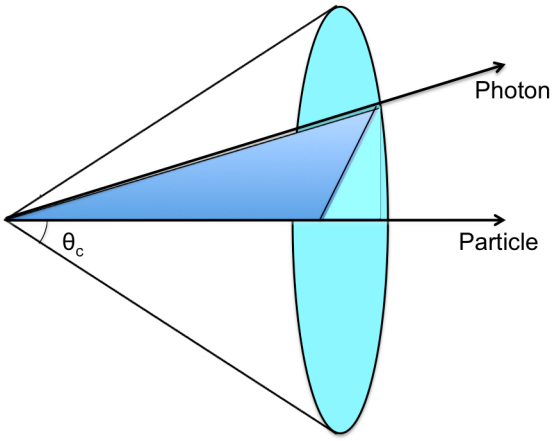


Figure 2.12: Emission of Cherenkov radiation cone.

This cone can be collected and focussed by an optics arrangement onto a photo-sensitive surface. By measuring the light cone produced, the Cherenkov angle θ_c can be determined using

$$\cos\theta_c = \frac{c}{nv}. \tag{2.5}$$

Each species of particle can be determined by the radius of the light cone detected. As, for a given p, each charged particle will produce a ring with a different radius. Thus the species can be inferred using this information.

2.7.1 RICH Design

The LHCb incorporates two RICH detectors into its design. The first RICH detector, RICH 1, is located between the VELO and the Trigger Tracker (see Figure 2.3) with the second RICH, RICH 2, located further downstream after the magnet and the tracking stations (see Figure 2.3). The two detectors are required so as to cover a wide momentum spectrum range and different polar angle regions. The momentum spectrum covered is determined by the radiator medium used, with each RICH using a different medium in order to cover a

specific range which overlaps with the other. RICH 1 (see Figure 2.13) uses a combination of aerogel and fluorobutane (C_4F_{10}) as a radiator medium, this allows it to cover the softer momentum region of approximately $1 \rightarrow 60\,\mathrm{GeV}/c$. This detector also covers the larger polar angle range of $\pm 25\,\mathrm{mrad}$ to $\pm 300\,\mathrm{mrad}$ in the horizontal plane and $\pm 250\,\mathrm{mrad}$ in the vertical plane. The harder momentum spectrum is covered by RICH 2, it uses the radiator medium CF_4 to cover the momentum range $15 \rightarrow 100\,\mathrm{GeV}/c$, where the smaller polar angle acceptance of $\pm 15\,\mathrm{mrad}$ to $\pm 120\,\mathrm{mrad}$ (horizontal) and $\pm 100\,\mathrm{mrad}$ (vertical) is covered.

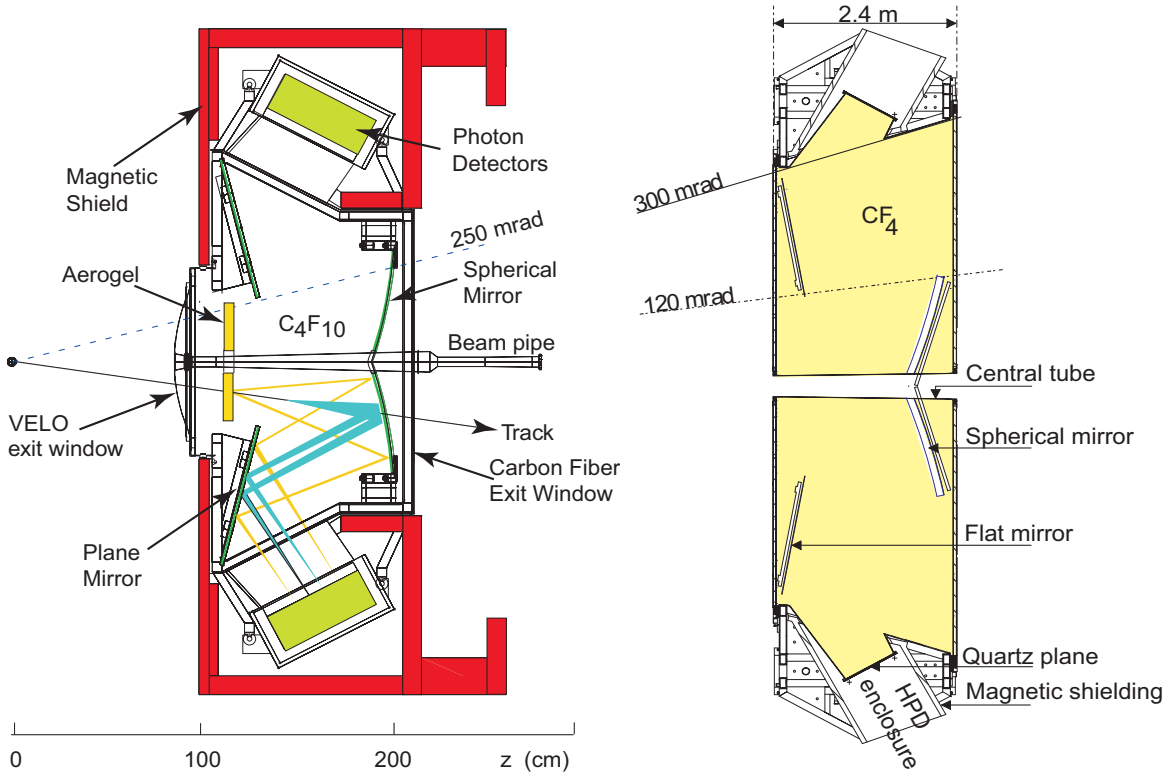


Figure 2.13: RICH 1 (left) and RICH 2 (right) schematics [44].

To collect the emitted Cherenkov light, both RICH detectors use a combination of flat and spherical mirrors to reflect and focus the photons out of the acceptance of the spectrometer; this will be discussed in further detail subsequently in Chapter 3. To detect the Cherenkov photons an array of Hybrid Photon Detectors (HPD), which are able to detect radiation in the wavelength region $200 - 600\,\mathrm{nm}$, are used. These HPDs have been specifically developed for use in the LHCb RICH and are composed of a pixellated reverse-biased silicon detector embedded in a vacuum photon tube (Figure 2.14). In them, a photoelectron is emitted from a conversion process and accelerated through a potential in the vacuum tube towards the silicon detector. The HPDs have been pixellated so as to be able to provide spatial positioning of the detected Cherenkov photons, it is from this that the spherical cones emitted by the particle

species are determined and resolved allowing a Particle IDentification (PID) to occur. The HPDs are arranged and shielded from magnetic field strengths of up to 50mT. This allows them to operate effectively in the LHCb environment.

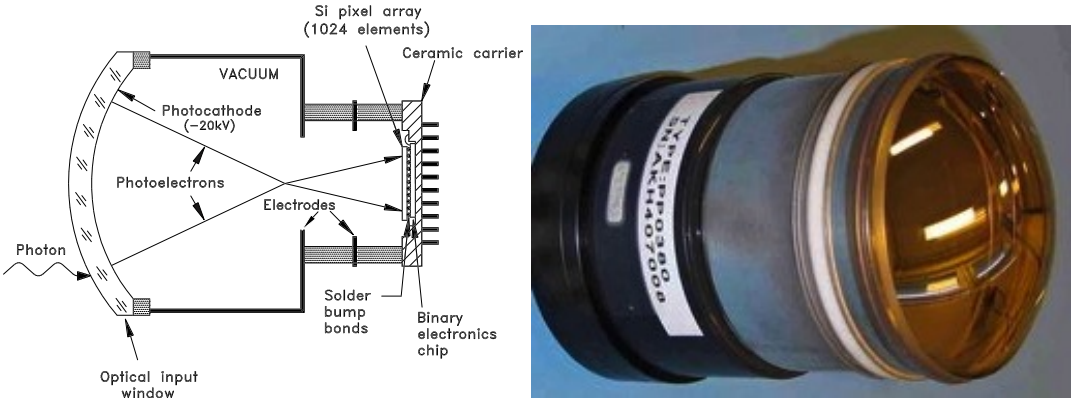


Figure 2.14: RICH HPD schematics [44].

2.7.2 RICH Performance

The performance of the RICH detectors is conducted using data collected up to the close of the 2011 data taking period. The principal purpose is to determine the Cherenkov angle resolutions for both RICH 1 and RICH 2. This is done by selecting high momentum saturated pion tracks, fitting the observed Cherenkov angle θ_c, and determining the offset ($\Delta\theta_c$) of the peak from the its expected nominal value (discussed further in Chapter 3). The θ_c resolutions achieved are 1.618 ± 0.002 mrad for RICH 1 and 0.68 ± 0.02 mrad for RICH 2. The distributions of $\Delta\theta_c$ for RICH 1 and RICH 2, calculated for each photon with respect to the measured track, are displayed in Figure 2.15.

Full PID is performed using the combined information from the two RICHs, the calorimeters (Section 2.8) and muon chambers (Section 2.9), with hadron identification using specifically the RICH. The method used to identify the final state particle species is referred to as the *global pattern-recognition*, this is because it considers all the radiators used in the two RICH detectors and all found tracks for an event. The implemented algorithm is based on a log-likelihood approach, which uses expected distributions from reconstructed tracks under different particle hypotheses. These are compared with the hit pixels in the RICH photon detectors to determine a match. While constructing this likelihood, the method is also required to calculate the effective emission angle for all the pixel-track combinations that could be associated through Cherenkov radiation. By varying the particle hypotheses between $\mathrm{e}, \mu, \pi, \mathrm{K}$ and p for each track, the particle species hypothesis giving the most probable match to the

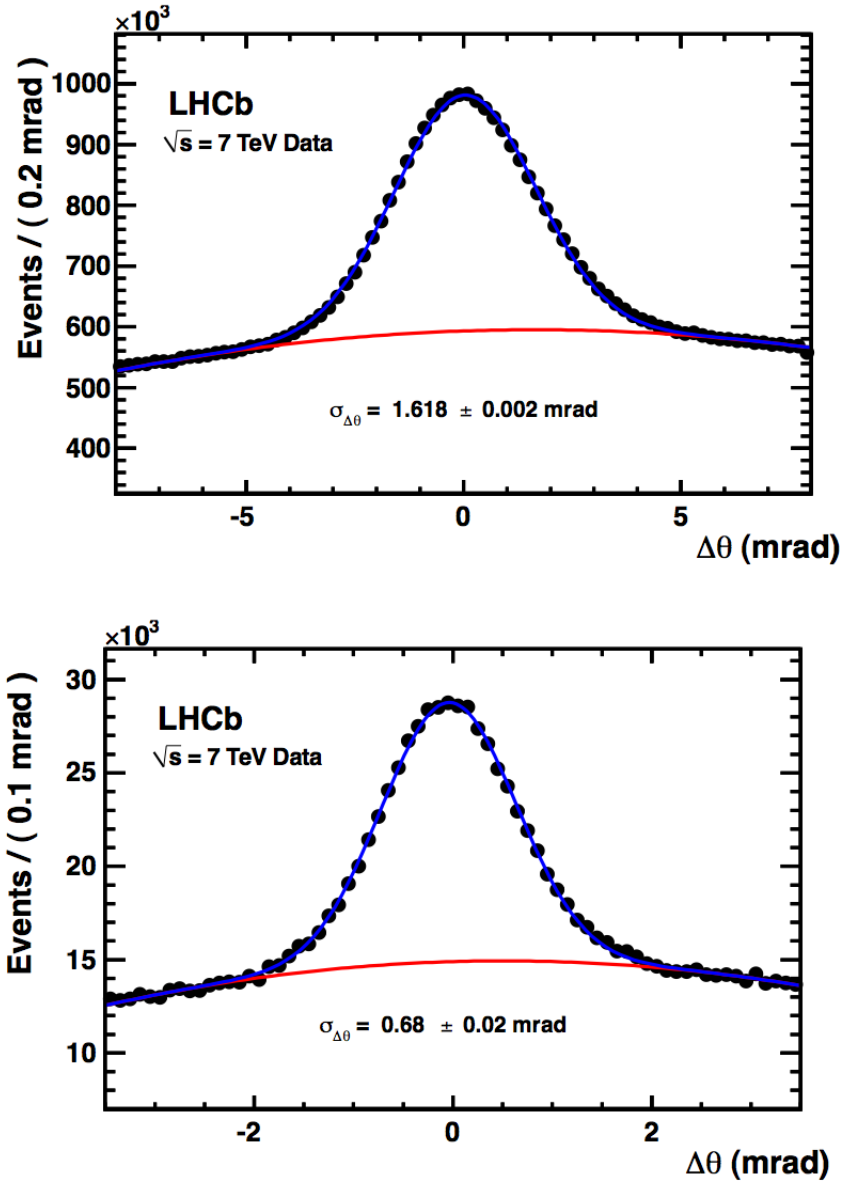


Figure 2.15: Cherenkov angle resolution for RICH 1 (top) and RICH 2 (bottom), determined from high momentum charged particle data collected during the 2011 data taking run. The red line describes the background and the blue line describes the signal [57].

observed Cherenkov ring is chosen. The PID performance of the RICH is demonstrated in Figure 2.16. The requirements for each track with the kaon hypothesis is that the likelihood of each track be greater than those with the pion hypothesis, so $\Delta\log\mathcal{L}(\mathrm{K}-\pi) > 0$. The average efficiency for the identification of a kaon in the range 2 to 100 GeV/c is $\approx 95\%$ with the corresponding pion mis-ID of $\approx 10\%$. By increasing the previous requirement on the tracks to $\Delta\log\mathcal{L}(\mathrm{K}-\pi) > 5$, the kaon efficiency becomes $\sim 85\%$ with the pion misidentification rate decreasing to $\sim 3\%$ [57].

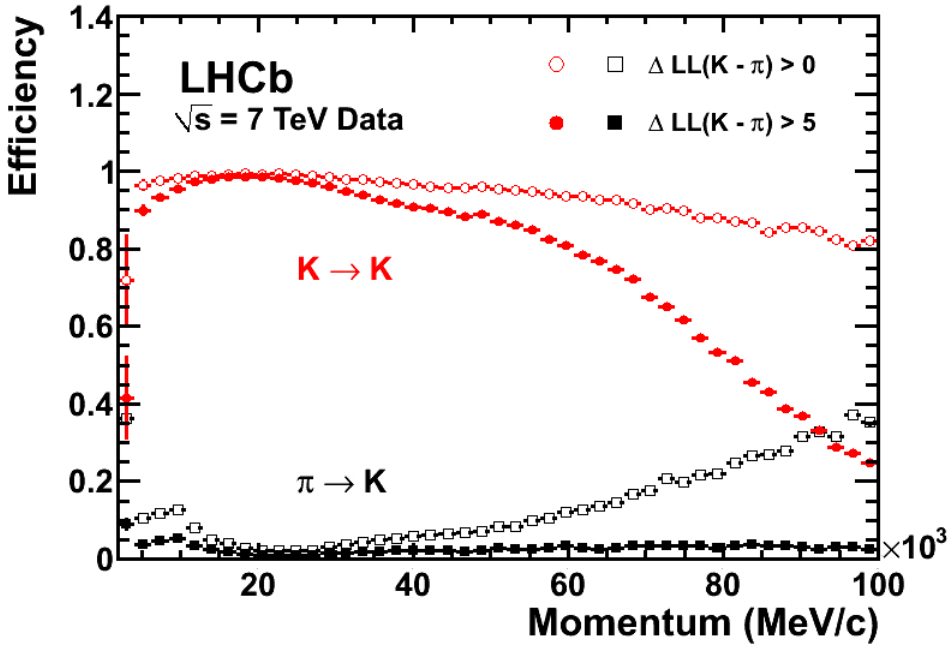


Figure 2.16: Kaon identification efficiency and pion misidentification rate measured on data as a function of track momentum [57].

2.8 Calorimeters

The LHCb Calorimeters play an important role in the physics analysis as good reconstruction efficiency is required for π^0 and prompt photons, in order to perform flavour tagging in CP measurements and for reconstruction of electromagnetic final states. This also enforces the need for good background rejection and B-decay efficiency. The Calorimeters are located downstream of the tracking and RICH subdetectors previously discussed, and are only succeeded by the Muon chambers as displayed in Figure 2.3. The Calorimetry system has been classically designed to incorporate the Electromagnetic Calorimeter (ECAL) followed by the Hadron Calorimeter (HCAL). The function of the Calorimeters is to identify electrons, photons and hadrons which it is able to do by measuring their energy and position. By selecting transverse energy hadron, electron and photon candidates to pass the first trigger level (L0) decision, a decision on the candidate can be made within a 4 μs window after the interaction has occurred. The full performance of the Calorimeters is determined from offline analysis.

2.8.1 Calorimeter Design

Particle physics calorimeters are all designed to function using the same basic method. The ECAL is designed to measure the energy of particles that interact primarily via the electro-

magnetic interaction, whereas the HCAL is designed to measure the energy of particles that interact via the strong force. The design of the ECAL and HCAL allows these functions to be performed, employing a variety of methods even with the imposition of demanding conditions. The Level 0 (L0) electron trigger is required to remove 99% of the inelastic pp interactions to enrich the data sample, this is due to the limitation of the bandwidth allocated to it. To do this, electrons are selected with large transverse energy, E_T. In order to reject the other large background composed of charged pions a Pre-Shower (PS) detector, which is an electromagnetic shower detector segmented longitudinally, is used and located before the main section of the ECAL. In front of the PS, is a Scintillator Pad Detector (SPD) section and between them both is placed a thin lead converter. The SPD is used to select charged particles that allow the rejection of the background π^0's of high E_T, with the lead converter used to convert the particles to photons that can be distinguished later. To obtain the optimal resolution, the full EM showers from the high energy particles must be contained, which is why the optimal ECAL thickness was determined to be 25 radiation lengths. The HCAL resolution does not impose any such constraints on the containment of the hadronic shower, so its thickness is set to 5.6 interaction lengths.

The scintillation light produced in the electromagnetic and hadronic showers is transmitted through WaveLength-Shifting (WLS) fibres to photomultiplier tubes. The SPD/PS cells use MultiAnode PhotoMultiplier Tubes (MAPMT) wheres the ECAL and HCAL require individual PMT tubes. To keep the transverse energy, E_T, constant between both the ECAL and HCAL phototubes, the gain is varied to be proportional to the distance between them and the beampipe, which means the HCAL operates at a higher gain to the ECAL phototubes. The prototype PS was tested at CERN using a test beam of electrons and pions in the energy range 10 - 50 GeV/c momentum to study the e/π separation performance. The deposited energy in the PS is shown in Figure 2.17.

2.8.2 Electromagnetic CALorimeter (ECAL)

The electromagnetic calorimeter was built using shashlik technology [58] [59]. This technology was chosen not only due to its known reliability, but also due to its energy resolution, fast time response and resilience to radiation. In particular, the energy resolution gives a B mass resolution of $65\,\mathrm{MeV}/c^2$ for the $B \to K^*\gamma$ penguin decay, and a $75\,\mathrm{MeV}/c^2$ resolution for the $B \to \rho\pi$ decay, where the π^0 mass resolution is $\approx 8\,\mathrm{MeV}/c^2$.

Shashlik technology refers to alternating 'slices' of absorber and scintillator material, where in general the absorber material has a small radiation length (X_0) such that the particle radiates energy over a short distance. This energy is then picked up and converted into visible light by the scintillator material. The ECAL is located 12.5m downstream from the interac-

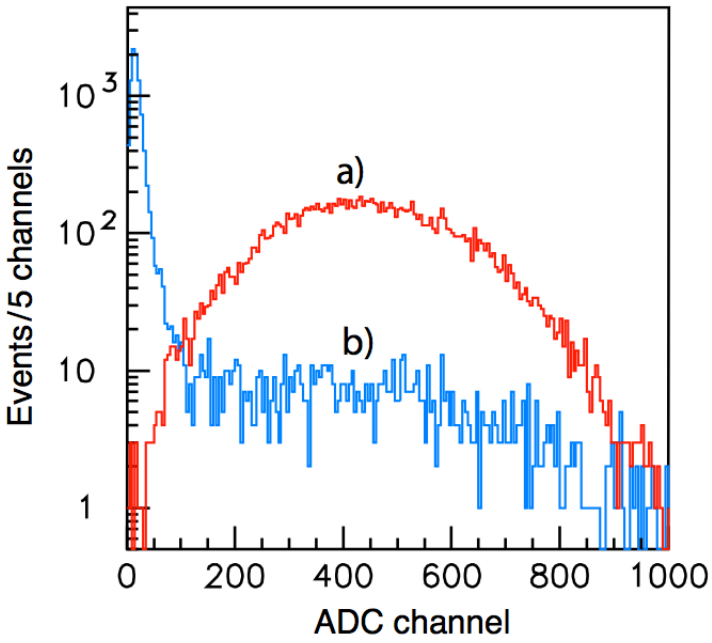


Figure 2.17: Energy Deposition of (a.) 50 GeV electrons in red and (b.) pions in blue at the PS [44].

tion point, with the outer dimensions following the projection of the tracking system components. These were discussed in Section 2.5, where $\theta_x < 300\,\mathrm{mrad}$ and $\theta_y < 250\,\mathrm{mrad}$. The hit density varies as a function of distance to the beampipe, because of this, the calorimeter is split into three sections the inner, middle and outer sections. Each of these module sections is composed of $2\,\mathrm{mm}$ thick lead followed by $120\,\mu\mathrm{m}$ thick reflecting TYVEK [60] paper and $4\,\mathrm{mm}$ thick scintillator tiles produced from polystyrene. They are then wrapped in black paper to keep them light tight. The scintillator light is focussed into the WLS fibres by the positioning of small holes in the scintillator tiles and lead plates. The light is then absorbed, re-emitted and transported down $1.6\,\mathrm{mm}$ diameter fibres, which are looped through the entire module twice so as to increase the light collection efficiency. The fibres are then bundled at the rear end of the module, with the ends being polished, and the collected light being read out by Hamamatsu R7899-20 phototubes. The number of photoelectrons differ for each phototube due to variations in the WLS fibre density within each module and also the cell sizes.

The performance of the ECAL is based largely on its light collection ability and energy resolution. Of all the scintillation light produced through the electromagnetic showers, only a few percent is actually registered by the phototubes after it is captured and re-emitted by the WLS fibres. This is due to a large number of factors including the dependence on the emission point of the light to the fibres, which itself depends on the mean flight path of the

light through the scintillation tile, and the imperfect reflection of the surface of the tile edges. The efficiency was modelled using a dedicated MC simulation that took into account the light absorption, and propagation through the materials, in the ECAL. The inner, middle and outer modules were scanned with a 100 GeV/c muon beam, this was performed in transverse steps of 1 mm. The results of this showed that the global non-uniformity of the response was measured to be negligible between the module types. The results of the full test beam with a lateral scan of 50 GeV/c electrons are displayed for the inner and outer modules in Figure 2.18. The module response is found to be uniform within 0.8%.

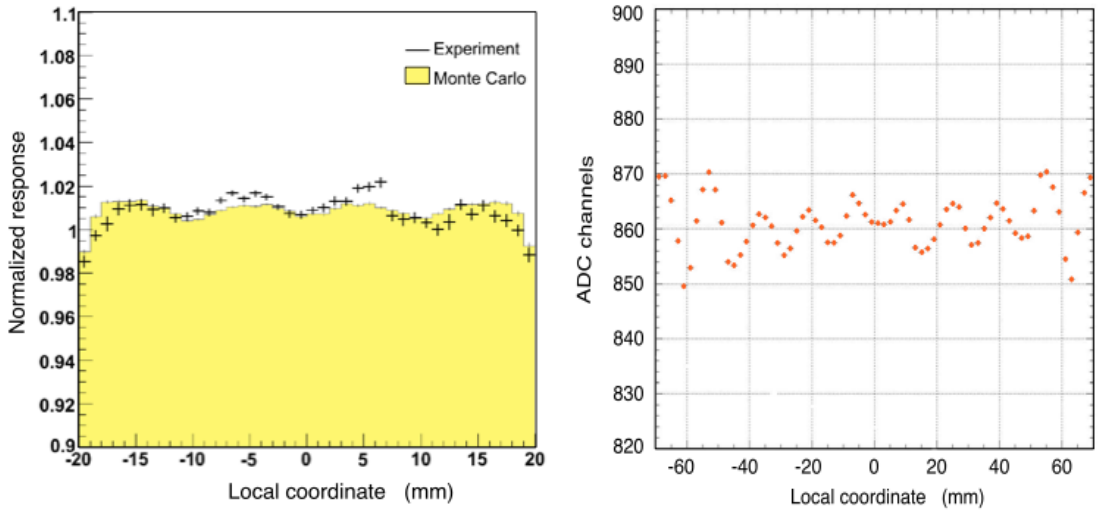


Figure 2.18: Uniformity of response to 50 GeV/c electrons of the inner (left) and outer (right) modules. The scan was made in 1mm wide slices through the fibre positions [44].

2.8.3 Hadron CALorimeter (HCAL)

The hadron calorimeter uses a similar principle to the ECAL, except this sampling device uses iron as the absorber material and scintillator tiles as the active material to convert the scintillation light to photons. The unique feature of the organisation of the detector arrangement is that the scintillating tiles run parallel to the direction of the beam. The absorber and scintillation materials are placed such that there is 1 cm of iron placed between each tile as the distance from the beam increases (Figure 2.19). Longitudinally the length of the tiles and iron spaces corresponds to the hadron interaction length λ_I in steel, approximately 17 cm. Similar to the ECAL, WLS fibres run longitudinally through the structure towards the back region where they are brought together into light guides which focus the detected scintillator light into photomultiplier (PMT) tubes. The general layout of the HCAL resembles a wall,

which is located at a distance of 13.33 m from the interaction point and has dimensions of 8.4 m in height, 6.8 m in width and 1.65 m in depth. To allow ease of access to the detector, it has been vertically divided into two symmetric parts which are positioned on movable platforms that can move away from the beam pipe for maintenance. The tile calorimeter consists of three main components: the scintillating tile, the WLS fibre and a small square light mixer located in front of the photomultiplier window. The light produced via scintillation propagates through the 3 mm thick tile to the edges where it is collected by the WLS fibres. Each WLS fibre collects light from 3 scintillator tiles arranged along the shower direction, with the light then propagating via total internal reflection to the light mixer and PMT. Light collected from tiles that are further away from the PMTs yield less light due to the attenuation and scattering within the fibre. To normalise this and create a level of light uniformity reaching the PMTs, the tile to fibre optical contact is progressively reduced based on the different tile layers and depth. As an example, the last tile layer closest to the PMT has its optical contact reduced by 22% compared with the first layer at the HCAL entry. The PMTs are connected to the ends of the WLS fibres by a $35\,\mathrm{mm}$ long light mixer of square shape. The PMTs, including the light mixers, are also shielded from the magnetic field of the dipole magnet as it could have an effect on the performance. The shielding consists of a $3\,\mathrm{mm}$ thick iron tube and MuMetal [61] foil.

The performance of the HCAL has been tested with the use of prototypes and test beams. The detector response when the angle of the incoming particle beam changes has been tested by rotating the detector around the beam axis through an angular change of 0° and 15° to be close to the average shower maximum. From a lateral scan using a pion beam of energy $10\,\mathrm{GeV}$ to $80\,\mathrm{GeV}$, the uniformity in response is measured to be within $\pm 3\%$. Further beam tests and comparisons with MC simulation have allowed the energy resolution to be determined, with the resolution fitted at several energies being

$$\sigma_E/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\% \text{ (E in GeV)}. \tag{2.6}$$

2.9 Muon System

The Muon identification system is fundamental to the LHCb, as many CP-sensitive decays have muons present in their final states, specifically the so-called golden channels, $B_d \to J/\psi(\mu^+\mu^-)K^0_S$ and $B^0_S \to J/\psi(\mu^+\mu^-)\phi$. Decays with muons in their final states play a large role in oscillation and CP asymmetry measurements, as the muons are used to tag the initial flavour state of B-mesons. They are also used in the study of rare decays, such as the $B^0_S \to \mu^+\mu^-$ which may reveal NP BSM and predict which extension models could be more

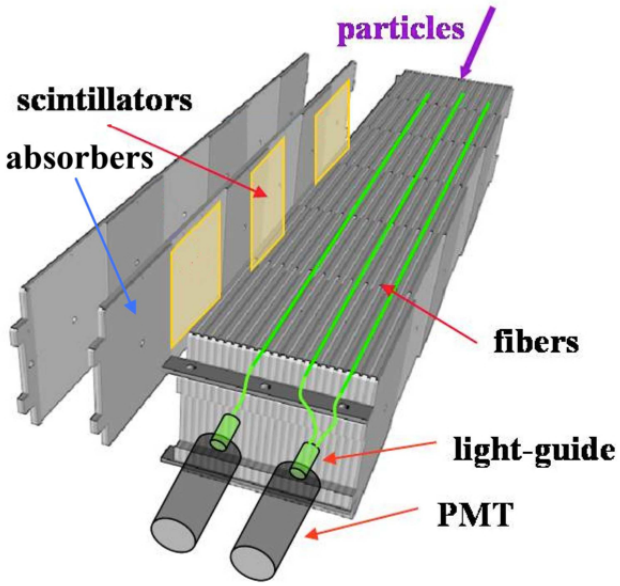


Figure 2.19: HCAL schematic [44].

reliable than others.

The muon system schematic is visualised in Figure 2.20, with the system seen to be composed of five separate stations. The first, M1, is located before the calorimeters with the remaining stations, M2-M5 being located downstream of the calorimeters. The M1 station is placed here to improve the p_T measurement of the trigger. To select penetrating muons, the M2-M5 stations have 80 cm thick iron absorber sheets placed between them, with the minimum momentum of muons allowed to traverse the stations being 6 GeV/c. The detectors themselves are partitioned into rectangular logical pads which define the x and y space point resolution of the tracks, this also provides binary information to the muon trigger and DAQ system. The muon trigger mentioned is based on an independent muon track reconstruction algorithm and a p_T measurement which requires aligned hits in all five muon stations. The track direction is calculated primarily from the information acquired by the M1-M3 stations, as these have a high resolution along the x-coordinate. These stations can also calculate the p_T of these particles with a resolution of 20%. The remaining stations, M4 and M5 do not have as good a resolution and so are used to identify the particles penetrating to their depth. As can be seen in Figure 2.20, each muon station is also separated into four regions R1-R4 of increasing distance from the beam pipe. The dimensions of these four regions, R1, R2, R3 and R4 and their segmentations scale in the ratio 1:2:4:8. Another aspect of this division into regions is that different detector technologies are required due to the need for radiation hardness closer to the beam, there are two main technologies employed: Multi-Wire Pro-

portional Chambers (MWPC) and Gas Electron Multiplier (GEM) detectors. These will be discussed in more detail in the following sections. The full system covers an area of $435\ \mathrm{m}^2$, with the inner and outer acceptances of the system being 20(167) mrad and 306(258) mrad respectively. As of the time of this thesis, the muon detection efficiency is above the design requirement of 99% [62] in all 5 muon stations. This is mainly determined by the chamber time resolution [63].

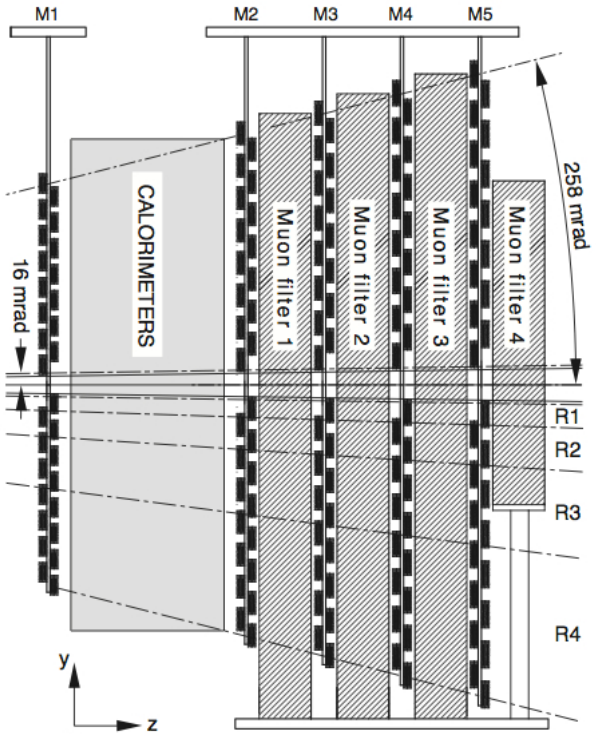


Figure 2.20: Muon chamber schematic [44].

2.9.1 Multi-Wire Proportional Chambers (MWPC)

There are 1368 MWPCs in the LHCb muon system, with studies showing that the time resolution that can be achieved is 5 ns. Each chamber consists of two cathode plates spaced at a distance of 5 mm apart and kept at a potential of 2.5-2.8 kV. Between the cathodes is a gas mixture containing 40% Ar, 55% CO_2 and 5% CF_4. Also in this 'gas gap' is contained gold-plated Tungsten wire of 30 µm diameter at a spacing of $\approx\ 2.0\ mm$. These collect and read out the accelerated electrons produced from the charged particle induced ionisation in the gas. A cross section view of the wire chambers is shown in Figure 2.21, with this representing the M2-M5 chambers which have four equal gas gaps stacked together. The M1 chambers consist of only two gas gaps stacked together.

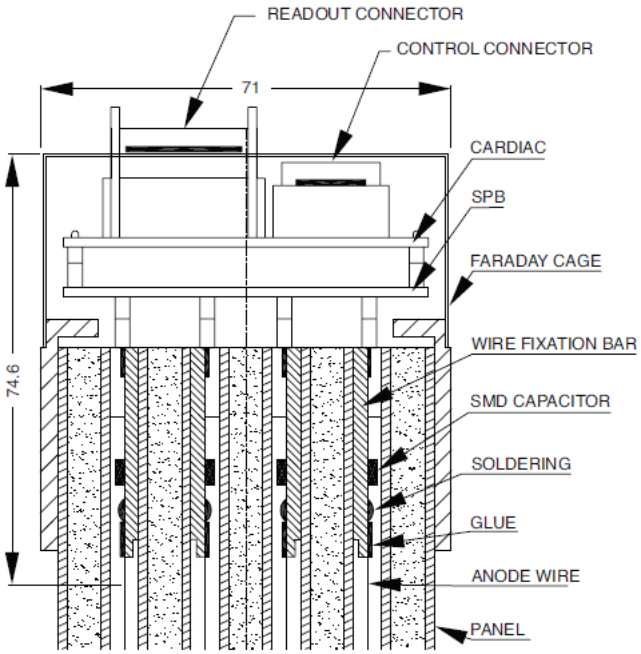


Figure 2.21: MWPC schematic [44].

2.9.2 GEM Detectors

The position within the subdetector that has the highest flux of generated particles is in the region R1 of muon module M1, seen on Figure 2.20. It is here that the harshest requirements are needed due to the high rate of charged particles crossing the detector, which can be up to ≈ 500 kHz/cm^2. This high rate means the technology used in this location must be particularly radiation hard and long lived in such conditions to survive up to 10 years without significant ageing effects. Using wire chambers in this region is difficult, so an alternative must be considered. The option chosen in this high intensity region was to use Gas Electron Multiplier (GEM) chambers. The chambers consist of a triple-GEM detector that is comprised of three GEM foils sandwiched between an anode and cathode plane pair with a gas mixture between them (see Figure 2.22).

The layout allows the detector to be used as a tracking detector with good timing and position resolution. The technology works by having ionisation electrons, produced by traversing charged particles, drift towards the cathode and the first GEM foil. The electrons are accelerated by the electric fields applied to each of the GEM foils, which then multiply them at each step. After the final GEM foil they drift towards the anode which induces a current on the pads. The gas composition of 45% Ar, 15% CO_2 and 40% CF_4 gives the detector a time resolution of better than $3\,\mathrm{ns}$ in comparison with the time resolution of $\approx 10\,\mathrm{ns}$ obtained from the standard gas mix of 70% Ar and 30% CO_2. The optimised gap field and related

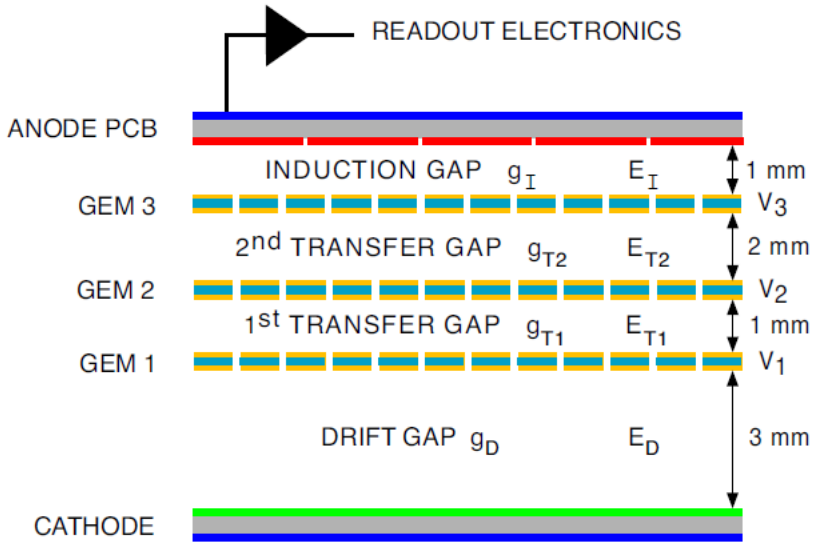


Figure 2.22: Triple GEM sensor schematic [44].

voltages across the GEM foils are: $\mathrm{E_D} = 3.5\ \mathrm{kV/cm}$, $\mathrm{E_T} = 3.5\ \mathrm{kV/cm}$ and $\mathrm{E_I} = 5\ \mathrm{kV/cm}$ and $\mathrm{V_1} = 440\ \mathrm{V}$, $\mathrm{V_2} = 430\ \mathrm{V}$, $\mathrm{V_3} = 410\ \mathrm{V}$. The completed GEM component of the muon chambers comprises of 12 chambers, each of which has two triple-GEM detectors superimposed to make two sensitive layers that are logically OR-ed. The individual chambers have a total active area of $20 \times 24\ \mathrm{cm}^2$.
The performance of the GEM chambers has been conducted using cosmic rays and using a dedicated 40 MHz test beam.

2.10 Triggers

The purpose of the LHCb trigger system is to reduce the amount of data collected by the experiment down to a manageable amount to write to storage. The current bunch structure of the LHCb, in combination with the luminosity, crossing frequency and visible cross section give an event rate of $10\ MHz$. The trigger deployed must be able to reduce this down to about $2\ kHz$, although this has been increased to $5\ kHz$ for 2012 data taking, which is the maximum rate at which data can be written to storage for further offline analysis. The reduction proceeds through two different levels: The hardware-based Level 0 (L0) trigger and the software Higher Level Trigger (HLT1 and HLT2). The flow of the trigger sequences is given in Figure 2.23. Overall the Trigger is optimised to achieve the highest signal efficiency for events selected for the offline analyses [64]. The specific L0, HLT1 and HLT2 triggers used for the lifetime analyses presented in this thesis are discussed in Section 5.3.1.

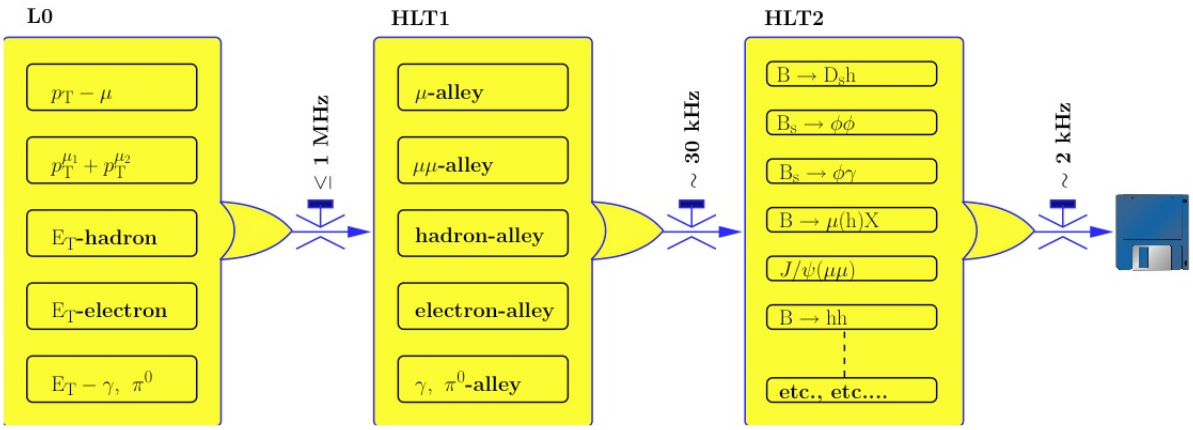


Figure 2.23: Flow of the Trigger sequences for 2011 data taking. This has since been increased to 5 kHz output for 2012 data taking [44].

2.10.1 Level-0 System

The L0 trigger system is designed to reduce the rate at which the detector reads out data down to 1 MHz, as the full event rate is too large to record. The large mass of the B meson means that daughter decay particles produced from collisions, often have large transverse momentum (p_T) and energy (E_T). Knowing this information, the L0 attempts to reconstruct the hadron, electron and photon clusters with the largest E_T and the two highest E_T muons in the muon chambers. These, coupled with further information from the VELO and calorimeters, allow a large number of events that would otherwise be triggered by large combinatorics to be rejected thus increasing the fraction of useful events flowing through the available bandwidth. The L0 trigger is composed of three discrete components: the L0 calorimeter trigger, the L0 muon trigger and the L0 pile up trigger. The calorimeter trigger system looks for high E_T particles such as electrons, γs, π^0 s and hadrons while the muon trigger selects the two muons with the highest p_T for each quadrant of the muon detector. The pile up trigger is not used to select flavour physics events, but instead aids the determination of the luminosity. The L0 decision unit can then perform a simple logic on the signatures from each component to come to a pass/fail decision in ≈ 2.5 µs. The full flow of the trigger sequences can be seen in Figure 2.23

2.10.2 HLT System

The HLT system is implemented to reduce the data flow rate from the 1 MHz coming through the L0 trigger, down to the more manageable 2 kHz. To do this, the HLT utilises the full event data in order to pre-select candidates and filter them down pathways to specific data sets of

interest. These paths are determined by previous L0 decisions, with the requirement that candidates are reconstructed in the tracking stations and that they have high p_T and/or large impact parameters. These criteria are able to reduce the data flow down to $30\ \mathrm{kHz}$ during the first stage. Selection cuts are then imposed in the second stage to reduce this rate from $30\ \mathrm{kHz}$ to $2\ \mathrm{kHz}$. The HLT itself is subdivided into two parts, HLT1 and HLT2, which will now be discussed.

2.10.3 HLT 1

The purpose of the HLT1 Trigger is to reconstruct particles from the tracking station information that relate to L0 events. The HLT1 consists of 'alleys', where each alley is related to different trigger types in the L0. Of all the events that pass the L0 conditions, $\approx 15\%$ of events will pass more than one trigger of the HLT1. To confirm the events that pass these alleys a set of algorithms are used, with each HLT1 alley using a sequence of these alleys to reduce the rate. The algorithms will not be discussed here as they are beyond the scope of this thesis. An example of one of the alleys used is the hadron-alley. The hadron-alley running at a luminosity of $2 \times 10^{32}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$, with an input rate of $\approx 600\ \mathrm{kHz}$, contains $\approx 1.3\ \mathrm{L0}^{\mathrm{hadron}}$ objects per event. Applying a subset of the available algorithms incrementally reduces the input rate of the alley until it its reduced to a rate of $11\ \mathrm{kHz}$, with ≈ 3.2 candidate secondary vertices per event. Other HLT1 alleys use analogous methods to this.

2.10.4 HLT 2

The secondary HLT trigger, HLT2, is designed to perform a full reconstruction on the events which have been accepted by the HLT1 alleys. This is possible due to the L0 and HLT1 event rate reduction. The HLT2 step uses cuts on both kinematic and geometrical variables, with the resultant selections being both exclusive and inclusive. Selections that are 'exclusive' require the mother particle to be fully reconstructed *e.g.* $\mathrm{B}^0_d \rightarrow \mathrm{K}^+\pi^-$, whereas 'inclusive' selection only require the mother to be partially reconstructed *e.g.* $\mathrm{B} \rightarrow \mathrm{h}^+\mathrm{h}^-$. The aim, with either selection type, is to reduce the data rate to a much more manageable $2\ \mathrm{kHz}$ which is the rate at which the data is written to storage for offline analysis. Each selection in HLT2 is taken in a logical OR to decide whether to keep or discard an event, thus reducing the event retention rate. Events passing HLT2 are sent to permanent storage, to be fully reconstructed offline and used in physics analyses.

2.11 Offline Reconstruction and Data Processing

In preparation for analysis, the raw data collected by the experiment must be processed into a useable format. To accomplish this, LHCb employs two key pieces of software, these are BRUNEL [65] and DAVINCI [66].

The event reconstruction is performed offline by the BRUNEL package, which takes into account information from all the subdetectors. The tracks and primary vertices of an event are reconstructed using the raw hits and clusters registered by the VELO, with these tracks then extrapolated upstream to be associated with the Cherenkov rings detected by the RICH. Information from the calorimeters and muon stations is then included to determine a likelihood for each track being either a π, K, e, μ or p. The output files from the BRUNEL reconstruction contain all the information regarding the fully reconstructed PVs, tracks and their PID likelihoods. These are then saved to permanent storage.

Once the reconstruction of events is completed, the data can be searched for particles of interest, such as B_s^0 mesons, using the DAVINCI software package. The DAVINCI software is able to combine sets of tracks under the hypothesis that they were made by the daughters of a single mother particle. This uses the mother particle's momentum, and trajectory, which is defined as the sum of those of the daughters. Immediately after the BRUNEL reconstruction is completed, DAVINCI fully reconstructs offline events and runs a set of algorithms that are designed specifically to select different decay channels. These criteria assist in confirming the trigger decisions made previously, Section 2.10, while also effectively removing background from the channels of interest. This process is know as 'stripping', with many stripping lines used at LHCb for different analyses. The specific $\mathrm{H_b} \to \mathrm{h^+h^-}$ stripping line used for the lifetime measurement in this thesis, is discussed in Section 5.3.2. Additionally, the software can be rerun over these stripped events, applying different selection criteria, to improve the signal purity of the data sample for individual analyses. The end user is able to select the parameters of interest for their analysis, such as the mass and lifetime of a candidate, which is demonstrated in Section 5.3.3.

2.12 Summary

This chapter has provided an overview of the LHCb experiment, one of the four primary experiments at the LHC complex, CERN. A short overview of the LHC accelerator and the main experiments it hosts is provided, before detailing the one most relevant to this thesis. The LHCb experiment was designed to make high precision measurements in heavy flavour physics, this requires many specialist subdetectors to contribute to this aim. A detailed de-

scription of the major subdetectors is given, in the order of distance from the interaction point. The detail of their purpose, design and, where possible, their performance is provided. The performances given indicate the excellent performance of the detector as a whole. This allows high precision data analysis, discussed further in this thesis, to be performed. Another key component, the trigger, which is a combination of both hardware and software that reduces the stream of raw data down to a manageable amount, is discussed. Finally, the reconstruction and offline processing software that provides the link between the collection of raw data, and the extraction of useful signal events, is presented.

Chapter 3

RICH 2 Mirror Alignment Stability and Gas Refractive Index Monitoring

3.1 Introduction

The daily operation of the RICH subdetectors requires constant monitoring of many variables within the detector environment. Many of the monitored environmental variables allow re-calibrations to be automated, which allow for continuous precise measurements of important values such as the refractive index, n. Others that are monitored are related to the infrastructure and hardware of the RICH detectors, and so directly and indirectly affect performance. This chapter provides details of two studies used to monitor the performance of the RICH subdetectors. Section 3.2 presents a study of the Laser Alignment Monitoring System (LAMS) of RICH 2 that provides an independent monitor of the mechanical stability of the mirrors. The study will describe the system setup, method of calibration for the mirrors, analysis of possible sources of mirror movement and calculations to verify movement origin hypotheses. The second study, Section 3.3, will detail an alternative method to determine the refractive index used in the RICH 2 reconstruction. This method also allows a determination of the Cherenkov angle θ_c directly from the refractive index, and the molecular mass $< M >$ of the gas used as a medium in both RICH detectors. The unique element to this method is that it provides continuous monitoring of these values, with a greater frequency in time than is currently performed.

3.2 Laser Alignment Studies

The RICH mirror alignment monitoring system uses a set of mirror-camera pairs in order to measure the respective mirror movements within the RICH subdetector. The RICH design is

described in detail in Section 2.7, however the beam splitter and monitoring camera locations in RICH 2 are provided in Figure 3.1.

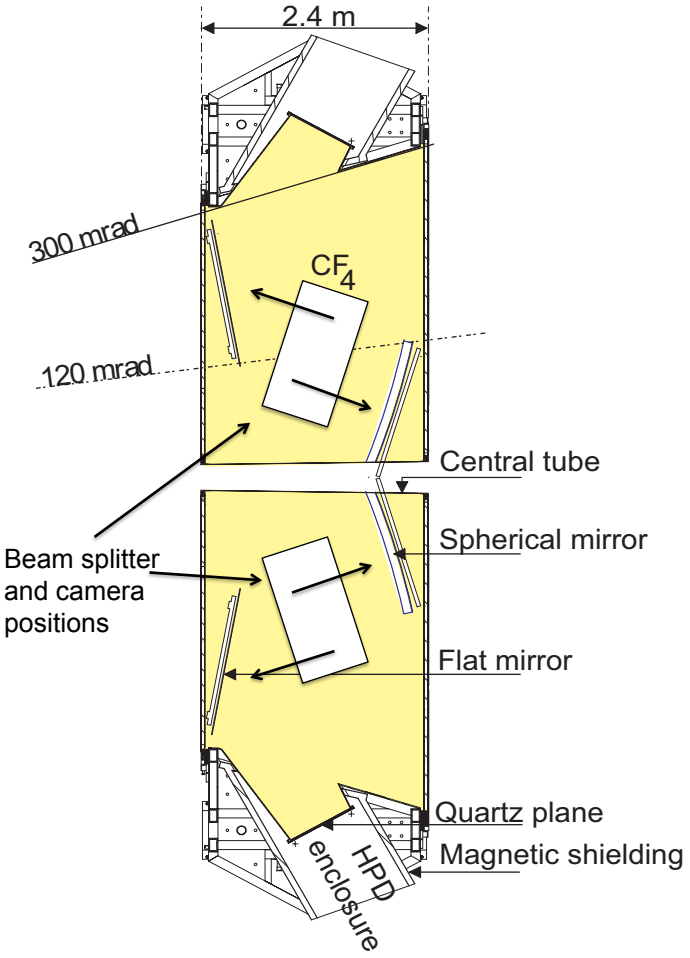


Figure 3.1: Top down RICH 2 schematic displaying the relative positioning of the beam splitters and CMOS monitoring cameras. The beam splitters are mounted on the floor, with the cameras mounted on the ceiling.

The central positions for the beam splitters and their corresponding monitoring cameras, are provided in Appendix A.

The mechanical stability of the flat and spherical mirrors (displayed in Figure 3.1) is of high importance as this directly affects the Cherenkov angle resolution that is able to be obtained. Due to the size of the mirrors, which have radii of approximately 0.5 m, securing them to the supports without causing mechanical deformations is a difficult procedure. The mirror supports are made of an Aluminium honeycomb structure $40\,\mathrm{mm}$ thick (see Figure 3.2). This material and structure was chosen because it is non-magnetic, so will not be affected by the strong magnetic field, and it fulfils the requirements to be mechanically rigid and lightweight [67].

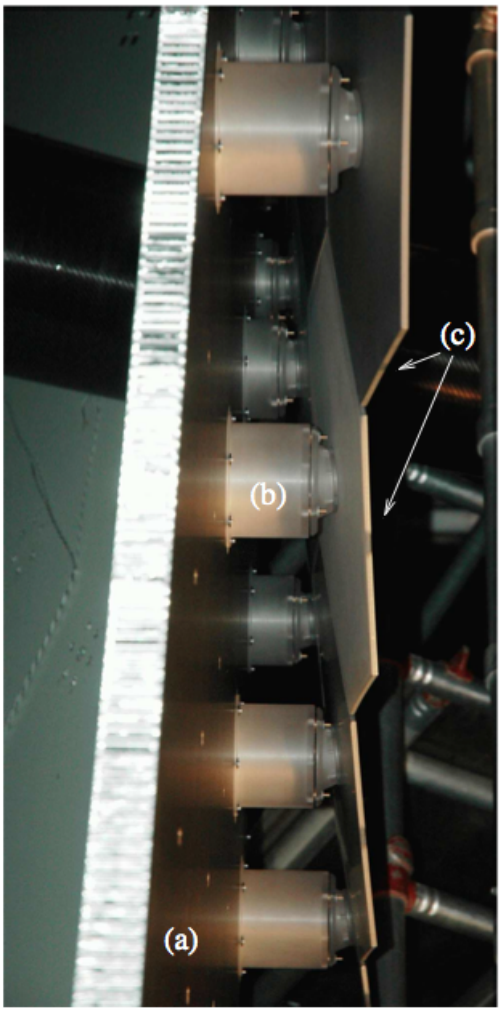


Figure 3.2: The support structure for RICH mirrors: (a) the aluminium honeycomb panel, (b) the polycarbonate supports, (c) the mirrors [68].

To monitor potential movement of the mirrors, an alignment system has been installed in both RICH 1 and RICH 2. The organisation of the flat and spherical mirrors of RICH 2 on these supports are given in Figures 3.3 and 3.4

The alignment system uses a pixellated Complimentary Metal-Oxide Semiconductor (CMOS) camera, which has pixel pitch of 6 μm, matched up to a mirror partner. A laser light source is then fed through a bunch of optical fibre cables to a beam splitter matched to each mirror-camera pair, eight in RICH 1 and sixteen in RICH 2. The laser light is then split by each beam splitter into two beams, the first beam travels directly from the laser source to the camera and is used as a reference beam, with the second spot directed via the partner mirror onto the camera (see Figure 3.5). Each laser beam and CMOS detector combination is associated to one mirror only (eight flat mirrors and eight spherical mirror segments in RICH 2).

Two beam spots are focussed on the focal plane of the camera, Figure 3.6, which are sepa-

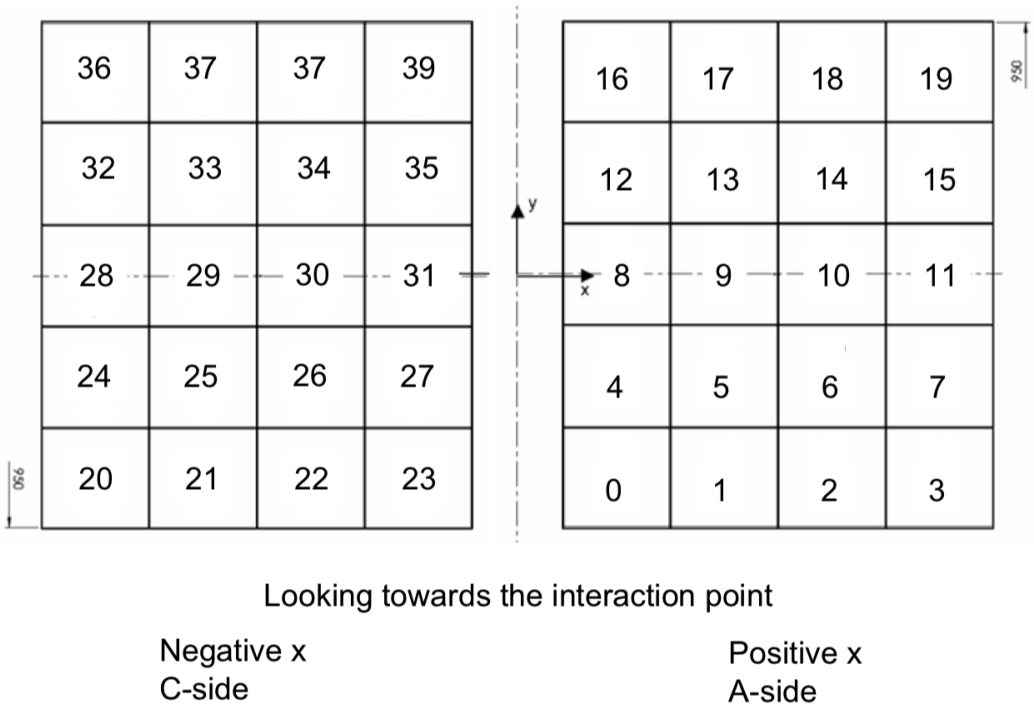


Figure 3.3: Flat mirror map schematic for RICH 2, including the mirror labelling.

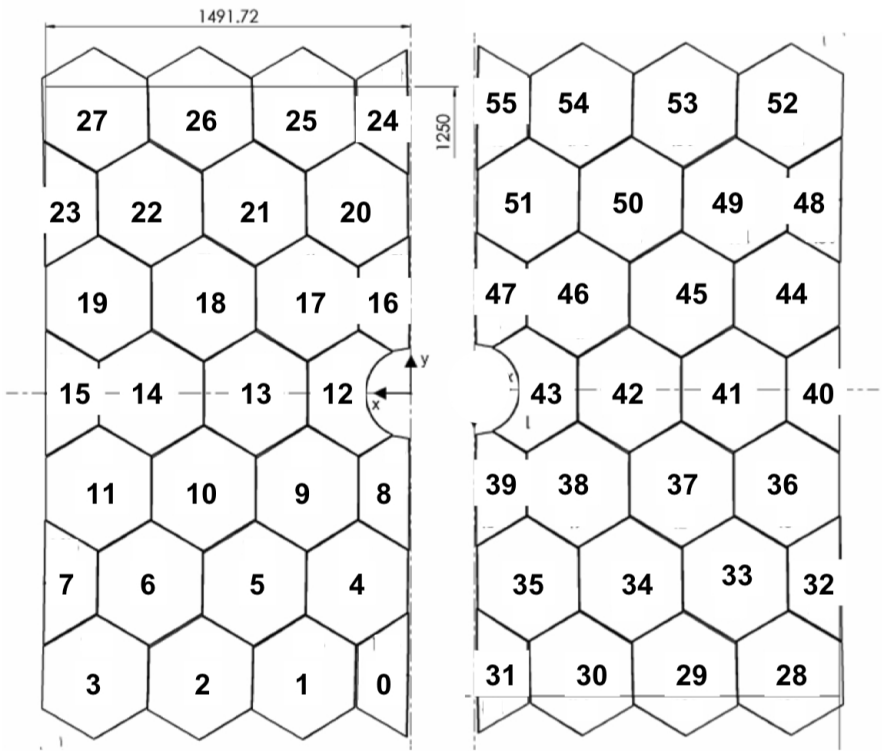


Figure 3.4: Spherical mirror map schematic for RICH 2, including the mirror labelling.

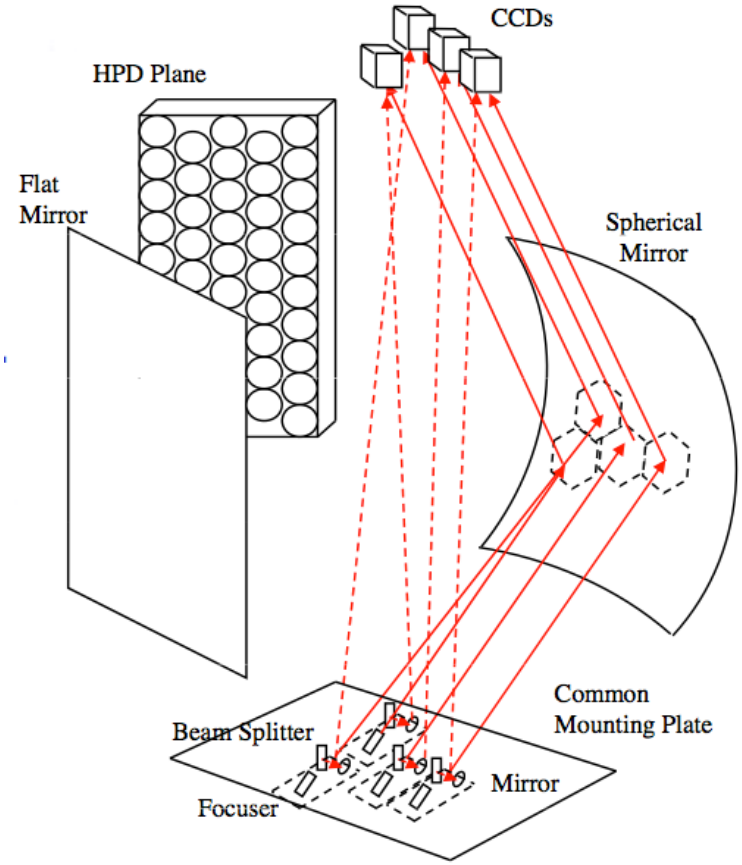


Figure 3.5: Schematic of the mirror alignment system for RICH 2 [69].

rated from each other by a small distance. Any change in the relative position of these two beam spots would indicate a mirror mechanical misalignment. To accurately determine the central position of the beam spots, a 2D gaussian fit algorithm is employed. The fit, detailed in [70], uses an established method of fitting to a beam spot that has a Gaussian profile, which is then performed for both the x and y dimensional variates.

There is a linear transformation relationship between changes in the mirror tilt $(\Delta\theta_x, \Delta\theta_y)$ and the observed relative movement of the beam spot positions $(\Delta x, \Delta y)$ on the CMOS camera [44], under the assumption of small mirror tilts $(\Delta\theta_x, \Delta\theta_y << 1)$:

$$\Delta\theta_x = A\Delta x + B\Delta y \tag{3.1}$$

$$\Delta\theta_y = C\Delta x + D\Delta y \tag{3.2}$$

The A, B, C and D parameters are fixed constants for each mirror-camera setup, which are dependent on the geometry of the system. The constants and the method used to calculate them will be discussed in Section 3.2.1. Currently the system is able to track the movement of these spots to within an accuracy of $0.01\,\mathrm{mrad}$.

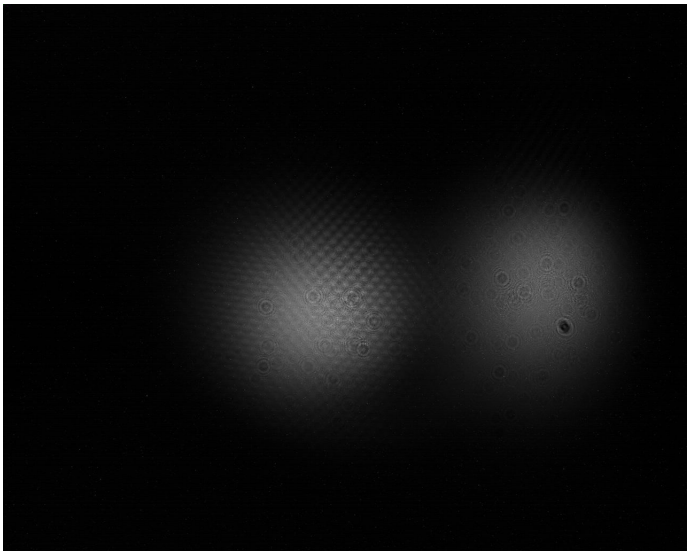


Figure 3.6: Reflected and reference beam spots from camera 4, collected from data in the period June 2012.

3.2.1 Mirror Coefficient Calibration

The calibration of the mirrors is performed through a variety of steps that are detailed here. The basis for this is a ray tracing code developed for use in the LHCb software environment, which uses RICH objects from the LHCb computing database to reconstruct the correct position of the mirrors. The beam splitter and camera positions, along with their respective vectors to the mirrors related to them, are used as input to calculate the coefficients in Equations 3.1 and 3.2. Appendix A includes a summary of all the beam splitters, camera positions and their respective vectors. There are two coordinate systems to account for at this stage; the global LHCb coordinate system and the local coordinate system of the RICH mirrors (Figure 3.7).

The beam splitter positions and vectors are all calculated in the LHCb global coordinate system (Figure 3.7, top), as are the central camera position and camera vectors. The positions of these are given in Figure 3.8 with the relative positions in RICH 2 given in Figure 3.1. The corresponding coordinates and ray-trace vectors are detailed in Appendix A.

The vectors from each beam splitter to a partner mirror were all trialled to ensure they traverse a position within the partner mirrors location. The reflected beam is then optimised to intersect the camera plane close to the (0,0) position in the (x, y) coordinates of the camera. The rotations of the RICH mirrors are performed about the axis in their local coordinate system, given in Figure 3.7 (bottom). These rotations are performed within the limits $-5 \rightarrow +5\,\mathrm{mrad}$ in steps of $1\,\mathrm{mrad}$. The qualitative results of each rotation are measured by observing the movement of the simulated beam on the defined camera plane. The observed move-

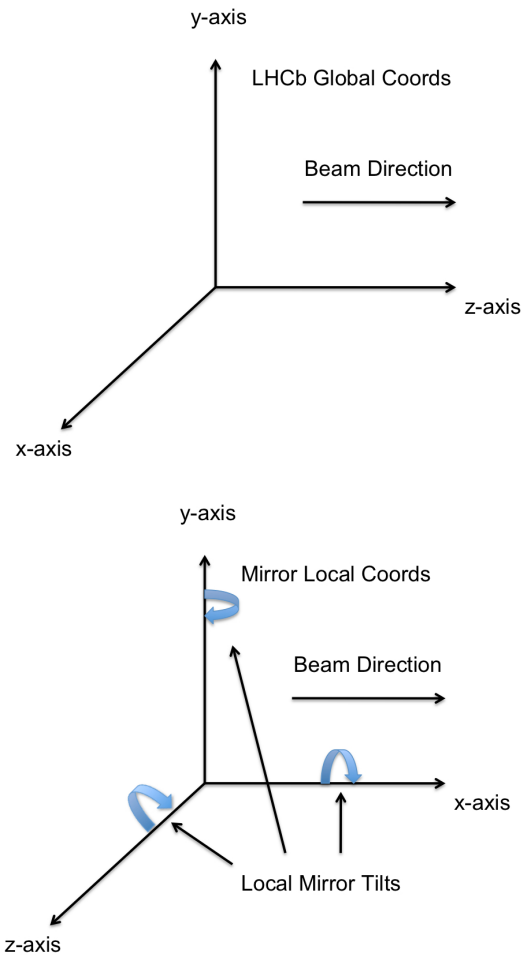


Figure 3.7: LHCb global coordinate system (top) and RICH mirror local coordinate system (bottom). The axes corresponding to the mirror tilt rotations used in the ray tracing code are also shown on the mirror local coordinate system diagram (bottom).

ments are heavily dependent on the mirrors, ray-trace vectors and rotations applied, and will vary based on the mirror studied. The resultant distributions for the spherical mirrors are similar, with the gradients of the distributions representing the A, B, C and D parameters for each mirror.

Figure 3.9 displays the expected variations on each coordinate on the defined camera plane after mirror tilts around the Local Y-axis (top plots) and Local Z-axis (bottom plots). Once these gradients are fitted based on the Δx and Δy values, the camera plane coordinate system must then be transformed to the corresponding CMOS camera coordinates. This is simply a case of transforming the camera plane z-axis to the CMOS y-axis and keeping the camera plane x-axis aligned with the CMOS x-axis, see Figure 3.10. The gradients that correspond to the A, B, C and D parameters in Equations 3.1 and 3.2, are shown for one mirror (mirror

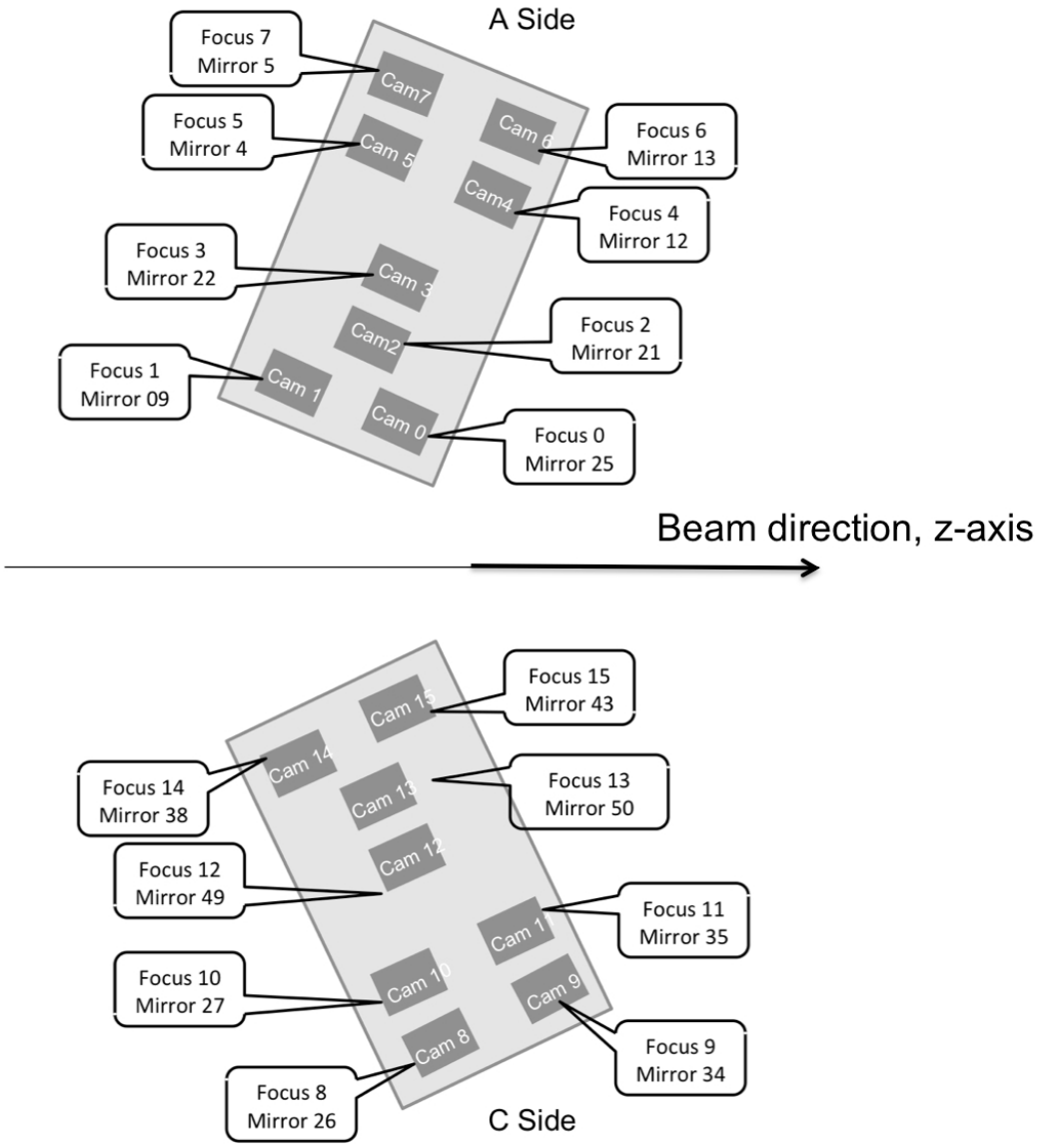


Figure 3.8: Beam splitter and focusing unit map.

25, camera 00) in Figure 3.9.

The resultant fitted parameters for each of the mirrors used are tabulated in Tables 3.1 and 3.2.

3.2.2 Data Calibration

Once the calibration parameters have been determined, they can be applied to the data taken by the LAMS. Unfortunately, due to the strength of the magnetic field and particle fluence during the LHCb runs, the RICH 1 cameras have all ceased to function. A preliminary analysis based on the RICH 1 camera is found in [71]. Further discussions in this section

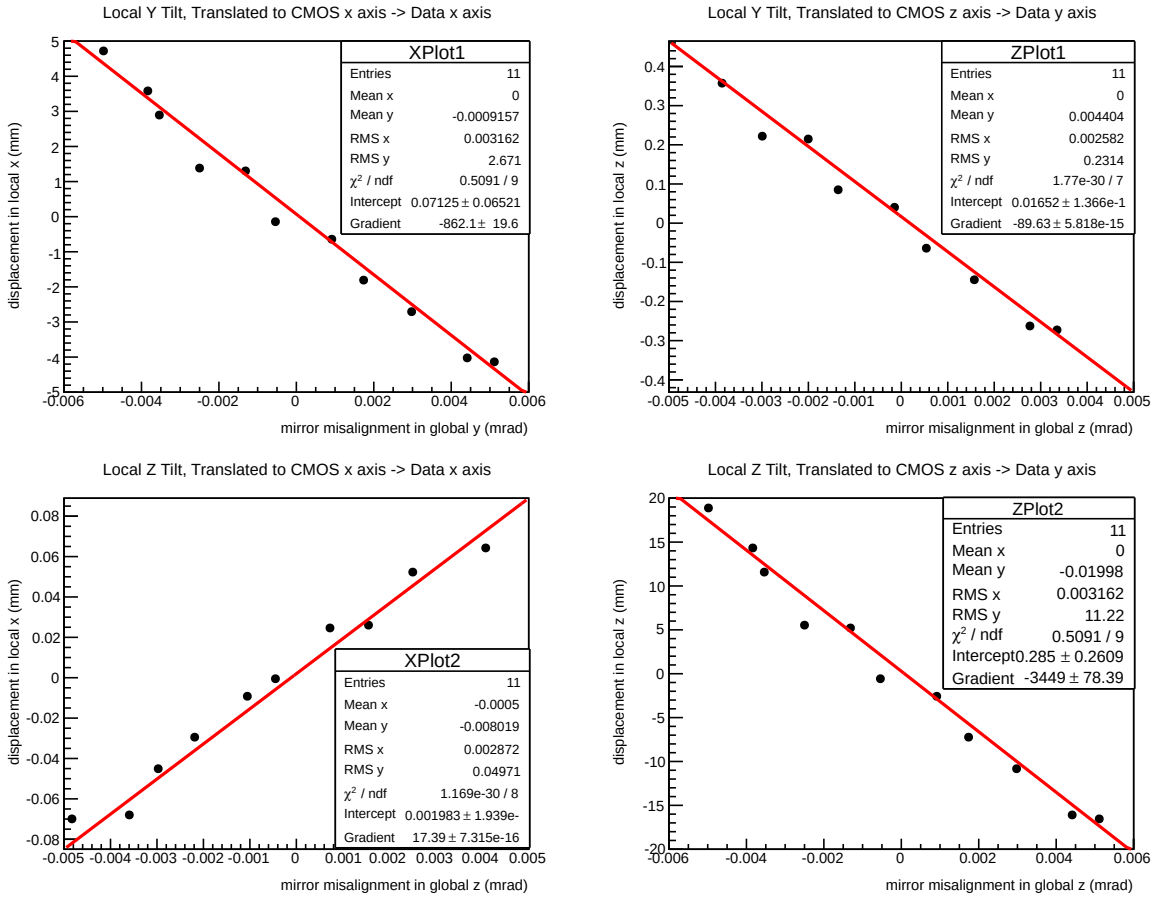


Figure 3.9: Calculated spherical mirror movements for Mirror 25-Camera 00 pair for the key coordinates based on rotations in the mirror local y and z-axis.

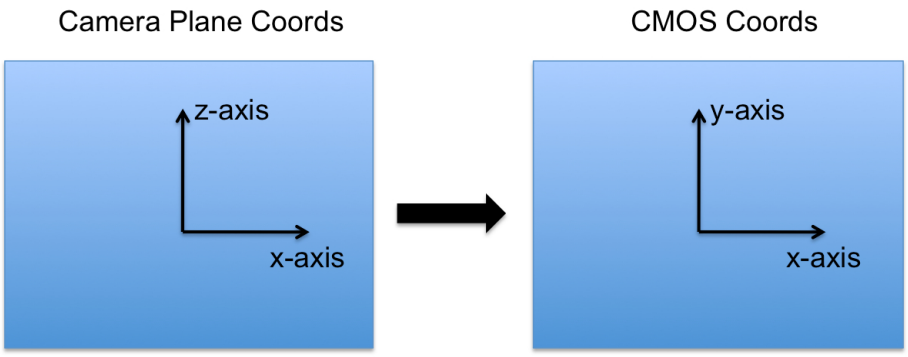


Figure 3.10: Transforming from the defined Camera Plane to CMOS camera coordinates.

will detail only the RICH 2 cameras. Due to the same circumstances as mentioned for the RICH 1 cameras, a subset of the RICH 2 cameras are unavailable, which are detailed in Table 3.3 and 3.4.

Mirror	A ($\mu rad/\,\mathrm{mm}$)	B ($\mu rad/\,\mathrm{mm}$)	C ($\mu rad/\,\mathrm{mm}$)	D ($\mu rad/\,\mathrm{mm}$)
25	-1.16×10^{-3}	3.01×10^{-5}	-5.85×10^{-6}	-2.90×10^{-4}
09	-3.87×10^{-4}	-6.54×10^{-6}	2.42×10^{-6}	-1.45×10^{-4}
21	-5.80×10^{-4}	1.26×10^{-5}	-3.37×10^{-6}	-2.30×10^{-4}
22	-5.80×10^{-4}	8.51×10^{-6}	-2.58×10^{-6}	-2.30×10^{-4}
49	-5.80×10^{-4}	-8.51×10^{-6}	2.32×10^{-6}	-2.31×10^{-4}
50	-5.80×10^{-4}	-1.18×10^{-5}	3.46×10^{-6}	-2.31×10^{-4}
38	-3.86×10^{-4}	7.00×10^{-6}	-2.31×10^{-6}	-1.50×10^{-4}
43	-5.80×10^{-4}	1.63×10^{-5}	-2.53×10^{-7}	-1.45×10^{-4}

Table 3.1: RICH 2 spherical mirror calibration constants.

Mirror	A ($\mu rad/\,\mathrm{mm}$)	B ($\mu rad/\,\mathrm{mm}$)	C ($\mu rad/\,\mathrm{mm}$)	D ($\mu rad/\,\mathrm{mm}$)
12	4.53×10^{-4}	0.0	-1.52×10^{-7}	1.73×10^{-4}
04	6.13×10^{-4}	0.0	-1.53×10^{-7}	1.46×10^{-4}
13	4.06×10^{-4}	0.0	-1.46×10^{-7}	1.70×10^{-4}
05	5.66×10^{-4}	0.0	-1.51×10^{-7}	1.46×10^{-4}
26	5.59×10^{-4}	0.0	1.22×10^{-8}	1.45×10^{-4}
34	4.06×10^{-4}	0.0	3.59×10^{-8}	1.70×10^{-4}
27	6.14×10^{-4}	0.0	3.59×10^{-9}	1.46×10^{-4}
35	4.58×10^{-4}	0.0	3.10×10^{-8}	1.74×10^{-4}

Table 3.2: RICH 2 flat mirror calibration constants.

To demonstrate the calibration of the data, one of the working cameras with a large period of uninterrupted data taking is used. Camera 00 on the A-side of the RICH system, which monitors the stability of spherical mirror 25 is used. The full 2011 and 2012 monitoring data for each available camera is downloaded from the RICH database for analysis. This is downloaded in the form of the reference and reflected beam spot data for both the x and y coordinates on the camera. The movements for each position coordinate can then be recorded as a function of time.

During the calendar year there are periods where the cameras stop taking data for technical reasons, this causes periods where there is no observed movement, which can be seen in Figures 3.11 and 3.12 for the reference and reflected beam spots respectively. To remove this from the analysis it is useful to only select periods where there is a constant stream of data taking, this can easily be done by selecting periods pertaining to specific months.

RICH A-Side	
Camera 00	Available
Camera 01	N/A
Camera 02	Available
Camera 03	Available
Camera 04	Available
Camera 05	Available
Camera 06	Available
Camera 07	Available

Table 3.3: RICH 2 A-side currently functioning cameras.

RICH C-Side	
Camera 08	Available
Camera 09	Available
Camera 10	N/A
Camera 11	N/A
Camera 12	N/A
Camera 13	N/A
Camera 14	Available
Camera 15	Available

Table 3.4: RICH 2 C-side currently functioning cameras.

To observe the absolute movement of the mirrors it is necessary to remove effects that are independent of the mirrors. By subtracting the movement observed on the camera from that of the reference beam spot, it is possible to leave only the mirror movement. Trivially, this is done by performing the subtractions:

$$\Delta\theta_x = \Delta\theta_{x\ reflected} - \Delta\theta_{x\ reference}. \tag{3.3}$$

$$\Delta\theta_y = \Delta\theta_{y\ reflected} - \Delta\theta_{y\ reference}. \tag{3.4}$$

As an example, the data distributions after the subtractions of Equations 3.3 and 3.4 have been applied are shown in Figure 3.13.

The movement on camera 25 in Figure 3.13, which is of the order of $10 - 40\ \mu\mathrm{rad}$, is of similar order to that shown by the other available spherical mirror cameras (see Table 3.5). The discussion on the source of this movement is detailed further.

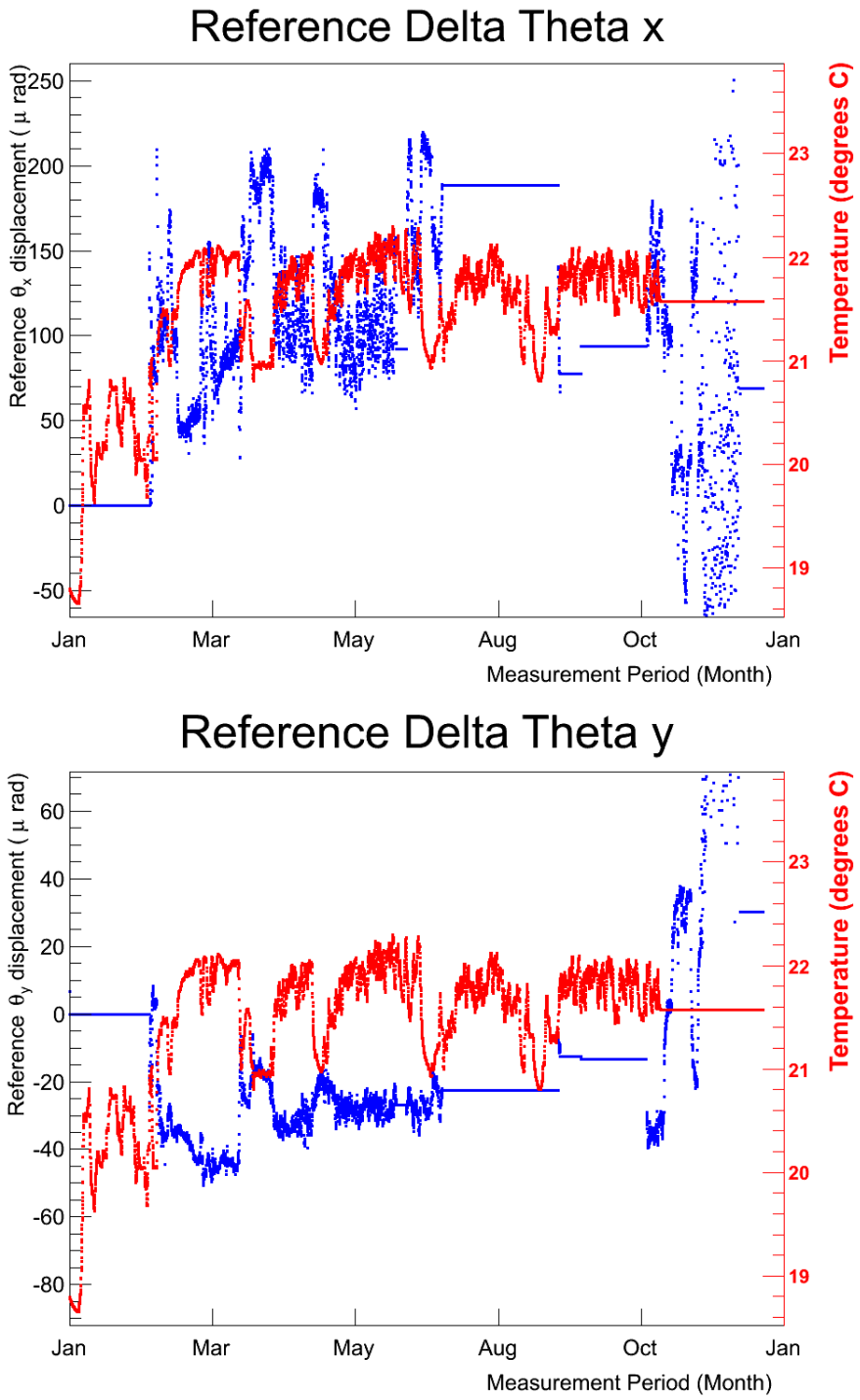


Figure 3.11: Reference x (top) and y (bottom) beam spot coordinate data from Camera 00 off Mirror 25 for the 2011 data taking period.

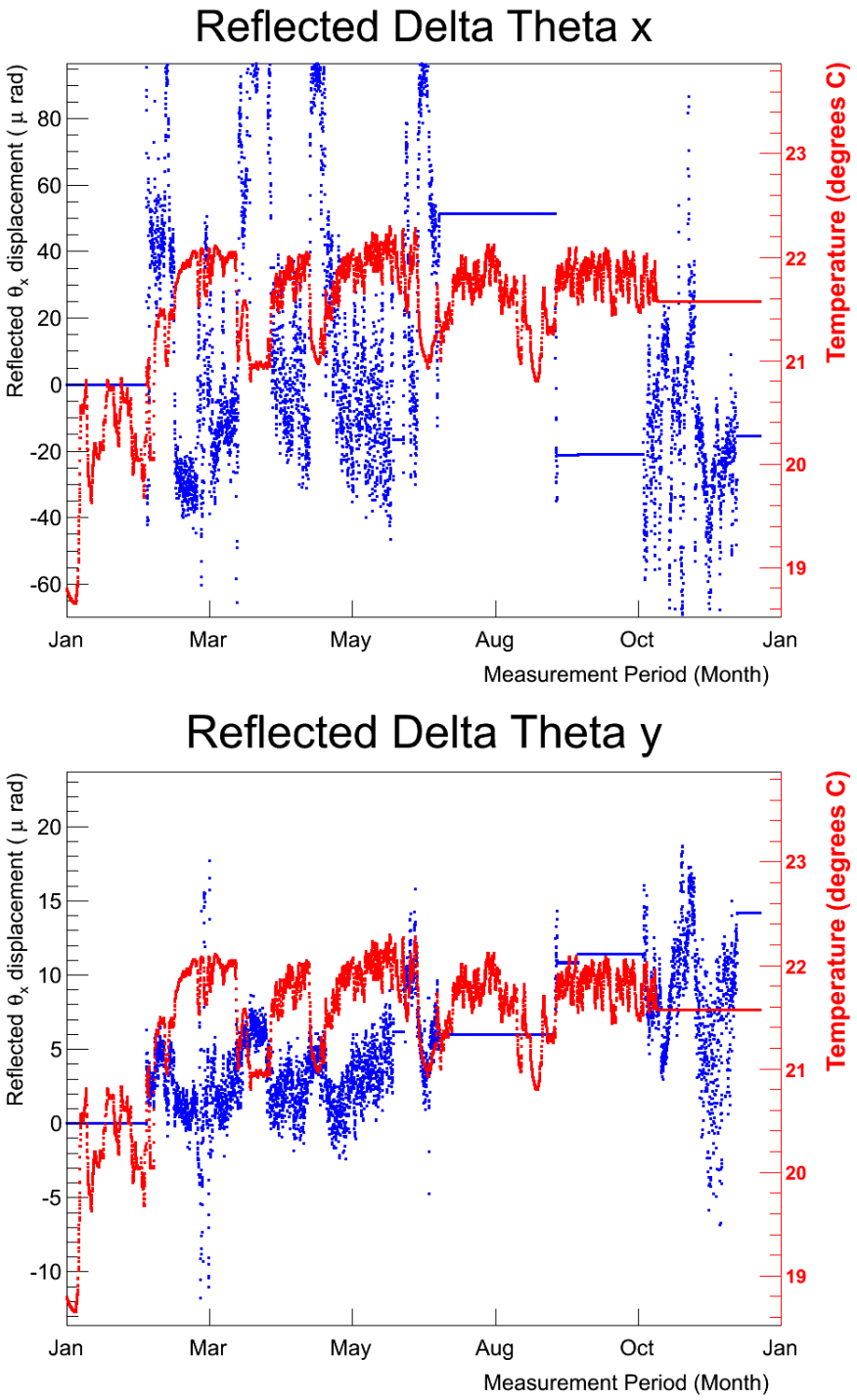


Figure 3.12: Reflected x (top) and y (bottom) beam spot coordinate data from Camera 00 off Mirror 25 for the 2011 data taking period.

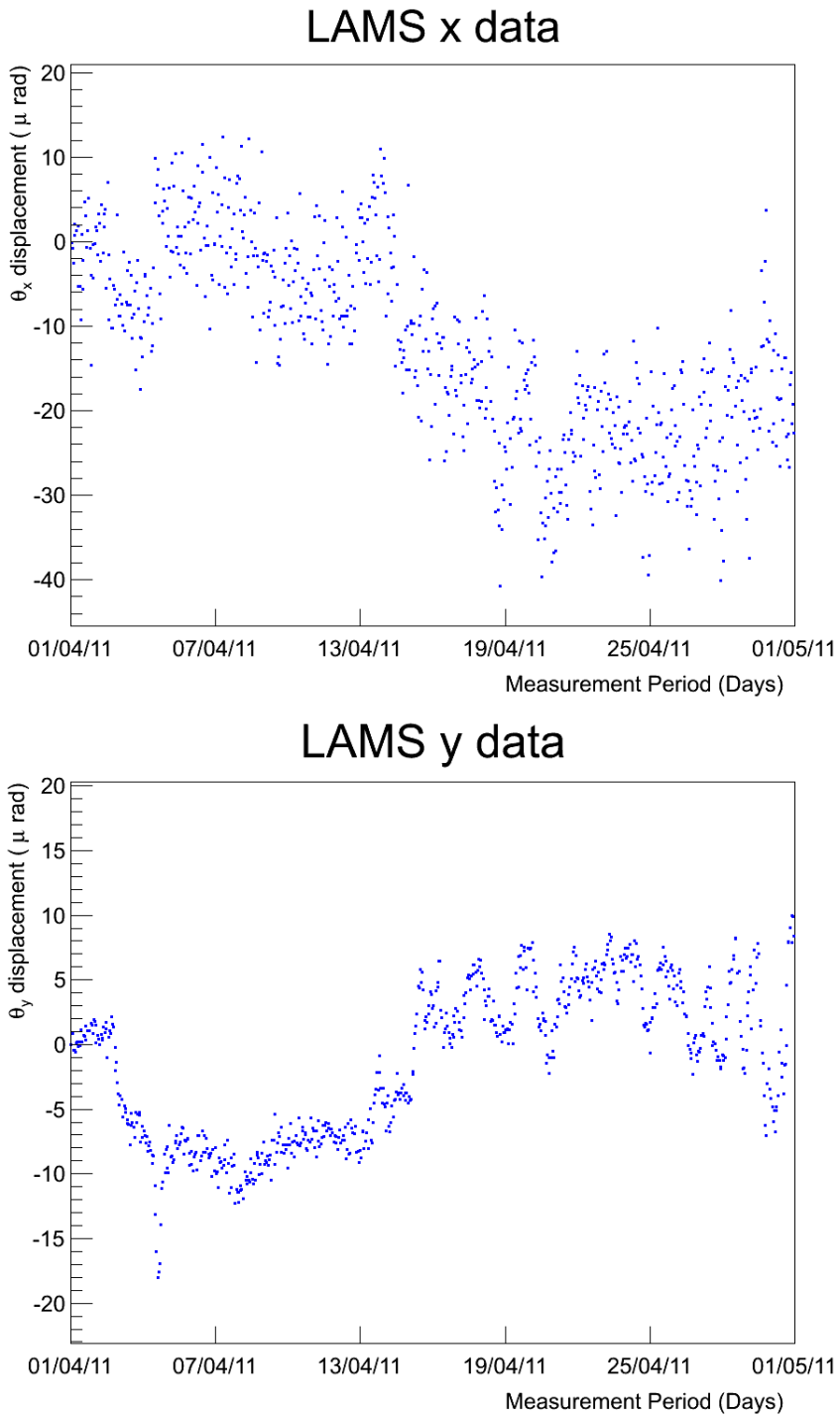


Figure 3.13: $\Delta\theta_x$ (top) and $\Delta\theta_y$ (bottom) movement of mirror 25 from April 2011 data.

Side	Mirror Type	Mirror	Fitted $\Delta\theta_x$ deviation (μrad/K)	Fitted $\Delta\theta_y$ deviation (μrad/K)
A-Side	Spherical	25	46.1 ± 1.1	19.6 ± 0.3
		09	-	-
		21	-	-
		22	53.8 ± 2.0	44.2 ± 1.0
	Flat	12	17.6 ± 0.3	65.8 ± 1.3
		04	-	-
		13	17.3 ± 0.1	44.8 ± 0.6
		05	-	-
C-Side	Spherical	26	49.5 ± 1.5	71.4 ± 2.0
		34	-	-
		27	-	-
		35	-	-
	Flat	49	-	-
		50	-	-
		38	52.1 ± 1.6	29.1 ± 0.7
		43	40.5 ± 1.6	11.7 ± 0.1

Table 3.5: Fitted relationships between $\Delta\theta_x$ and $\Delta\theta_y$ vs the change in temperature, ΔT, for each mirror. Only statistical errors from the fit are provided. Mirrors with no values attributed did not have sufficient data, or stable periods, to perform a fit.

3.2.3 Sources of Mirror Movement

The observed movement of the reference and reflected beam spots can be converted to an absolute movement of the mirrors, as was discussed in section 3.2.2. Therefore, the true mirror movement may be calculated as a mirror rotation in μrad around the y-axis to obtain the $\Delta\theta_x$ and the x-axis to obtain $\Delta\theta_y$. To account for this movement three external factors have been considered, these were

- pressure effects,
- effect of magnet polarity shift,
- the effect of temperature variations within RICH 2.

Each of these possible sources of movement are discussed independently.

3.2.3.1 Pressure Effects on Mirror Movement

Pressure is not expected to contribute much to the movements observed on the monitoring cameras. Previous studies [71] have shown that large variations can have an effect, with the relationship given by

$$\Delta r = aT + bP, \tag{3.5}$$

where r can be the x or y coordinate of the light spot, T and P are the temperature and pressure respectively. The constants a and b can be extracted and used to disentangle the movements due to the temperature and pressure. The range of pressure monitored over the LHCb operational periods 2011-2012 should, however, have negligible impact on the beam spot movement. To illustrate this observation, the period of April - May 2011 for mirror 25, with the corresponding monitored values of alignment $\Delta\theta_x$, $\Delta\theta_y$ and ΔP (change in pressure) is provided in Figure 3.14.
For the data range given, the pressure does increase linearly. However, no correlation between pressure variation and mirror alignment is observed.

3.2.3.2 Magnet Polarity Shift on Mirror Movement

The materials used to construct the two RICH detectors are non-magnetic. This is to avoid distortions that would be caused by the high magnetic fields to which they are exposed. However, correlations between the changing of the magnet polarity and observed angular movements of the mirrors are a useful cross check to ensure that the RICH detectors remain stable as a function of magnetic field. A limited amount of magnet polarity data corresponding to the period May - July 2012 was available, which is presented in Figure 3.15 together with the alignment deviations recorded.
There does not appear to be an obvious correlation between the changes in magnet polarity and movements observed on the camera. This confirms that there is no magnetic dependent movement occurring, which is what would be expected.

3.2.3.3 Temperature Effects on Mirror Movement

To determine if the movement deviation observed is related to temperature changes, it is important to study the correlations between change in temperature ΔT and change in beam position Δx or Δy. A correlation can be seen when examining Figures 3.16, with a clear correlation between the y coordinate movement and the temperature change over the corresponding period. The x coordinate tends to show an anti-correlation but is less clear than the y data.

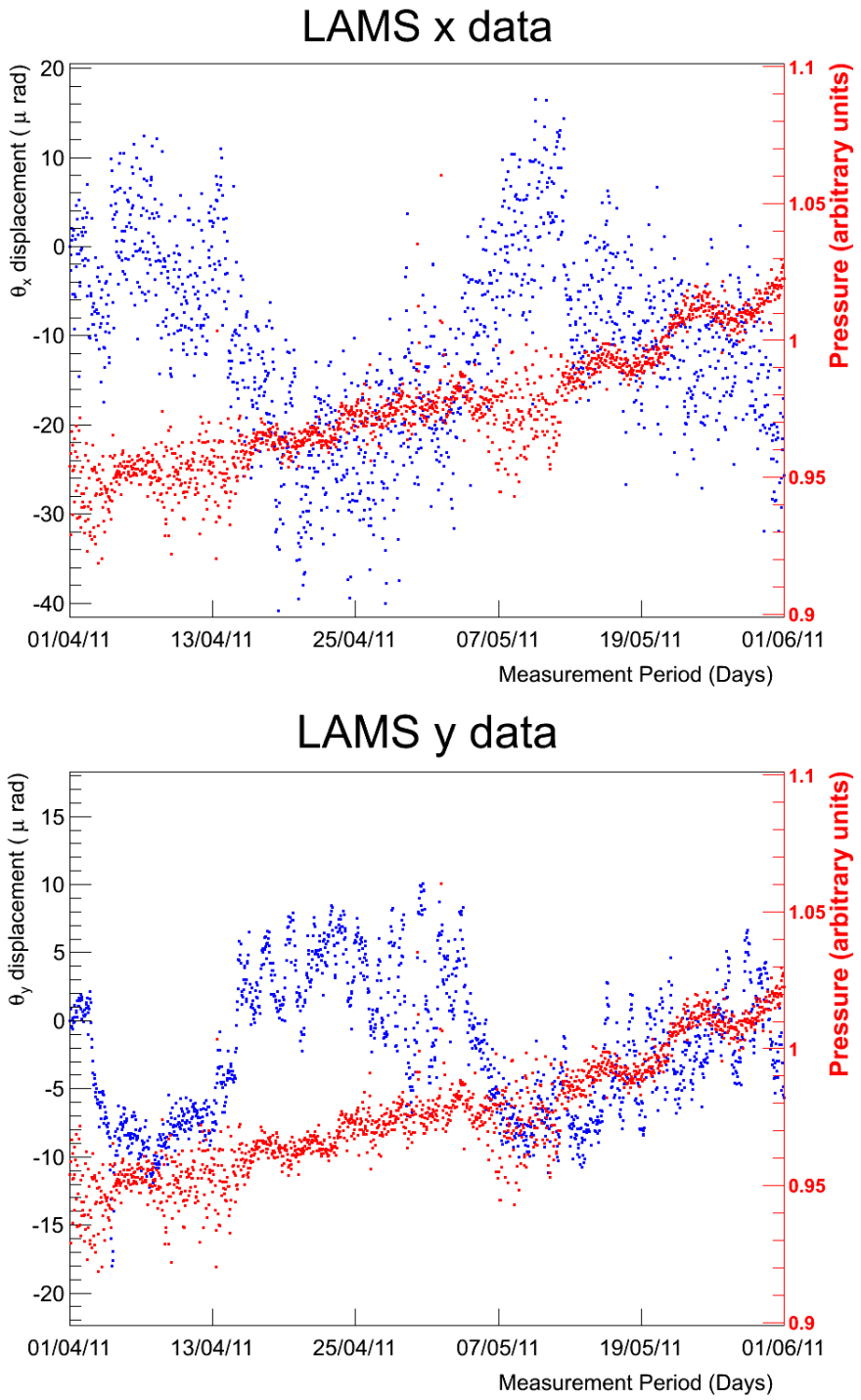


Figure 3.14: $\Delta\theta_x$ with ΔP data (top) and $\Delta\theta_y$ with ΔP data (bottom). Both datasets are from mirror 25 and correspond to the monitoring period April-May 2011.

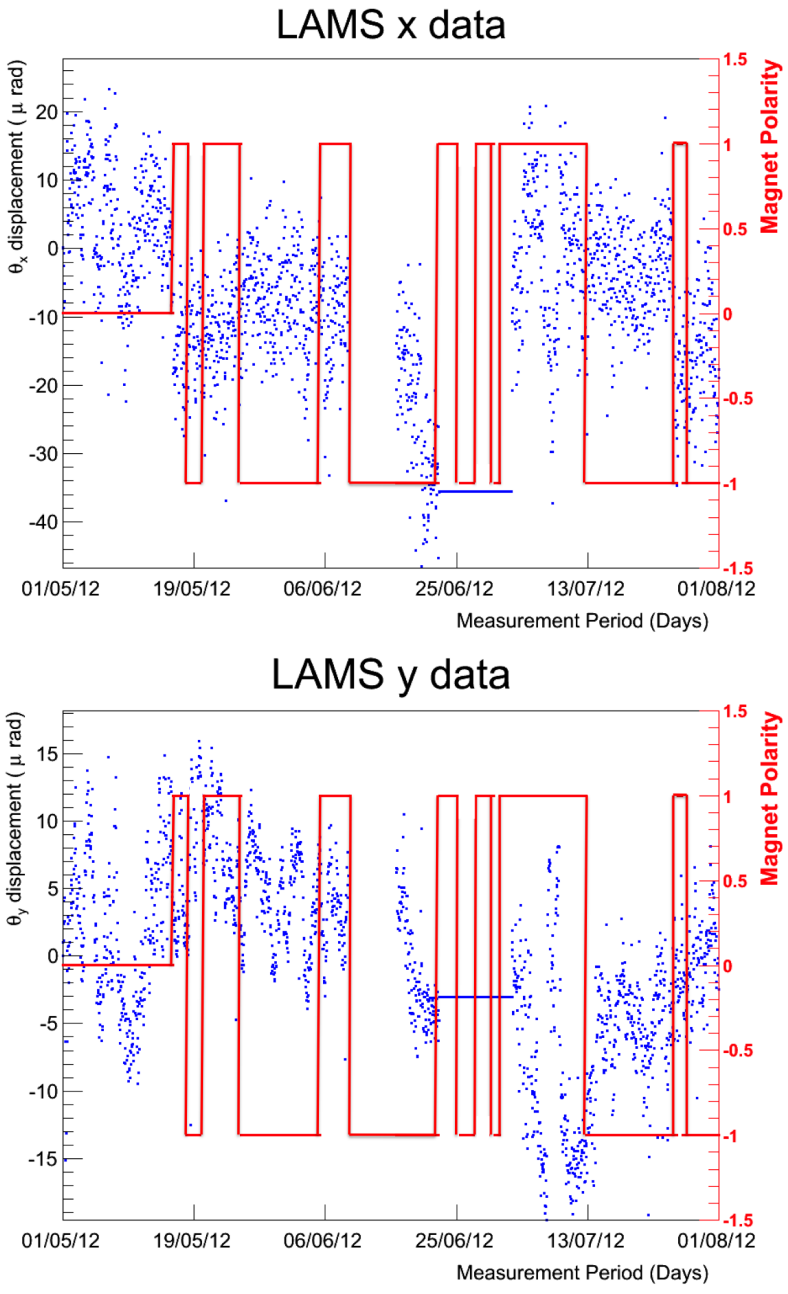


Figure 3.15: $\Delta\theta_x$ with magnet polarity data (top) and $\Delta\theta_y$ with magnet polarity data (bottom). Both datasets are from mirror 25 and correspond to the monitoring period May-July 2012.

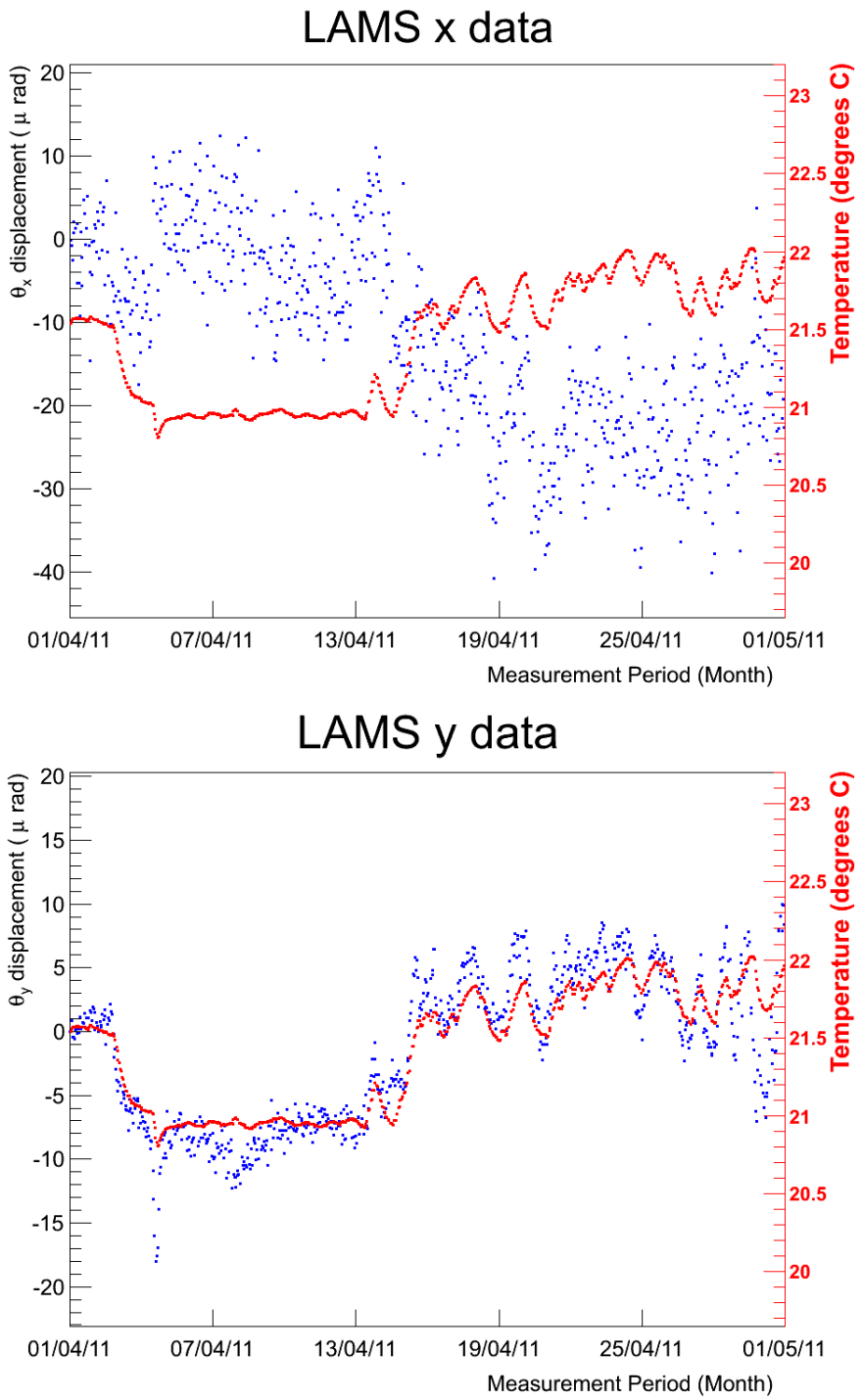


Figure 3.16: $\Delta\theta_x$ with ΔT data (top) and $\Delta\theta_y$ with ΔT data (bottom). Both datasets are from mirror 25 and correspond to the April 2011 monitoring period.

The correlation distributions of $\Delta\theta_x$ and $\Delta\theta_y$ against ΔT display this more clearly and can be seen in Figure 3.17.

One of the possible sources for this movement could be due to the thermal expansion and

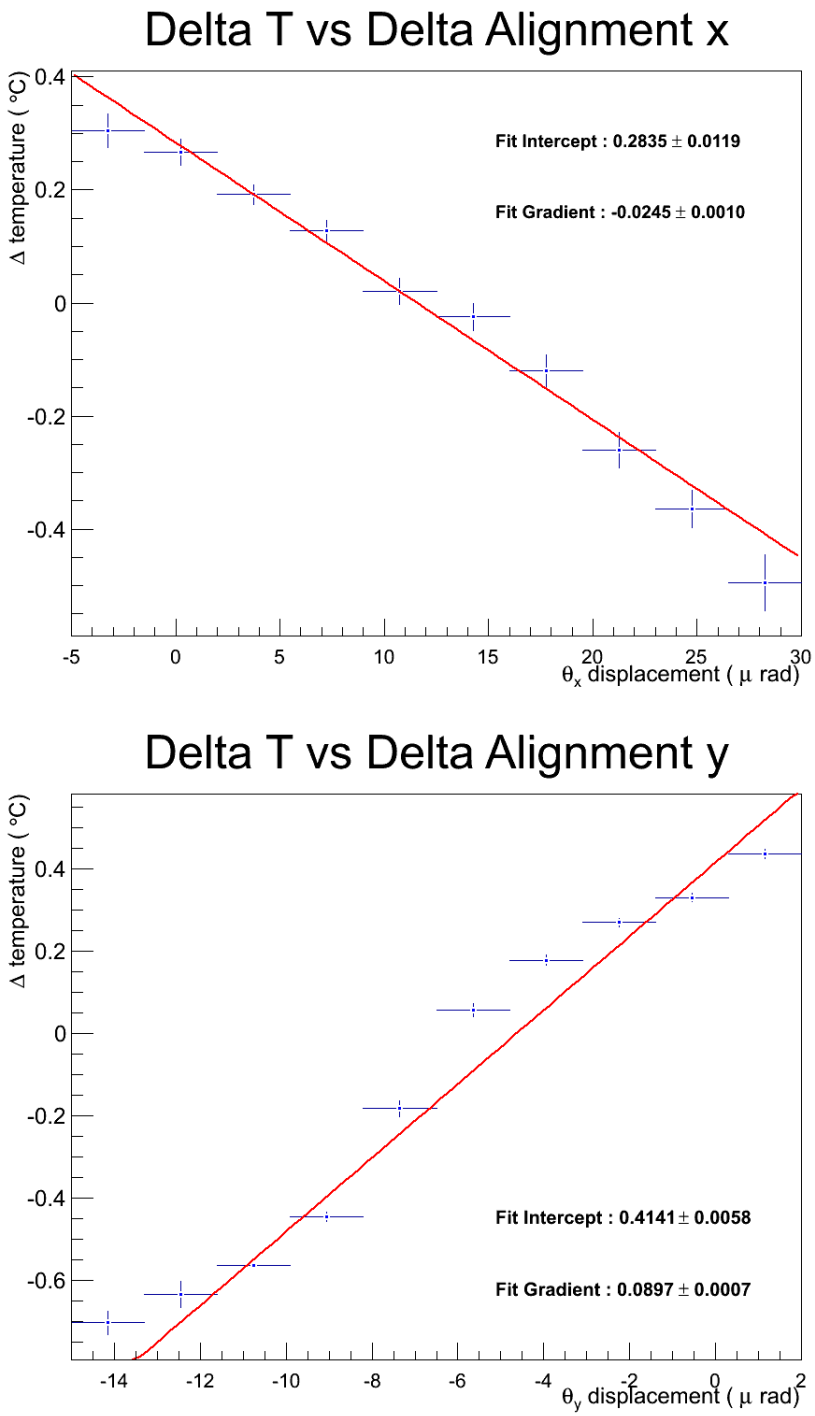


Figure 3.17: $\Delta\theta_x$ vs ΔT data (top) and $\Delta\theta_y$ vs ΔT data (bottom). Both datasets are from mirror 43 and correspond to the April to May 2011 monitoring period.

contraction of the materials. The mirror material, pyrex, has a thermal expansion coefficient of $\approx 4.0 \times 10^{-6}$ m/K. The expansion coefficient of the aluminium mirror support is 5 times larger at 22×10^{-6} m/K. To test this hypothesis, the ray tracing code that was previously

used to obtain the calibration coefficients of the mirrors is used for this purpose. To simulate the expansion of the material in the mirror plane, the y-component of each ray-trace vector is varied such that the y position at which it intersects with the mirror plane is shifted by $\pm 1\,\mathrm{mm}$ (see Figure 3.18). The code is then run to determine the intersection of the correlating beam with the defined camera plane and the movement observed.

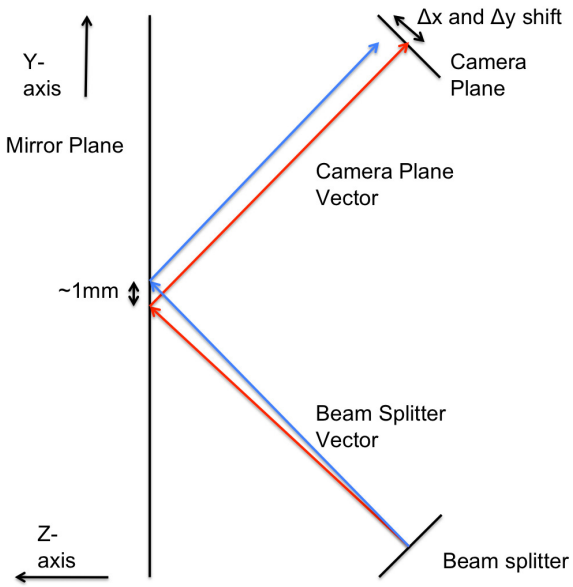


Figure 3.18: Schematic of how the beam splitter vectors were varied to simulate mirror movement due to thermal expansion.

Making the assumption that the movement is primarily originating from material expansion, it is reasonable to assume that the dominant expansion will be due to the aluminium support plate. By combining the movement obtained from the ray tracing code with the aluminium expansion coefficient, the resultant calculations show an expected angular variation as shown in the 'Al only' column of Table 3.6 (for $\Delta\theta_x$) and Table 3.7 (for $\Delta\theta_y$) for each mirror of RICH 2 monitored. Additionally, these calculations can be expanded to include the expansion of the mirror material, which causes the expected angular variation to increase slightly and is shown in the 'Glass only' column of Tables 3.6 and 3.7. A combination of the aluminium and glass expansions is shown as the column 'Al and Glass' of Tables 3.6 and 3.7. These values are of a similar order of magnitude as what is observed in the calibrated data, which is shown in Figure 3.16 for mirror 25, and is tabulated for all monitored mirrors in Table 3.5.

Side	Mirror Type	Mirror	Al only (μrad/K)	Glass only (μrad/K)	Al and Glass (μrad/K)
A-Side	Spherical	25	67.3	12.1	79.6
		09	19.8	3.6	23.4
		21	40.4	7.3	47.7
		22	40.3	7.3	47.6
	Flat	12	28.8	5.2	34.0
		04	30.0	5.5	35.5
		13	26.0	4.7	30.8
		05	28.0	5.1	33.1
C-Side	Spherical	26	27.7	5.0	32.7
		34	23.0	4.2	27.2
		27	30.0	5.5	35.4
		35	27.6	5.0	32.7
	Flat	49	40.3	7.3	47.7
		50	40.4	7.3	47.7
		38	19.8	3.6	23.3
		43	37.8	6.9	44.6

Table 3.6: Expected movement in $\Delta\theta_x$ for each mirror dependent on material expansion of support, mirror and both combined.

Side	Mirror Type	Mirror	Al only (μrad/K)	Glass only (μrad/K)	Al and Glass (μrad/K)
A-Side	Spherical	25	11.3	2.1	13.4
		09	4.7	0.9	5.6
		21	22.5	4.1	26.6
		22	24.8	4.5	29.3
	Flat	12	13.8	2.5	16.4
		04	8.3	1.5	9.8
		13	13.7	2.5	16.2
		05	8.4	1.5	9.9
C-Side	Spherical	26	8.4	1.5	9.9
		34	13.7	2.5	16.2
		27	8.3	1.5	9.9
		35	13.9	2.5	16.4
	Flat	49	24.9	4.5	29.4
		50	24.8	4.5	29.3
		38	24.2	4.4	28.6
		43	24.0	4.4	28.4

Table 3.7: Expected movement in $\Delta\theta_y$ for each mirror dependent on material expansion of support, mirror and both combined.

3.2.4 Discussion and Summary

A study of the Laser Alignment Monitoring System (LAMS), which monitors the stability of the flat and spherical mirrors in the RICH structure, is presented. Ray tracing software was used to determine calibration constants for each mirror, this is then used to convert the observed beam movement on the cameras into a $\Delta\theta_x$ and $\Delta\theta_y$ movement for their respective mirrors. Potential sources of the observed movement are then considered, these include the monitored pressure, temperature and magnet polarity data. The deviation observed is seen to correlate heavily with temperature, with measured values being of the order of $10 - 70\ \mu\mathrm{rad}/\mathrm{K}$. These are approximately within the same order as calculated values from simulation studies. The angular resolution of RICH 2 determined from data is $0.68 \pm 0.02\ \mathrm{mrad}$, which was described in Section 2.7.2 and shown in Figure 2.15. As such, the observed angular mirror movement of $\sim 50\ \mu\mathrm{rad}$ is negligible. It can therefore be stated that the mirror stability does not affect the RICH 2 resolution and performance.

3.3 Refractive Index Studies

Knowledge of the refractive index, n, of the RICH gases is needed to perform the reconstruction of particle events detected in the RICH subdetectors. It is currently determined using a method that calculates its expected value from the Cherenkov angle, θ_c, measured from high momentum track data and corrected by values of pressure and temperature, which are monitored via the RICH software. Additionally, the gas composition is determined to a high precision approximately once every month using a gas chromatograph, but these data have not been used by the RICH reconstruction so far. The Cherenkov angles are measured routinely, but not over specifically defined intervals. This study proposes an alternative method to determine the refractive index of the gas radiators in the RICH subdetectors, primarily applying this method to data from RICH 2. This alternative method allows for finer granularity in the determination of the refractive index. By using data for the Cherenkov angle, the absolute molecular mass of the gas, pressure and temperature, the values of n can be interpolated between calibration runs more accurately. This section will begin by discussing in further detail the current method of determining the refractive index, followed by the application of the alternative method and the differences observed between the new values of n and those currently used from the RICH database.

3.3.1 Current Determination of Refractive Index

The refractive index stored in the RICH database when performing reconstructions is determined using a number of variables. Currently the value used from the LHCb database is determined based on the wavelength dependence of the refractive index, assumed to follow the Sellmeier equation [72], given by

$$n^2(\lambda) = 1 + \frac{B\lambda^2}{\lambda^2 - C}, \tag{3.6}$$

where n is the refractive index, λ is the wavelength and B and C are experimentally determined Sellmeier coefficients [73]. This value of the refractive index also needs to be corrected for changes measured in the environment of the RICH detectors, which will naturally change the true value, in particular, changes in pressure (P) and temperature (T). The RICH Experimental Control System (ECS) performs a variety of tasks [57]. These include using predefined sequences for normal detector operation and automating actions to protect the equipment when monitored parameters fall outside their specified ranges. The RICH ECS also collects environmental information, in particular it logs the temperature and pressure of the gas radiators in each RICH. When the values of these environmental variables change, it is necessary to propagate them to the RICH conditions database where they are

used to re-evaluate the refractive index of the gas radiators, which is also dependent on the exact composition of the gas mixture. Variations of this composition in time can affect the calculated value of n, which can then affect the performance of the particle identification algorithms.

The primary method of calibrating the refractive index is performed using high momentum saturated pion tracks, fitting the observed Cherenkov angle, θ_c, and determining the offset ($\Delta\theta_c$) of the peak from the expected value for the nominal refractive index $n(\lambda)$. The nominal value of $n(\lambda)$ is determined from Equation 3.6 and other components, such as the quantum efficiency, which are described in [57]. Any deviation of $\Delta\theta_c$ from zero implies a change in the refractive index. Examples of this for RICH 1 and RICH 2 are given in Figure 2.15.

The offset observed, $\Delta\theta_c$, requires a correction factor to be applied in order to pull the value back to be equal to that of the expected value. In this case $\Delta\theta_c$ equals zero. The correction factor is determined from simulation where a small shift in the generated θ_c is applied. This allows a linear function to be used to determine the correction factor from the offset to the expected value. This procedure is performed on a run by run basis.

3.3.2 Alternative Method to Determine the Refractive Index

The alternative method being proposed to determine the refractive index, n, is by relating its value to monitored data only. This allows the value to be re-evaluated based on data logged directly by the ECS for any given time, and automatically incorporates measurement effects. The ideal gas law equation gives the state of any hypothetical gas with respect to its pressure, volume and temperature, and is given by

$$PV = nRT, \tag{3.7}$$

where P is the pressure in $\mathrm{kg/m^3}$, V is the gas volume in $\mathrm{m^3}$, n is the number of moles of the gas, R is the ideal gas constant in $\mathrm{J\ K^{-1}\ mol^{-1}}$ and T the temperature in K. The density of a gas medium may be related to its refractive index, n_{data}, by the following relationship,

$$\rho = a(n_{data} - 1), \tag{3.8}$$

where ρ is the gas density and a is a proportionality constant. The expression relating the refractive index of the gas radiator to the environmental variables can then be given by

$$(n_{data} - 1) = \frac{1}{a}\frac{PM}{RT}, \tag{3.9}$$

where the constant a needs to be determined from data to be able to calculate the value n_{data}, and M is the mass of one mol of the gas medium. The extraction of this constant is discussed

in the next section.

3.3.3 Evaluation of the Refractive Index Constant

In order to accurately determine the refractive index constant a, precise measurements of each variable at a specific time are required. Due to the ECS regular logging of temperature and pressure, the primary data restrictions are due to limitations in data points from the gas molecular mass and refractive index. The refractive index used in Equation 3.9 is the value determined from the high momentum tracks as displayed in Figure 2.15. This is calculated by using the relationship

$$\cos\theta_c = \frac{1}{n\beta}, \tag{3.10}$$

where β has the standard meaning

$$\beta = \frac{v}{c}, \tag{3.11}$$

and we assume $\beta = 1$ for high momentum tracks. Days during the 2012 data-taking when measurements of molecular mass of the gas, and the Cherenkov angles, were both available are selected in order to calculate the refractive index. The molecular mass of the gas varies in time and depends on the relative concentrations of the main radiator CF_4 and other gases in the mixture, such as CO_2, N_2 and O_2. These measurements were performed using a gas chromatograph during specific calibrations carried out by CERN personnel throughout data taking [74]. The Cherenkov angles were determined by high momentum track reconstruction calibrations carried out regularly throughout the LHCb runs, as has been discussed in Sections 2.7.2 and 3.3.1. The gas chromatograph is a chemical analysis instrument that is able to separate out the relative fractions of a composite gas into its constituent components. This allows accurate measurements of the gas components listed in Table 3.8. The refractive index, determined from the Cherenkov angle data, for the same dates is provided in Table 3.9.

Date	O_2 %	N_2 %	CO_2 %	CF_4 %	$< M >$ (kg/mol)
19/04/2012	0.03	0.8	5.8	93.3	0.0849
08/05/2012	0.02	0.8	4.9	94.3	0.0854
08/06/2012	0.03	0.8	4.2	95.0	0.0857
06/07/2012	0.03	0.8	5.0	94.2	0.0853
22/08/2012	0.03	0.6	5.0	94.2	0.0853
03/09/2012	0.03	0.6	4.9	94.4	0.0854
02/10/2012	0.03	0.6	3.9	95.4	0.0859

Table 3.8: Gas composition fractions measured via chromatograph and the date at which they were measured. These measured data points have corresponding refractive index values from data and P and T data. The values do not add exactly to 100% due to how the fractions were determined by the chromatograph. The final column is the molecular mass of the gas determined from the composition.

Date	RICH database n	Fitted n from θ_c data
19/04/2012	1.00043082	1.00041993
08/05/2012	1.00043253	1.00043073
08/06/2012	1.00043092	1.00042861
06/07/2012	1.00043170	1.00042931
22/08/2012	1.00043348	1.00043079
03/09/2012	1.00043385	1.00043120
02/10/2012	1.00043036	1.00043215

Table 3.9: Refractive index values for use in the calibration. The values extracted from the RICH database as determined by the method discussed in Section 3.3.1 and the values determined from fitting the Cherenkov angle of high momentum saturated track data discussed in Sections 2.7.2 and 3.3.1 and displayed in Figure 2.15.

The average value for a given day of the second component of Equation 3.9 can then be evaluated. Fitting a linear function to the distribution obtained from the $(n_{data} - 1)$ vs $\frac{PM}{RT}$ gives a value for the required constant, a. This fit is displayed in Figure 3.19.

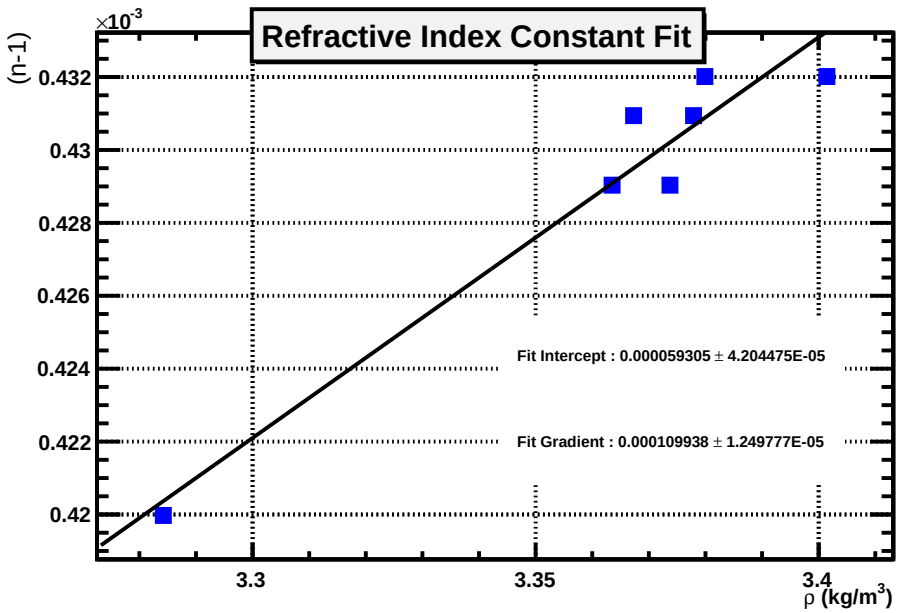


Figure 3.19: Linear fit to the components of Equation 3.9 using data from Tables 3.8 and 3.9 to determine the constant of proportionality. Errors included are smaller than size of data point symbols.

The results of the linear fit in Figure 3.19 introduces an extra term into Equation 3.9, such that the relationship becomes

$$(n_{data} - 1) = \frac{1}{a}\frac{PM}{RT} + C, \tag{3.12}$$

where $\frac{1}{a} = (1.10 \pm 0.13) \times 10^{-4}\ \mathrm{m}^3/\mathrm{kg}$ and $C = (5.93 \pm 4.20) \times 10^{-5}$. These constants allow the evaluation of the n, θ_c and the average molecular mass $< M >$ for any logged pressure and temperature with a valid time stamp.

3.3.4 Application of Alternative Method to Data

The full implementation of the alternative method on data allows finer granularity in the movement of three key variables; θ_c from data, refractive index n and the average molecular mass $< M >$ of the gas. These three variables are displayed respectively in Figures 3.20, 3.21 and 3.22.

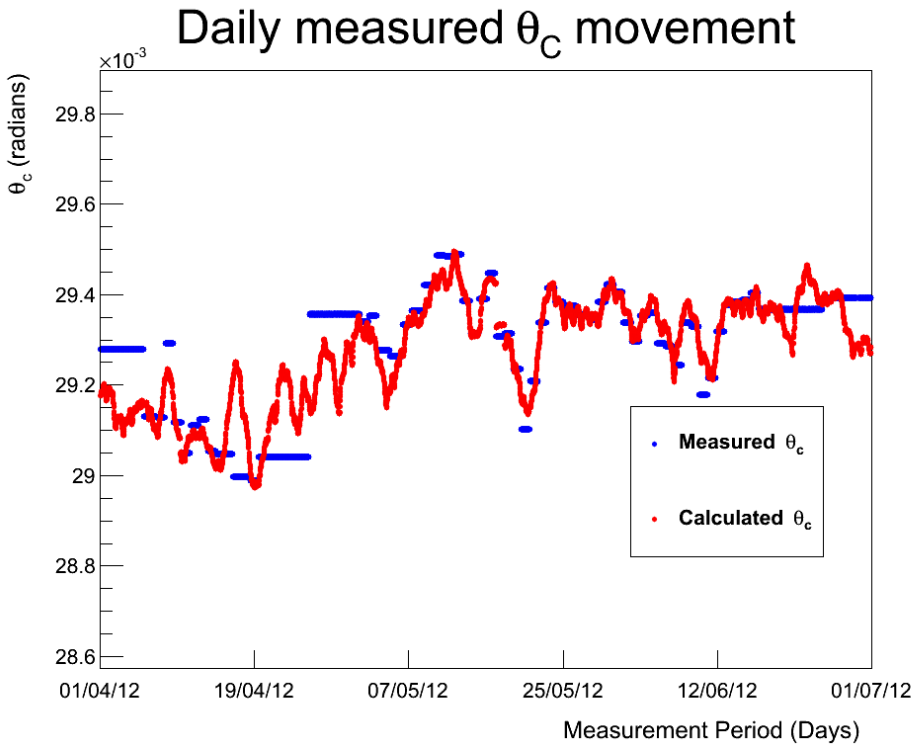


Figure 3.20: Raw Cherenkov angle, θ_c, measured from saturated track data. The blue distribution shows the values measured directly from data and the periods for which they are valid, the red data shows the values determined using the method detailed in section 3.3.2 and calibration discussed in section 3.3.3.

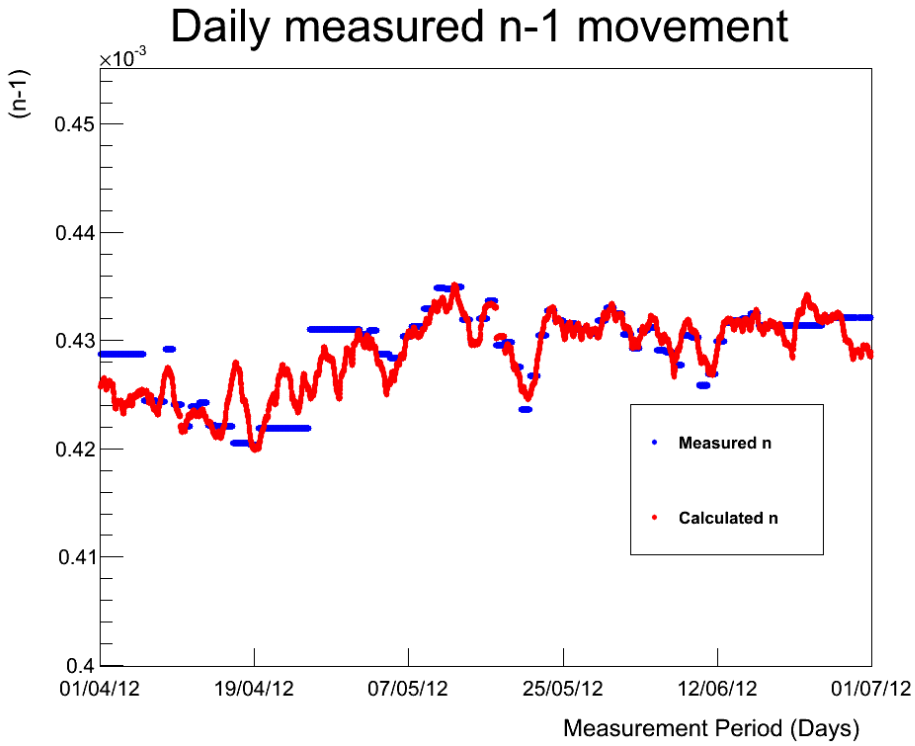


Figure 3.21: Raw refractive index, n, measured from data. The blue distribution shows the values measured directly from data and the periods for which they are valid, the red data shows the values determined using the method detailed in section 3.3.2 and calibration discussed in section 3.3.3.

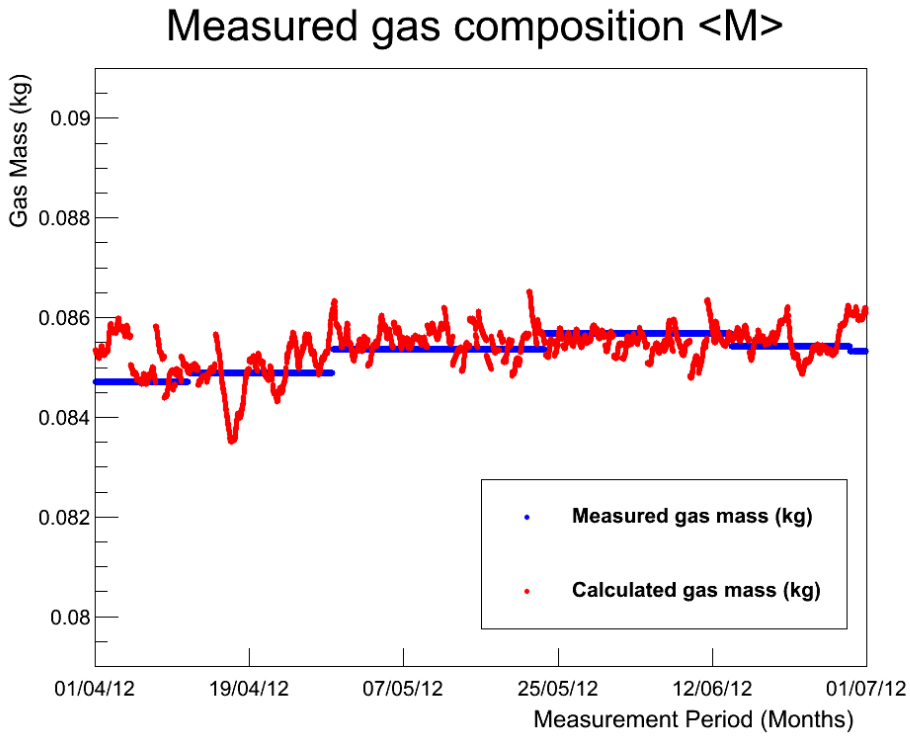


Figure 3.22: Average molecular mass, $< M >$, of the gas measured from data. The blue distribution shows the values measured directly from data and the periods for which they are valid, the red data shows the values determined using the method detailed in section 3.3.2 and calibration discussed in section 3.3.3.

It can be observed in the figures that changes in the RICH environment do cause the variables to fluctuate between actual calibration measurements. These are much easier to monitor when employing the alternative calibration method. If large fluctuations in particular variables were logged by the ECS over a short period of time, the alternative calibration method would prove a faster method to reconstruct these distributions and would allow to re-evaluate values of n, θ_c and $< M >$ to incorporate into the database.

It is also possible to determine the offset between the values calculated using the current method and the values determined using this alternative method. The relative difference between the refractive index and θ_c values determined from the alternative method against the values stored in the RICH database are provided in Figures 3.23 and 3.24 respectively. The distributions of these differences over the data range in the figures given is relatively constant, and the differences between the refractive indices and Cherenkov angles are less than 1%.

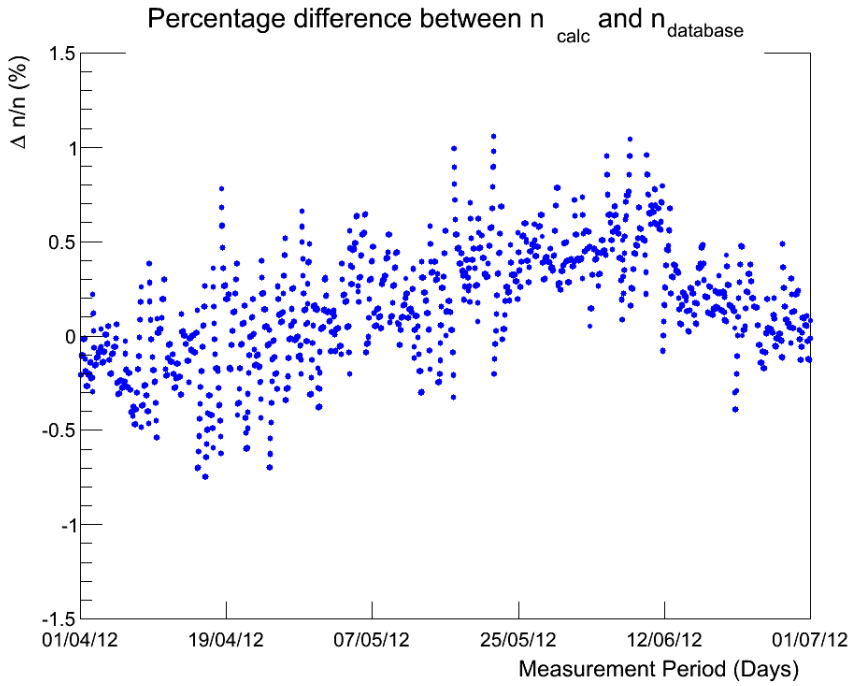


Figure 3.23: Distribution of the value of $\Delta n/n_{database}$ as a percentage for the period of April to June 2012

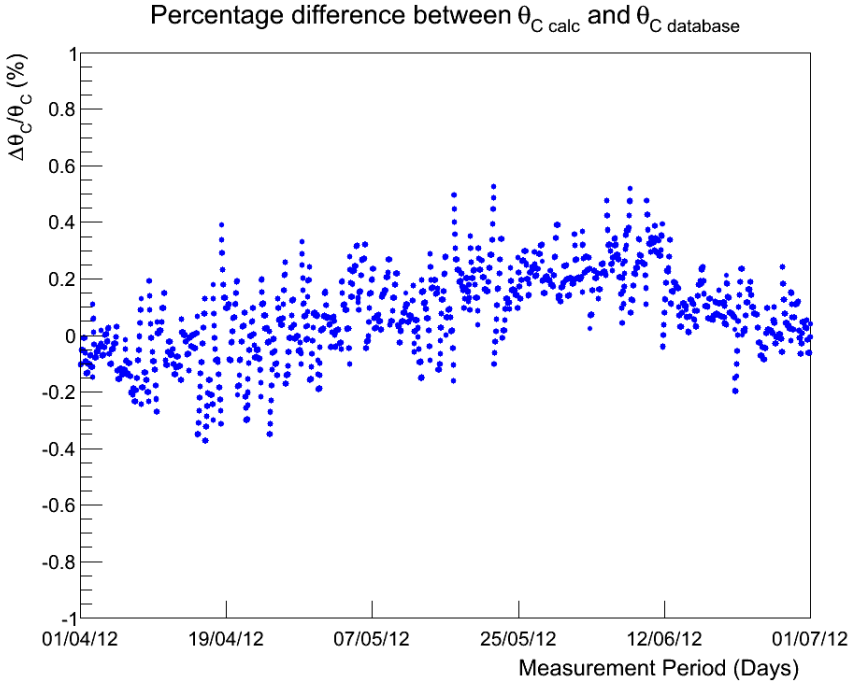


Figure 3.24: Distribution of the value of $\Delta\theta_c/\theta_{c\ database}$ as a percentage for the period of April to June 2012

3.3.5 Summary

A second study of RICH monitoring data proposes a purely data driven method to evaluate the refractive index of the gas radiators in the RICH. The method suggested differs from the current determination of the refractive index by using only environmental monitoring data, with the prior assumptions being the ideal gas law and relationship between gas density and refractive index. An additional advantage is the ability to track, using finer measurements in time, the fluctuations in refractive index n, expected Cherenkov angle θ_c and molecular gas mass $< M >$ that occur between current measurements of these quantities. After initial calibration with data samples we find the values of the constants that relate $n - 1$ linearly with $\frac{PM}{RT}$ to be $\frac{1}{a} = (1.10 \pm 0.13) \times 10^{-4}$ m^3/kg and $C = (5.93 \pm 4.20) \times 10^{-5}$. The measurements displayed excellent agreement between methods, with differences between measured n values being of the order $\approx 1\%$ and θ_c of the order $< 1\%$. Fluctuations of the molecular gas mass, $< M >$, that were not previously observed, are now much easier to track.

Chapter 4

Lifetime Determination Method

4.1 Introduction

A particle lifetime measurement is essentially performed by fitting some function $f(x)$ to the observed decay time distribution from data and extracting the parameters. In the simplest model, the particle decay time will follow the exponential decay distribution given by

$$N(t) = N_0 e^{-\frac{t}{\tau}} \tag{4.1}$$

and shown in Figure 4.1. Where $N(t)$ is the number of decay particles at the time t, N_0 is the number at the time $t = 0$ and the lifetime constant τ which is the mean decay time of the particle.

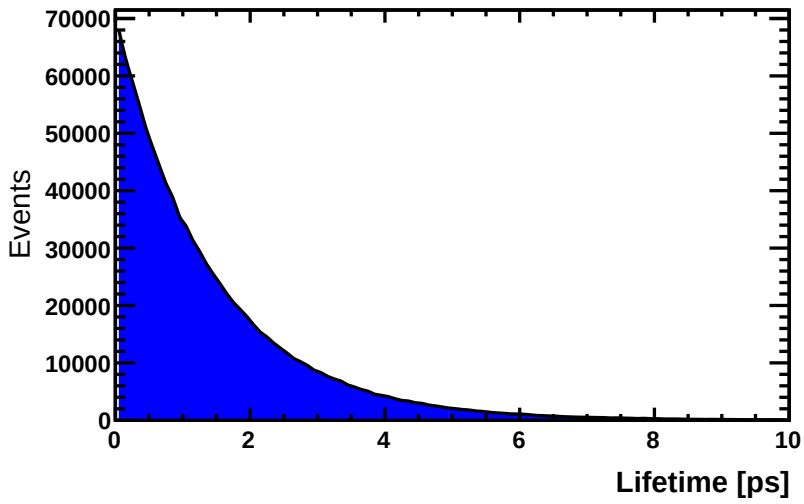


Figure 4.1: General unbiased decay time distribution, following the exponential decay distribution given by Equation 4.1.

In High Energy Physics (HEP) experiments, this simplistic model can be assumed to be the basis for determining the value of the lifetime of a particle.

In the lifetime measurements detailed in Chapters 5 and 6 of this thesis we deal with both B^0_d and B^0_s meson decays. These particles have relatively long lifetimes in particle physics terms, with current world averages from the Heavy Flavour Averaging Group (HFAG) [36] of $\tau_{\mathrm{B}^0_d} = 1.519 \pm 0.007\,\mathrm{ps}$ and $\tau_{\mathrm{B}^0_s} = 1.466 \pm 0.031\,\mathrm{ps}$ respectively. To select these longer decay time events, we use cuts on quantities that are generally lifetime biasing, such as Impact Parameter IP, Impact Parameter Chi-squared $IP\chi^2$ and Flight Distances FD. Descriptions of these variables are provided in Section 5.3.2.

In a hadronic environment such as that at the LHC, we are forced to cut harder on such variables due to the much larger and varied backgrounds that are produced in the collisions. These would not be as prevalent in a cleaner leptonic collider environment, which would allow us to apply looser cut criteria to select signal-like events.

In general, the imposition of these cuts causes the observed decay time distribution to deviate from the standard exponential distribution given by Equation 4.1, such that we observe a distribution similar to that shown in Figure 4.2.

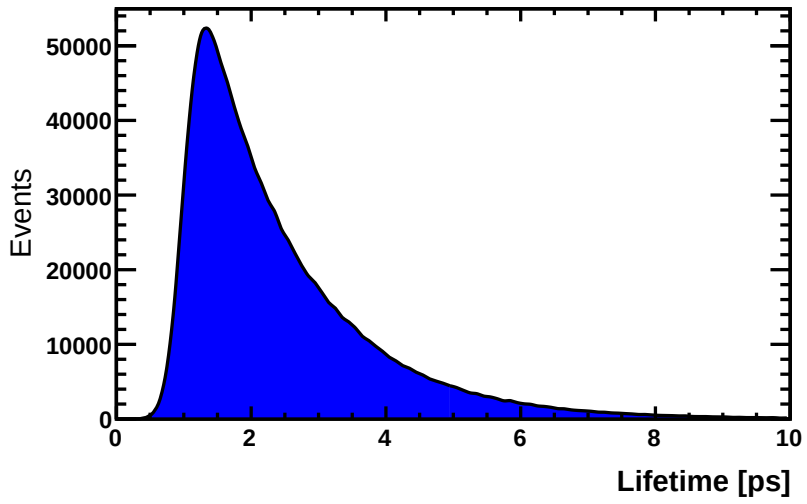


Figure 4.2: Example of an observed biased decay time distribution generated via a simplified Monte Carlo simulation event generator (discussed in Section 5.6.1).

The event selection procedure moves through a number of stages, with each stage accruing a bias. These stages are the trigger, pre-selection and offline selection. The conditions for each of these stages will differ depending on the analysis, with the specific conditions used for the analysis of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ using 2010 and 2011 data detailed in this thesis, found in Chapter 5. Particle Identification (PID) criteria are only available for offline analysis and so cannot be used in the online selections, however these

cuts are generally not lifetime biasing. The PID cuts allow the separation of the signal modes via the final state decay particles, which are described in further detail in Section 5.3.3.
There are a variety of methods that have been used to remove the lifetime bias introduced by a selection. These include taking the ratio of the signal mode lifetime distribution with a control mode lifetime distribution, this is illustrated in the relative lifetime measurement of the $B_s^0 \to K^+K^-$, detailed in a recent LHCb analysis [75, 76]. The ratio method makes the assumption that the acceptances of two kinematically similar channels should be identical and will cancel within a high degree of precision. By performing this ratio, the acceptance functions do not need to be calculated, leaving the lifetime of the reference channel as the only dependency. In general this lifetime should be well known and with small uncertainties.
Another method that has been used is to determine the acceptance function from fully generated Monte Carlo (MC) events. This relies heavily upon having very accurate models for the detector geometry combined with the hardware and software cuts that are applied to the events. The decay time used in the generation of the signal events must also be known such that the acceptance function (see Figure 4.3) for the signal can be determined.
The method used for the analyses performed in this thesis (Chapters 5 and 6) uses a data-driven approach. This method calculates event-by-event acceptances that are used in the lifetime fit to build up average acceptance functions per decay mode. The advantage of this over other demonstrated methods is its reliance purely on the data provided, with there being no dependency on MC or external values to calculate the lifetime.

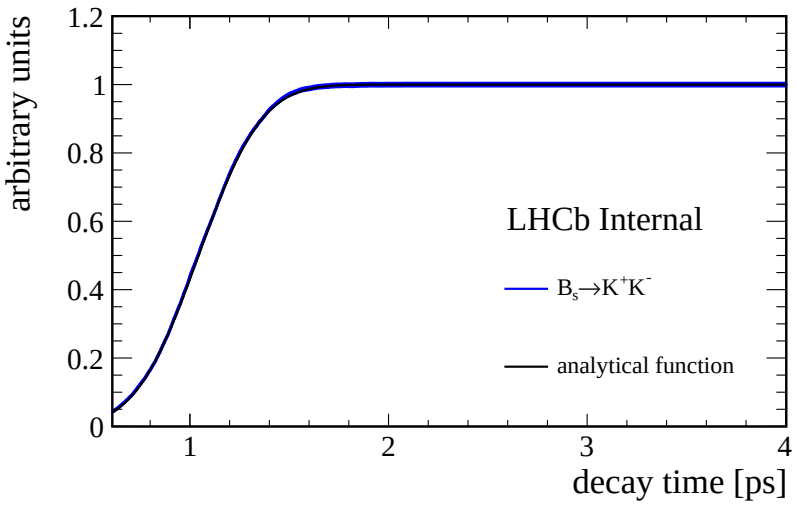


Figure 4.3: Average acceptance function from a $B_s^0 \to K^+K^-$ 2011 simplified simulation [76].

The methodology of this data-driven approach will be described in detail in this chapter. The chapter begins by describing the general lifetime fit method and the motivation for factorising

the fit into two components, the fit to the invariant mass and the fit to the reconstructed lifetime spectrum. This is provided in Section 4.2. The fit to the invariant mass distribution, which classifies the event type and determines the resultant event weightings, is detailed in Section 4.3. Following this, the details of how the acceptance function is determined from data is given in Section 4.4. Finally, the fit to the reconstructed lifetime distribution that results in the average lifetime of the signal decay is discussed in Section 4.5.

4.2 General Lifetime Fit Method

The lifetime fit, including reconstruction biases, of an individual decay is obtained using an unbinned maximum likelihood fit. This is analogous to the method used in the fit to the invariant mass spectrum described in Section 4.3. The total probability density for measuring the set of observables is given by

$$f(t, A, m) = f(t, A|m) \cdot f(m), \tag{4.2}$$

and is dependent on the reconstructed decay time t, the acceptance function A and the distinguishing variable mass m. The decay time reconstruction is performed by the LHCb software package DaVinci, introduced in Section 2.11. The algorithm employed uses the primary vertex, secondary vertex, particle momentum and measured mass as input to determine the candidates proper time. Equation 4.2 has two main components. The first, $f(t, A|m)$, will be described more succinctly in Equation 4.3. The second, $f(m)$, is the Probability Density Function (PDF) determined from the prior fit to the invariant mass where m is the mass and also the distinguishing variable. The lifetime part of the total probability function in Equation 4.2, can be factorised into two components given by

$$f(t, A|m) = \sum_{class} f(t, A|class) \cdot P(class|m), \tag{4.3}$$

where $P(class|m)$ is determined via a separate fit to the invariant mass, detailed in Section 4.3. The component $P(class|m)$ is the probability of an event belonging to a specific signal class given its mass, as defined in Equation 4.10 and described in Section 4.3.3.1. The remaining factor in the time probability density for a given class is given by

$$f(t, A|class) = f(t|A, class) \cdot f(A|class), \tag{4.4}$$

with the first term being the probability density of measuring time t given the lifetime acceptance function and a particular class, and the second term being the probability density of having this acceptance function for the given class. This factorisation allows the evaluation

of the probability per event of each event having an observed lifetime, given a particular signal class and a given per-event acceptance function. The determination of the acceptance function is given in Section 4.4.

4.3 Fit to Mass Spectrum

The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ channels are interesting for a variety of physics reasons overviewed in Section 1.4. One of the difficulties involved with analysing a channel from this family is that the invariant mass distributions tend to overlap when reconstructed. An example of this from the CDF collaboration [77] can be seen in Figure 4.4.

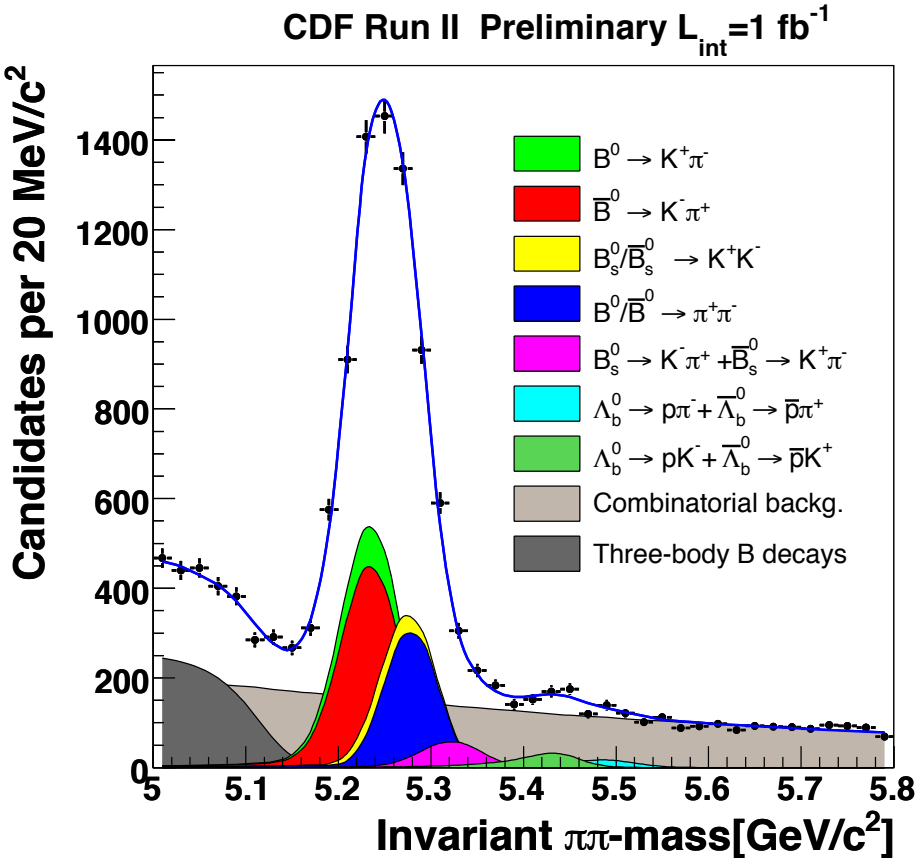


Figure 4.4: Invariant mass distribution from CDF data of $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ signal channels [77]. The data is separated by the application of energy deposition ($\partial E/\partial x$) constraints that are not as clean at distinguishing the signal channels as the RICH detector at LHCb.

As this family of decay channels are kinematically very similar, kinematic cuts designed for one of these channels will remove many backgrounds but not the other $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ modes.

This technical difficulty can be seen when applying kinematic cuts that are designed for the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ decay onto $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ selected data in Figure 4.5. The large signal peak in the figure is a superposition of all the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays in the collected data sample.

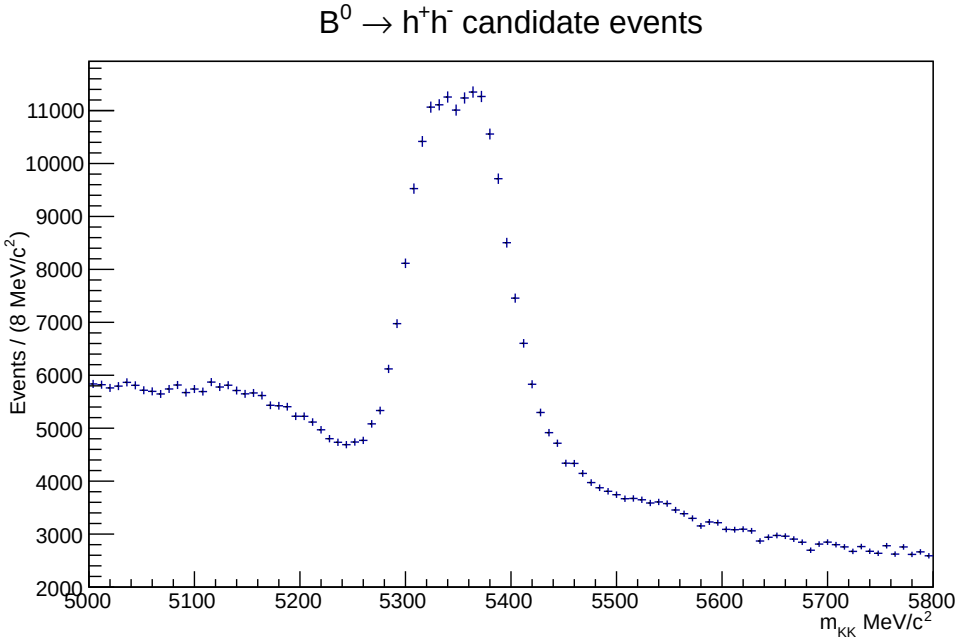


Figure 4.5: Full LHCb 2011 data after full trigger and kinematic cuts, but no PID imposed. The invariant mass distribution contains all $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ modes reconstructed under the KK mass hypothesis.

Applying PID specific cuts to the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ data, the resultant mass spectrum after fitting is given in Figure 4.6. The spectrum is much cleaner and, in this instance, is dominated by the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ decay channel, albeit with contamination from mis-identified $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ decays, partially reconstructed three-body decays and combinatorial background.

The first step of the general lifetime fit method detailed in Section 4.2, is a fit to the invariant mass spectrum. This is done to determine the relative fractions of each signal channel and also the signal and background probabilities per event. This process can be split into three parts:

- Determining analytical models that best represent the mass distribution of each signal class, and setting these as the models for their respective PDFs.
- Performing an unbinned maximum likelihood fit on the given data using the defined PDF models. This determines the parameters of the PDF models and the relative fractions for each signal class.

- Calculating event-by-event weightings per signal class.

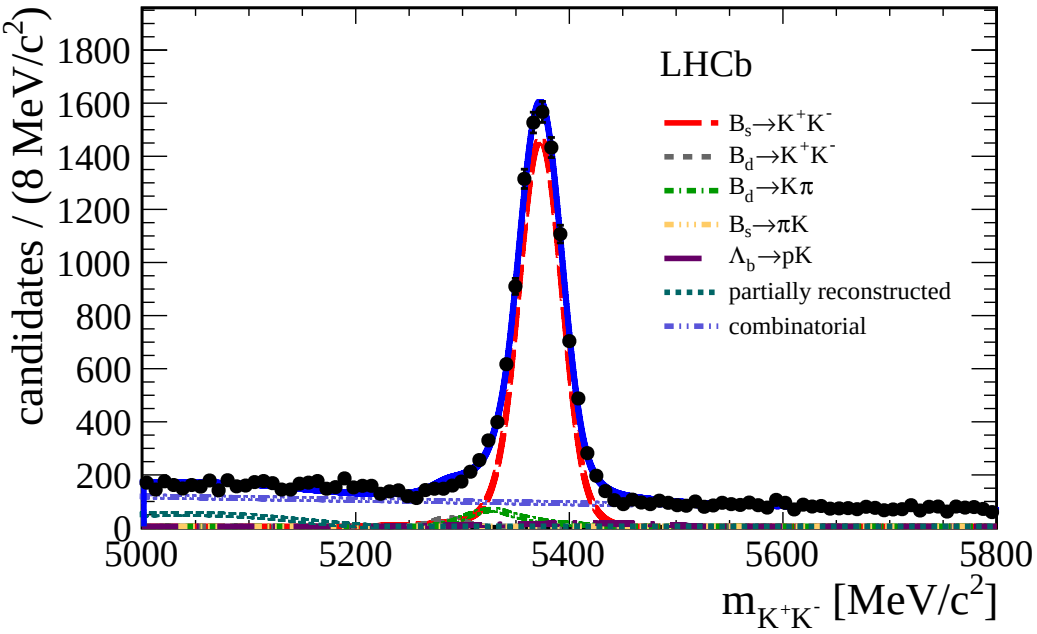


Figure 4.6: Selected $B_s^0 \to K^+K^-$ candidates from 2011 data with application of full kinematic and PID cuts. The invariant mass distribution is shown after the mass fit, displaying the remaining dominant signal channels under the KK mass hypothesis.

4.3.1 Mass Models

Whether or not a specific PDF is suitable for each signal class is dependent on the analysis being performed, as will be discussed in detail in Section 5.4. The PDFs implemented in the fits are either parametric or non-parametric. Non-parametric PDFs originate from distributions where the shape is not easily parameterised by an analytical function, whereas a parametric PDF is one which may be described by an analytical function. Parametric PDFs then require their respective parameters to be optimised for the observed data distributions.

4.3.2 Optimisation of Model Parameters

The optimum parameters of the invariant mass fit, which are the mass model PDF parameters and relative signal fractions, are determined by using the maximum likelihood method. Maximum likelihood is a statistical method for determining the parameters of a model, such that the values of the parameters are chosen in order to maximise the likelihood that the observed data agrees with the model. To understand this further it is useful to begin with describing

the use of PDFs.

In probability theory the probability density function, $f(x)$, of a continuous random variable, x, is a function that describes the relative likelihood for the given random variable to take a specific value. The probability of observing the variable x in the interval $[a, b]$ is:

$$P(x) = \int_a^b f(x)dx, \tag{4.5}$$

The normalisation of the PDF enforces the condition

$$P = \int_{-\infty}^{+\infty} f(x)dx = 1. \tag{4.6}$$

The PDF itself can take on many forms, with the most common used being a parametric PDF that has a shape defined by an analytical function determined by a set of free variables.

The likelihood function is given by

$$\mathcal{L}(\theta|x_1...x_n) = \prod_{i=1}^{n} f(x_i|\theta), \tag{4.7}$$

where x_i represents each of the observed values, n is the total number of observed values, θ is the vector of parameters related to the model used and $\mathcal{L}$ is the likelihood. The maximum likelihood states that the best estimator for the parameter vector is given by the set of values that maximises Equation 4.7. The method can be altered slightly such that the logarithm of the function, known as the log-likelihood function, is maximised instead. This is given by

$$\ln \mathcal{L}(\theta|x_1...x_n) = \sum_{i=1}^{n} f(x_i|\theta). \tag{4.8}$$

The fit is performed using the software package MINUIT [78], which is incorporated into the analysis software and offers a choice of several minimisation algorithms. Further details of the implementation are discussed in Section 4.6. To optimise the parameter values to attain a best fit to the data distributions, the MIGRAD algorithm is employed. This studies the evaluated values and rate of change of $\ln \mathcal{L}$ to try to obtain the maximum of the negative log-likelihood $\ln(\mathcal{L})$. The amount by which the parameters are varied is iteratively refined to give a high precision on their optimal values. The statistical uncertainties on these values are determined by finding the variation in the value of each individual parameter that results in a change in the log likelihood of $\Delta\ln(\mathcal{L}) = \frac{1}{2}$.

4.3.3 Event Weightings

There are three primary purposes of the mass fitter. The first two, which are performed simultaneously, are to determine the parameters of the PDFs and the relative signal fractions

from the data. The third is to calculate a per-event probability for each event belonging to a particular signal class. The parameters determined for each of the mass models used, allow the use of the correct mass PDFs to calculate these.

The two $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ analyses conducted in this thesis apply two different methods to calculate these weightings. The first, applied to the 2010 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analyses, calculates a per-event weighting using the Bayesian probability method. The second, which is used in the 2011 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analyses, uses the *sWeights* [79] method to determine the per-event weighting.

The event weights are used to separate out variable distributions in the lifetime fit, see Section 4.5, where the Bayesian weights are applied to 2010 data and the *sWeights* are applied to 2011 data. There is a subtlety in the method that actually means the Bayesian probabilities are used in both analyses, but not for variable separation in the lifetime fit. The method was updated to include the *sWeights* as these were found to provide a more accurate separation of the variable distributions [79], compared with the Bayesian probabilities.

4.3.3.1 Bayesian Probability

To determine the event weighting per signal class we must first understand the total mass distribution PDF. The total mass probability density function $f(m)$ can be written as the sum of the probability density functions $f(m|class)$ for each signal class present, multiplied by the corresponding relative fractions of that class $P(class)$. This is given by:

$$f(m) = \sum_{class} f(m|class) \cdot P(class). \tag{4.9}$$

The term signal class relates to all possible components of the distribution. This includes main signals, signal specific background and combinatorics. Using Bayes theorem, the probability of a single event to belong to a particular signal class can thus be expressed as

$$P(class|m) = \frac{f(m|class) \cdot P(class)}{f(m)}. \tag{4.10}$$

The values of $P(class|m)$ calculated per event are then stored and used in the lifetime fitter.

4.3.3.2 sWeights

The *sWeight* [79, 80] event weighting is used as an alternative to the Bayesian probability method detailed in Section 4.3.3.1. This approach is implemented as it is able to reproduce the original distributions with a higher accuracy than the Bayesian approach. The fit is performed on a discriminating variable, for this analysis the mass, which is assumed to be

uncorrelated with the control variable, in this analysis the decay-time. The *sWeights* are determined using

$$ {}_sP_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}, \tag{4.11} $$

where ${}_sP_n(y_e)$ is the sWeight per event per class with y_e the set of values for the discriminating variables of event e, $f_{j/k}$ the PDF of the discriminating variable, N_k the number of events expected for the signal class k, N_s the number of signal classes in the data sample and $\mathbf{V}_{nj}$ the covariance matrix given in [79].

4.4 Determination of the Acceptance Function

The equation for the single event probability density for measuring an event at decay time t, ignoring measurement errors, is given by

$$ f(t|A, class) = \frac{\frac{1}{\tau} e^{-t/\tau} A(t)}{\int_{-\infty}^{\infty} \frac{1}{\tau} e^{-t'/\tau} A(t') dt'}, \tag{4.12} $$

where $f(t|A, class)$ is the probability of the observed lifetime, conditional on the acceptance function $A(t)$ for this class and with τ the average decay time of the decay.

The lifetime bias that originates primarily from the selection cuts and gives a decay time distribution of the form displayed in Figure 4.2, may be corrected by using an acceptance function $A(t)$ in the fitter. Our method uses per-event acceptance functions, determined from data. To do this we employ the so-called 'swimming' algorithm. This method was initially developed and used at the NA11 spectrometer at the CERN SPS [81] and then further developed within CDF [82, 83] before subsequently being studied and implemented in LHCb [84, 85]. Previous use of the method has been complicated by the implementation of hardware based triggers that were used in experiments prior to LHCb becoming operational. This meant that the trigger hardware had to be simulated by custom written software in order to rerun the triggers to determine the event-by-event acceptances. This method is largely simplified in LHCb due to the triggers being almost entirely software based, with the hardware triggers employed applying cuts that are non-lifetime biasing. Thus all that is needed to rerun the trigger is simply to rerun the original software to determine these event-by-event acceptances.

The per-event acceptance function is characterised by a top hat function $\theta(t, t_{min}, t_{max})$ in time, where

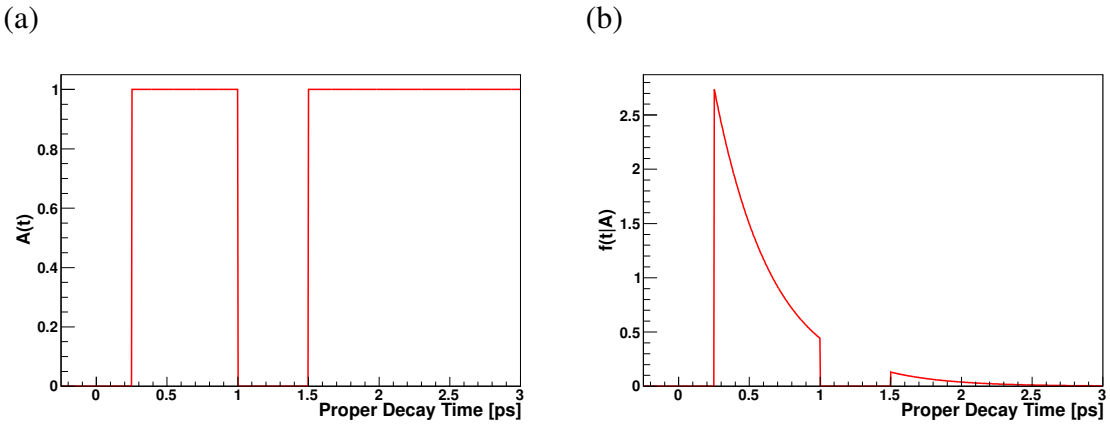


Figure 4.7: (a) An example of acceptance intervals generated for a single event. This example has three turning points and two acceptance intervals. (b) The decay time PDF resulting from the acceptance function in (a) [86].

$$\theta(t, t_{min}, t_{max}) = \begin{cases} 1 & t_{min} < t < t_{max} \\ 0 & \text{elsewhere.} \end{cases} \tag{4.13}$$

Implementing a single top-hat acceptance function with turning points t_{min} and t_{max}, into the lifetime calculation gives

$$f(t|A) = \frac{\frac{1}{\tau}e^{-t/\tau}\theta(t, t_{min}, t_{max})}{\int_{t_{min}}^{\infty} \frac{1}{\tau}e^{-t'/\tau}dt'}, \tag{4.14}$$

which can also be written as

$$f(t|A) = \begin{cases} \frac{\frac{1}{\tau}e^{-t/\tau}}{\int_{t_{min}}^{t_{max}} \frac{1}{\tau}e^{-t'/\tau}dt'} & t_{min} < t < t_{max} \\ 0 & \text{elsewhere.} \end{cases} \tag{4.15}$$

For events with more than one lifetime acceptance region, this may be generalised to a series of top hat functions in an analogous way. The swimming method determines the intervals of decay time where the lifetime is accepted or rejected based on the applied cuts. This is run for the trigger and offline selections independently to determine the intervals specific for each. The intervals for each of these components are then merged to determine the overall acceptance intervals per event.

The intervals are generated by re-applying the selection cuts for all hypothetical decay times, determining if the event would have been selected at that decay time. In practice, the swimming takes advantage of the fact that the decay time of the mother particle is independent of

the daughter particle kinematics, thus by moving the primary vertex along the momentum vector of the B particle, the intervals of decay time where the event is selected can be determined. This method is referred to as 'swimming'. The iterative movement determines the lifetime of the particle at each step, for which the decision on whether the candidate passes or fails the kinematic criteria of the trigger or selections is re-evaluated. The results of the pass/fail evaluation build up a set of binary step function(s) as a function of decay time, see Figures 4.7(a) and 4.7(b). The decay time values where this pass/fail acceptance changes from binary 0 to 1 or 1 to 0 are known as "turning-points". The basics of the method are illustrated in Figure 4.8, where as an example a lower IP cut is assumed, which translates into a simple step function as the decay time acceptance function.

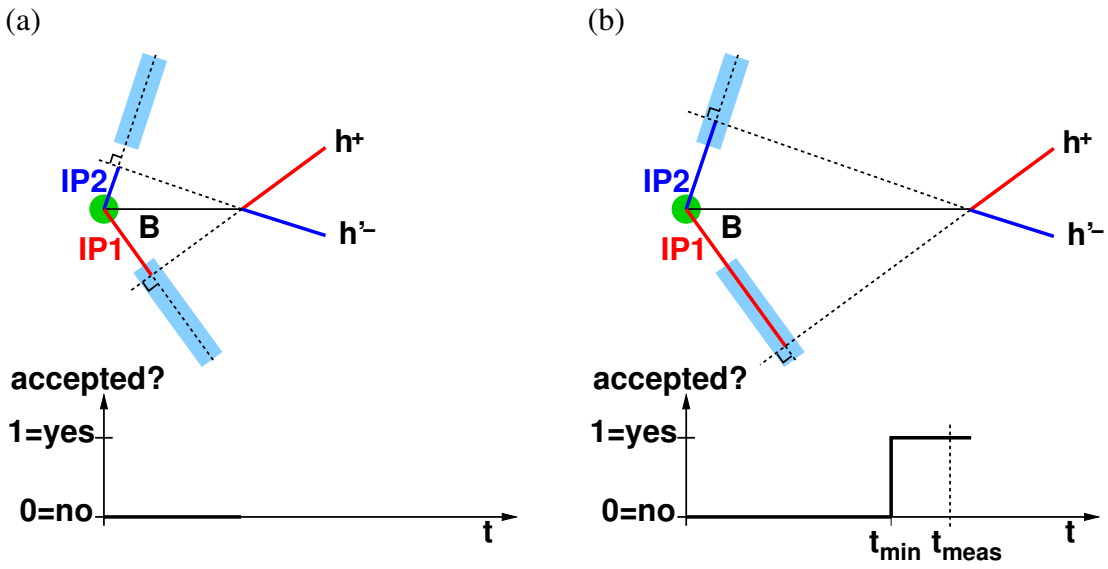


Figure 4.8: Decaytime acceptance function for an event of a two-body hadronic decay. The light blue (shaded) regions show the bands for accepting the impact parameter of a track. The impact parameter of the negative track (IP2) is too small in (a) and lies within the accepted range in (b). The actual measured decay time lies in the accepted region. The acceptance intervals give conditional likelihoods used in the lifetime fit.

The per-event acceptance functions can become more complicated when multiple primary vertices are reconstructed for a single event. Events which meet these conditions can have more than one step function representing their lifetime acceptance intervals, similar to the example in Figure 4.7(a). This, however, does not affect the performance of the fit.

4.5 Fit to Reconstructed Decay Time Spectrum

The general methodology for fitting the reconstructed decay time spectrum, including acceptance effects, is detailed in Section 4.2. The expressions for the main signal decay time PDFs are derived in Section 4.4, with the model parameters optimised using the methods discussed in Section 4.3.2. The effect of detector resolution on the lifetime measurement is taken into account in Section 4.5.1. Additionally for some of the backgrounds, for example the partially reconstructed and combinatorial, there are no analytical expressions which accurately describe the lifetime distributions. In these cases a non-parametric PDF is used, as described in Section 4.5.2. Furthermore, a description of how the average acceptance function is calculated is provided in Section 4.5.3. This details the general method and differences between the calculation performed in the analyses performed on 2010 and 2011 data. The full likelihood function for the lifetime fit is detailed in Section 4.5.4. Finally the technical details of the fit implementation are detailed in Section 4.6.

4.5.1 Detector Resolution Effects

As with all experimentally measured quantities, the value measured tends to be 'smeared' somewhat by the inherent resolution of the detector itself, and by the reconstruction software. To include this into the modelling of the decay time, the exponential decay distribution is smeared with a time dependent resolution function, $R(t, \sigma)$. The resolution function is assumed to be a single Gaussian with a width of σ, which is given using the standard Gaussian function

$$R(t, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}}. \tag{4.16}$$

The excellent VELO performance currently results in a decay time resolution of $\approx 50\,\text{fs}$ [87], this is included to accurately describe the physical model of the measured particle decay times. The measured decay time, t_{meas}, can thus be described by the 'true' decay time. This follows an exponential distribution plus an additional resolution term, δt, which follows a Gaussian with mean zero and σ equal to the average uncertainty on the measured proper decay time. This is given by

$$t_{meas} = t_{true} + \delta t. \tag{4.17}$$

The measured time distribution of the signal classes can then be determined analytically by convoluting the Gaussian resolution function with the exponential decay function of the lifetime, which includes the single step acceptance discussed in Section 4.4. This leads to

$$
\begin{aligned}
f(t_{meas}) &= \frac{1}{\tau} e^{-t_{meas}/\tau} \Theta(t_{meas}) \otimes \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{1}{2} t_{meas}^2/{\sigma_t}^2} \\
&= \int_{-\infty}^{\infty} \frac{1}{\tau} e^{-t'/\tau} \Theta(t') \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{1}{2}(t_{meas}-t')^2/{\sigma_t}^2} \mathrm{d}t' \\
&= \frac{1}{\tau} e^{-t_{meas}/\tau} e^{\frac{1}{2}\sigma_t^2/\tau^2} F\left(\frac{t_{meas}}{\sigma_t} - \frac{\sigma_t}{\tau}\right),
\end{aligned} \tag{4.18}
$$

where $\Theta(t)$ is the Heaviside function and F is the Gaussian frequency function defined by

$$
F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} \mathrm{d}y. \tag{4.19}
$$

Using this in Equation 4.15, which assumes an acceptance function that is a single top-hat function with limits t_{min} and t_{max}, results in the final equation for the lifetime probability density

$$
f(t_{meas}|t_{min}, t_{max}, signal) = \frac{\frac{1}{\tau} e^{-t_{meas}/\tau} e^{\frac{1}{2}\sigma^2/\tau^2} F\left(\frac{t_{meas}}{\sigma} - \frac{\sigma}{\tau}\right)}{N(t_{max}, \tau, \sigma) - N(t_{min}, \tau, \sigma)}, \tag{4.20}
$$

with

$$
N(t, \tau, \sigma) = -e^{-t/\tau} e^{\frac{1}{2}\sigma^2/\tau^2} F\left(\frac{t}{\sigma} + \frac{\sigma}{\tau}\right) - F\left(\frac{t}{\sigma}\right). \tag{4.21}
$$

The denominator in Equation 4.20 will change for events where the acceptance function consists of more than one top-hat function. The resulting equation to account for this becomes

$$
N_{tot} = \sum_{top-hats} N(t^i_{max}, \tau, \sigma) - N(t^i_{min}, \tau, \sigma). \tag{4.22}
$$

In the analysis performed in Chapters 5 and 6 we include events with up to six top hat acceptance function events, which can have up to ten primary vertices.

4.5.2 Non-Parametric PDFs for Lifetime Models

The decay time distributions of the main signal classes, $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$, are modelled using the single exponential function described at the beginning of this chapter by Equation 4.1. The combinatorial and partially reconstructed background classes present in the data, cannot trivially be described by an analytical PDF. These classes cannot be accurately modelled using the exponential lifetime decay model that is used for the main signal classes. In many lifetime measurements, finding a parametrisation for these background lifetime distributions has been difficult, so for the analyses conducted in this thesis and others [86] it has been deemed best to use non-parametric models.

The analyses applied on the 2010 and 2011 LHCb datasets have subtle differences between them. Both use the *kernel method* [88], however the per-event class weightings are calculated using different methods. The 2010 lifetime analysis uses a Bayesian probability to calculate the weightings, Sections 4.3.3.1, whereas the 2011 analysis uses *sWeights* [79], Section 4.3.3.2.

Another, very important, difference between the analyses is the determination of these non-parametric lifetime models. In the 2010 analysis, there is only one non-parametric lifetime model to determine. This is for the combinatorial background. The analysis takes advantage of the fact that the complete measured lifetime distribution is known, as well as the individual signal channel components which are modelled by exponentials. The background time distribution is thus obtained by subtracting all signal distributions weighted by the respective signal class probability from the total distribution. For a single event, this can be written in simplified form as

$$Total(t) - \sum_{class \neq bkg} f(t|class) \cdot A(t|class) \cdot P(class|X_i). \tag{4.23}$$

The per-event weightings calculated by the two methods discussed earlier, allow the separation of the total lifetime distribution into the constituent components of each signal class. The 2011 analysis, in particular, relies on the *sWeights* to determine the non-parametric lifetime models independently. In this analysis there is the combinatorial and partially reconstructed background models, so Equation 4.23 is no longer valid.

The *sWeights* can be used to make an *sPlot* of the reconstructed decay time, which normally takes the form of a histogram. As these are by definition binned, they can give rise to discontinuities that can be smeared by using kernels. The application of the *kernel method* allows us to smooth the decay time distribution so that it is well behaved in the maximum likelihood fit. The method assigns a Gaussian *kernel* for each event. This is centred at the measured decay time and with an area proportional to the Bayesian probability or *sWeight*, dependent on the analyses, of the particular background class for that event. The width of the Gaussian is a function of the total number of events (n), the resolution of the decay time (σ_τ) and the estimated density of events at that value of the decay time ($f_0(\tau_i)$). The width determines the level of smearing, and the value for event i suggested in [88] is

$$h(\tau_i) = \left(\frac{4}{3}\right)^{1/5} \sqrt{\frac{\sigma_\tau}{f_0(\tau_i)}} \cdot n^{-1/5}. \tag{4.24}$$

The density $f_0(\tau_i)$ is estimated differently depending on the analysis. The 2010 analysis used a histogram of the total decay time distribution, wheres the 2011 analysis uses a standard normalised *sPlot* of the decay time.

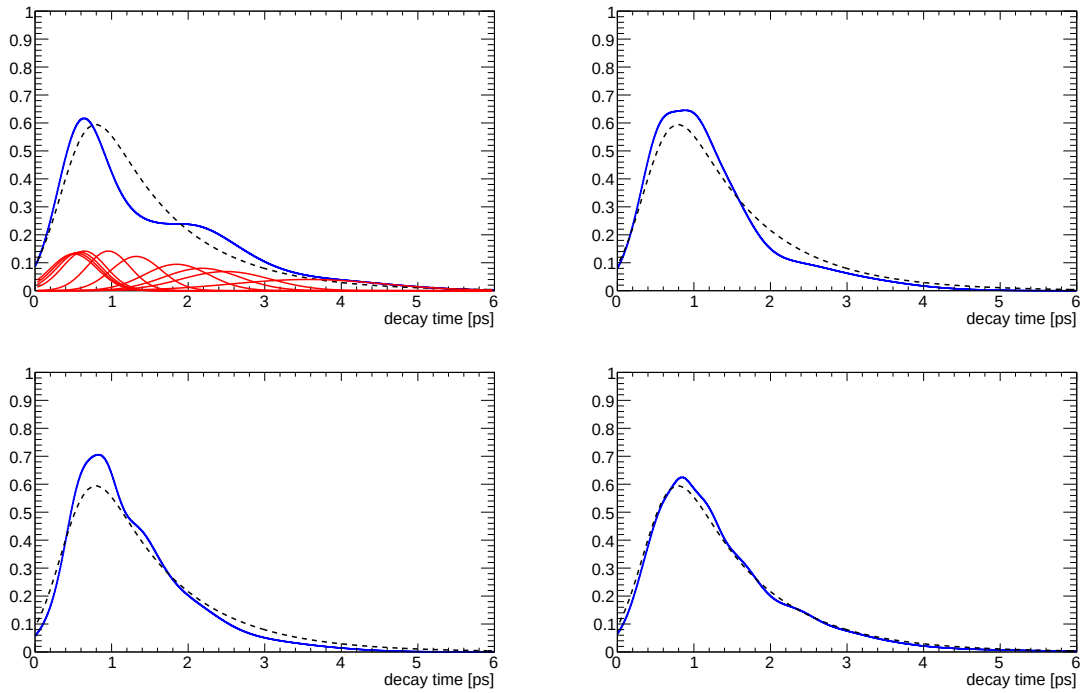


Figure 4.9: Illustration of how the kernel method reproduces the decay time distribution using 10 (top left), 50 (top right), 200 (bottom left) and 1000 (bottom right) events. The 'true' distribution is shown as the dotted black line is an exponential distribution multiplied with an acceptance function, modelled with an error function. The blue curve is the kernel estimate of the distribution and the red curves in the top left plot are the individual Gaussian kernels. The resolution parameter in Equation 4.24 is set to 100 fs [89].

The *kernel method* works over a large range of statistics for most distributions because of this adaptive smearing. Figure 4.9 shows for illustration how a decay time distribution is estimated using 10, 50, 200 and 1000 events. The decay time distribution is modelled as an exponential function that is then multiplied with an acceptance function, which is described by an error function.

When using sWeights, calculated in the 2011 analysis, the estimated distribution has to be positive since it is describing a probability density function. As the *sWeights* can be both positive and negative there is no guarantee that it will be positive for all values of the control variable. This problem is resolved with the aid of a test distribution estimated by a different method that guarantees it to be positive. This test distribution is made in the same way as the *sWeighted kernel* distribution, but using the Bayesian signal probability shown in Equation 4.10 as a weight instead of the *sWeight*.

If the signed ratio of the *sWeighted kernel* distribution and the test distribution is smaller than some value ϵ, then the ratio is mapped from the original region $[-\infty, \epsilon]$ to the region $[0, \epsilon]$. An exponential function is used for this mapping. The corrected distribution is assigned the

value of the re-mapped ratio times the value of the test distribution at these values of the control variable.
The value of the ϵ parameter is chosen to be 10^{-3} and a systematic uncertainty is assigned to the sensitivity of the exact value. Evaluation of this systematic for each analysis is detailed in Chapters 5 and 6.

4.5.3 Determination of the Average Acceptance Function

The decay time distributions estimated with the method described in Section 4.5.2 are biased by the trigger, stripping and offline event selections as discussed in prior sections. The fitter uses the per-event acceptance information for all signal classes, as described in Section 4.4. This requires an unbiased decay time distribution (f_{ub}) for the background classes that are described by the non-parametric biased distributions. The conditional PDF for these classes, corresponding to Equation 4.12 is

$$f(t|A, class) = \frac{f_{ub}(t|class)A(t)}{\int_0^\infty f_{ub}(t'|class)A(t')dt'}, \tag{4.25}$$

where $A(t)$ is a series of top-hat functions describing the per-event acceptance.
The unbiased distribution is determined by dividing the observed distribution from Section 4.5.2 with the average acceptance function for that signal class. The average acceptance is determined from the turning points calculated by the swimming, and is derived as follows. If $A_i(t)$ is the series of top hat functions describing the per-event acceptance for event i, then the observed distribution is

$$f_{obs}(t) = \frac{1}{N}\sum_{i\in Evts} f(t|A_i) \tag{4.26}$$

in the limit of infinite statistics. Expressing this in terms of the unbiased distribution gives

$$f_{obs}(t) = \frac{1}{N}\sum_{i\in Evts} \frac{f_{ub}(t)\cdot A_i(t)}{\int_0^\infty f_{ub}(t')\cdot A_i(t')dt'}. \tag{4.27}$$

The data sample consists of several classes with the Bayesian probability or *sWeights* used to classify the events. Hence the unbiased distribution for one particular class can be expressed as

$$f_{obs}(t|class) = \frac{1}{N}\sum_{i\in Evts} \frac{A_i(t)\cdot W_i^{class}}{\int_0^\infty f_{ub}(t'|class)\cdot A_i(t')dt'}\cdot f_{ub}(t|class), \tag{4.28}$$

where W_i^{class} is either the Bayesian probability or *sWeight* for event i for that class. By definition, the relation between the observed and unbiased distributions is

$$f_{obs}(t|class) = f_A(t|class) \cdot f_{ub}(t|class), \tag{4.29}$$

where $f_A(t|class)$ is the average acceptance function for that class. Hence the average acceptance function can be identified as

$$f_A(t|class) = \frac{1}{N} \sum_{i \in Evts} \frac{A_i(t) \cdot W_i^{class}}{\int_0^\infty f_{ub}(t'|class) \cdot A_i(t') dt'}. \tag{4.30}$$

As can be seen in Equation 4.30, the unbiased distribution ($f_{ub}(t|class)$) is required to determine the average acceptance. However, the reason why we need the average acceptance function is to turn the observed non-parametric proper time distributions into unbiased distributions for the background classes.
The method for determining a first estimate of the average acceptance function for the backgrounds is different for the 2010 and 2011 lifetime analyses.

2010 method

The definition of the average acceptance function is given by Equation 4.30. The analytical unbiased lifetime PDF is known for the signal and therefore used in the integral. However, it is a priori not known for the combinatorial background and so needs to be 'boot-strapped' using an initial approximation. An iterative method is employed where a single exponential is used in the calculation of the first acceptance function. The background time PDF is then determined as described in Section 4.5.2. A new acceptance function is calculated from this PDF which is used in the next step. The iteration stops when the difference between two determined background PDFs is smaller than a pre-defined value.

2011 method

The process is 'boot-strapped' initially by determining the average acceptance functions for one of the signal classes modelled with an exponential proper time distribution. This is then used as the first approximation to the average acceptance function for the non-parametric signal classes, to determine a first approximate unbiased distribution. This is then used in Equation 4.30 to determine a second approximation of the average acceptance function. The process is then repeated until no further change in the average acceptance function is observed.
An additional complication is that the use of *sWeights* can result in a locally negative average acceptance function. This is handled by simply setting the value of the average acceptance

function to zero in the regions where it is smaller than some value ϵ. The value of this parameter is the same as the ϵ parameter in the correction of negative PDF values in Section 4.5.2. The sensitivity to the exact value of this parameter is evaluated in Chapters 5 and 6.

4.5.4 Total Likelihood Function in the Lifetime Fit

The total likelihood function used in the lifetime fit is formed by combining the functions defined in the previous sections.

$$f(t, A|m) = \sum_{classes} f(t|A, class) \cdot f(A|class) \cdot \frac{P(class) f(m|class)}{f(m)} \tag{4.31}$$

The first factor comes from Equation 4.25 and the last factor comes from Equation 4.10. For the 2010 analysis we used a turning point distribution, $f(A|class)$, for each signal class weighted by the Bayesian weights detailed in Section 4.3.3.1. In the 2011 analysis, the turning point distribution is determined using the sWeights, Section 4.3.3.2.

4.6 Implementation of the Fit Method

The mass and lifetime fit are both implemented as an unbinned maximum likelihood fit. Their implementation is based on the MINT package [90] which provides an interface to MINUIT [91]. The fit to the invariant mass spectrum is set up such that the defined PDF mass models are fed into the maximum likelihood fit. The parameters for each PDF, and the relative fractions of each signal class, are then optimised for the observed data. The signal class weightings per-event per-signal class are then calculated based on the chosen method from Section 4.3.3.1 or 4.3.3.2. The fitter is set up such that the results of the mass fit are automatically propagated to the lifetime fit. The fit to the reconstructed lifetime spectrum begins by feeding in analytical PDF lifetime models for the main signal classes, with misreconstructed and combinatorial lifetime distributions determined using methods detailed in Section 4.5.2. As the evaluation of Gaussian kernels to describe distributions is computationally intensive, distributions are evaluated once whenever possible and stored as high resolution histograms to ensure minimal losses in accuracy. The fitter produces an output file with plots that serve as visual cross-checks of the fit convergence.

4.7 Blinding Procedure

In both the 2010 and 2011 lifetime analyses the measurements were performed as a blinded analysis, albeit using different methods. The 2010 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analysis only performed a measurement of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime. The fitter was verified on the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ control channel, prior to permission being granted to perform the full analysis on $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$. For the 2011 analysis we measure three lifetimes, $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$, $\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-}$ and $\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-}$, which are all required to be blinded. For this analysis, a fit parameter is implemented that is scaled internally with a random factor seeded with known input parameters. This random factor is sampled from a rectangular distribution between 0.9 and 1.1 and it scales both the fitted value and the uncertainty. The seeds for the random scaling of the three measured lifetimes, $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$, $\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-}$ and $\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-}$ are different. This allows the fitter to be run on the data to ensure its functionality is correct without biasing the result. Both unblinding of datasets were conducted once the analysis was fully verified on simplified and full simulations, and after agreement from the referees.

4.8 Summary

This chapter has presented the details of a method to determine the effective lifetime of a B decay, whilst removing the acceptance bias. An introduction is provided in Section 4.1 that discusses a general overview of lifetime measurement methods and how biases can enter into such analyses. The general fit method is then detailed in Section 4.2, which shows the full fit can be factorised into two independent fits, the first to the invariant mass and the second to the reconstructed lifetime spectrum. Each stage is then discussed in detail with respect to the two analyses, 2010 and 2011, that they are applied to.

The mass fit stage, given in Section 4.3.1, has three main purposes. The first is to determine the parameters of the PDF models, with the second being to determine the relative signal fractions in the dataset. These first two stages are performed simultaneously. The third and final purpose is to determine event-by-event weightings using whichever method is required for the analysis, Section 4.3.3. Prior to the lifetime fit, each event must have their per-event acceptance function calculated. These are used to correct for the lifetime biasing selection criteria that occur in the trigger and offline selection, and are determined via the novel 'swimming' method that is detailed in Section 4.4. The lifetime fit, Section 4.5, is able to differentiate the contributing signal channels and extract the effective lifetime by using the kernel method to fit for non-parametric backgrounds that occur in the data. Finally, a blinding method is described in Section 4.7, to ensure no user bias.

Chapter 5

$\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ Lifetime Data Selection and Verification

5.1 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ Lifetime Analyses

The lifetime measurements performed in this thesis will use the statistically independent 2010 and 2011 LHCb datasets. The 2010 dataset comprises a total integrated luminosity of $37\,\mathrm{pb}^{-1}$, with the complete 2011 data being much larger at a total integrated luminosity of $1.0\,\mathrm{fb}^{-1}$. Due to the size of the dataset collected during the 2010 run and the lifetime measurement method used, the only lifetime measured is the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$. The measurement of the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ lifetime acts as a control channel for the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime measurement.

The onset of the 2011 data taking that followed the 2010, provides a dataset that allows for a much higher precision measurement of the lifetimes to be made. The additional statistics, combined with improvements to the method detailed in Chapter 4, give the potential to make two more lifetime measurements, these are the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$.

The analysis of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ decay mode in particular provides constraints on parameters in the SM and probes for NP effects, this was discussed previously in Section 1.4. The details of the two analyses will be split over two chapters. The first part, detailed in this chapter, will describe the current experimental status (Section 5.2), the data sample selection (Section 5.3), the mass and lifetime models used (Sections 5.4 and 5.5) then finally verification studies of the fit method (Section 5.6). The second part detailing event yields, systematic studies and results, will be set out in Chapter 6.

5.2 Experimental Status

The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ family of decays were first observed by the ARGUS Collaboration [92], with the various final states not able to be distinguished until later studies. The $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ channel, observed by the CLEO collaboration [93], was the first final state decay from this family to be distinguished. This was followed by observation of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, together with evidence for the $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$, by the CDF Collaboration [94] that was subsequently confirmed by the BELLE [95] Collaboration.
The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays have been studied in detail by LHCb with measurements of the branching fractions [96], and time integrated [97] and time dependent [98] CP violation.
LHCb has published a measurement of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime using the 2010 dataset [76] using two independent methods, one of which is using the method described in Chapter 4 and given as a result in Chapter 6.
LHCb has also published a measurement with the full 2011 dataset using a method that introduces a minimal lifetime bias in the selection [99], which yields the lifetime

$$\tau_{K^+K^-} = 1.455 \pm 0.046 \text{ (stat)} \pm 0.006 \text{ (syst) ps}. \tag{5.1}$$

Currently this is the most precise measurement of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime, although the aim of the analysis presented in this thesis is to improve the precision on this measurement.
The current world average of the B^0_d lifetime is [36]

$$\tau_{\mathrm{B}} = 1.519 \pm 0.007\ ps, \tag{5.2}$$

and the current world average of the flavour specific B^0_s lifetime is [36]

$$\tau_{\mathrm{B}^0_s} = 1.466 \pm 0.031\ ps. \tag{5.3}$$

5.3 Data Sample

Each analysis is conducted using the complete dataset for that data taking period. The data is collected with both the magnet up and magnet down polarities, with the VELO in its fully inserted position. The energy per beam for both the 2010 and 2011 data taking periods is $3.5\,\mathrm{TeV}$. The trigger, stripping and offline selections differ between analyses, with the details of each given in Sections 5.3.1, 5.3.2 and 5.3.3 respectively. The b candidate events are selected via separation of the primary and secondary vertices and through cuts on their Impact Parameters. The drawback of applying these criteria is the preferential rejection of shorter lifetime candidates, such that the decay time distribution becomes biased as discussed

in Section 4.1 and shown in Figure 5.1. Before the analysis is performed, each dataset is processed to determine their per-event acceptance functions as detailed in Section 4.4.

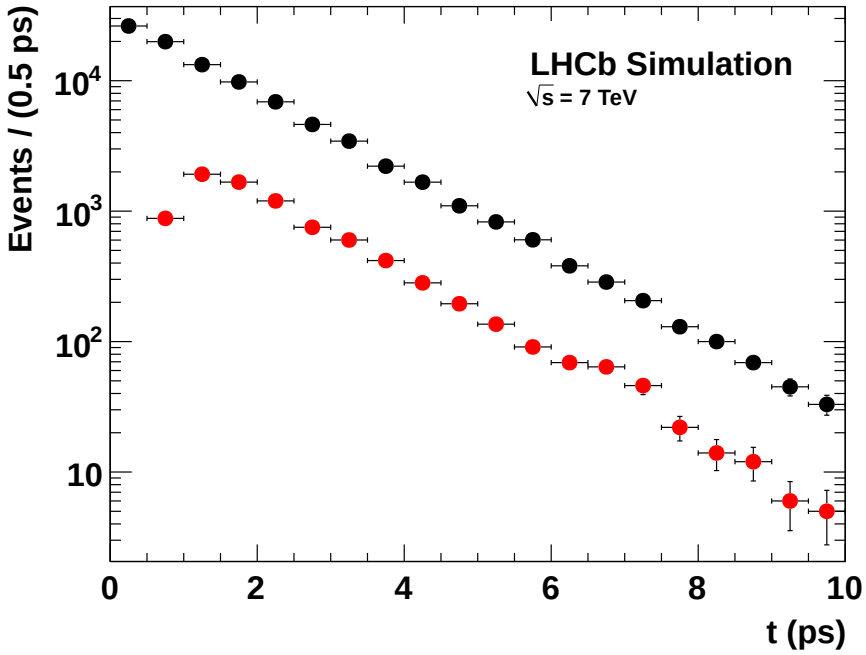


Figure 5.1: The decay time (t) distribution for true reconstructed $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ events from full 2010 LHCb Monte Carlo before (black) and after (red) the full selection has been applied [75, 89].

5.3.1 Trigger

As discussed previously in Section 2.10, the triggers are composed of three different levels: a lower level hardware trigger (L0) and two higher level software triggers (HLT1 and HLT2). During the data taking periods the luminosity of the LHC was increased, where this was particularly noticeable in the initial run during 2010. To optimise the data taking, the HLT trigger lines need to be varied with respect to this increase. These changes are recorded as different Trigger Configuration Keys or 'TCK'. The efficiency of the higher level triggers relies on the precision vertexing capabilities of the VELO. Triggered events whose tracks do not originate from the $\mathrm{B}_{\mathrm{d/s}}$ candidate are known as 'Triggered Independent of Signal' (TIS), whereas events that do are known as 'Triggered On Signal' (TOS). The TIS events are not subject to the proper time bias mentioned previously.

The measurements discussed in this chapter, and the next, apply different trigger conditions to select candidates for the analyses. The trigger names, provided in Tables 5.1 and Table 5.2, correspond to specific trigger algorithms that yield an independent trigger decision. A brief

description of the triggers used in the analyses is now provided.
The conditions for any Global trigger, L0, HLT1 and HLT2, is dependent on if the event is selected by any other trigger line within that stage. The L0Hadron trigger selects heavy flavour decays with hadronic final states. The HLT1TrackAllL0 trigger is executed on all trigger lines accepted by the L0 stage, and is designed to select hadronic decays that are significantly displaced from a PV. The final trigger line used, the HLT2B2HHDecision, is purposely designed to select two body hadronic B-decays. These triggers are separated depending on if they are TIS or TOS type events. The trigger criteria for both analyses are provided in Table 5.1 and Table 5.2.

Cut Type	Cut
HLT Cuts	HLT1Global_TIS=1 *and* HLT2Global_TIS=1 *or* HLT1Global_TIS=1 *and* HLT2B2HHDecision_TOS=1 *or* HLT1Global_T0S=1 *and* HLT2B2HHDecision_TOS=1

Table 5.1: Trigger conditions for 2010 $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime analysis.

Cut Type	Cut
HLT Cuts	L0Global_TIS=1 *or* L0Hadron_TOS=1 *and* HLT1TrackAllL0_TOS=1 *and* HLT2B2HHDecision_TOS=1

Table 5.2: Trigger conditions for 2011 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analyses.

5.3.2 Stripping

The triggering is only the first step in selecting potential signal candidates. These data are further processed to produce smaller, cleaner datasets for offline analyses with the process known as "stripping". The method involves imposing a set of optimised cuts on the trigger selected data to maximise the signal efficiency of interest. These are organised into a set of 'stripping lines' that pre-select a certain family of decays. In this analysis a line developed specifically to select two-body hadronic decays is used, with the cuts for both the 2010 and 2011 analyses detailed in Tables 5.3 and 5.4 respectively. The $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ candidates selected are reconstructed under the assumption that both daughter particles are π's. There are no requirements placed on the trigger lines in this selection. As with the trigger criteria,

the stripping cuts bias the decay time distribution of the selected candidates.
A brief description of the variables used in the stripping, and subsequent offline selections, is now provided in order to decipher the cuts. The variables m, τ and p_T have the usual meanings of invariant mass, decay time and transverse momentum respectively. The Impact Parameter (IP), which is defined as the distance of closest approach of a track to a certain point in space, i.e. the PV, is given by the variable IP with the χ^2 fit of this parameter given by $IP\chi^2$. The $\mathrm{Track}\ \chi^2/\mathrm{nDOF}$ is described as the χ^2 per degree of freedom of the track fit, and FD is the flight distance of the mother particle. Finally, the DOCA variable is defined as the distance of closest approach between all possible particle pairs, with the DIRA variable being described as the cosine of the angle between the momentum of the particle and the direction of flight from the best PV to decay vertex.

Cut Type	Cut
Daughter Cuts	$\min(p_T^{h^+}, p_T^{h'^-}) > 700\ \mathrm{MeV}/c$
	$\max(p_T^{h^+}, p_T^{h'^-}) > 2400\ \mathrm{MeV}/c$
	$\min(IP^{h^+}, IP^{h'^-}) > 0.08$ mm
	$\max(IP^{h^+}, IP^{h'^-}) > 0.2$ mm
	$\max((\mathrm{Track}\ \chi^2/\mathrm{nDOF})^{\mathrm{h}^+}, (\mathrm{Track}\ \chi^2/\mathrm{nDOF})^{\mathrm{h}'^-}) < 5$
Combination Cuts	4500 MeV/c^2 $< m_{comb} <$ 6500 MeV/c^2
	DOCA < 0.08 mm
Mother Cuts	4800 MeV/c^2 $< m_B <$ 5800 MeV/c^2
	$IP^B < 0.06$ mm
	$FD^B > 1.8$ mm

Table 5.3: Stripping 2010 Requirements for $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ candidates.

5.3.3 Offline Selection

An offline selection is required in addition to the previous two stages of data reduction, as the dedicated stripping line is designed to select all $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ type candidates. The offline selection is designed to specifically select $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ channels and maximise their significance with respect to the combinatorial background. The selection is designed using a set of kinematic cuts that are optimised using a multivariate analysis technique, introduced further in Sections 5.3.3.1 and 5.3.3.1, then followed by channel specific PID cuts for the separation of the KK and $K\pi$ final state candidates. The kinematic and PID selections between the 2010 and 2011 analyses are optimised independently and will be discussed

further in the following subsections.

Cut Type	Cut
Daughter Cuts	$\min(p_T^{h^+}, p_T^{h'^-}) > 1100\ \mathrm{MeV}/c$ $\min(IP^{h^+}, IP^{h'^-}) > 0.12$ mm OR $\min(IP\chi^{h^+}, IP\chi^{h'^-}) > 100$ $\max((\mathrm{Track}\ \chi^2/\mathrm{nDOF})^{\mathrm{h}^+}, (\mathrm{Track}\ \chi^2/\mathrm{nDOF})^{\mathrm{h}'^-}) < 5$
Combination Cuts	4800 MeV/c^2 $< m_{comb} <$ 5800 MeV/c^2 $\max(IP^{h^+}, IP^{h'^-}) > 0.25$ mm OR $\max(IP\chi^{h^+}, IP\chi^{h'^-}) > 200$ $\max(p_T > 2500\ \mathrm{MeV}/c)$ DOCA < 0.08 mm
Mother Cuts	4800 MeV/c^2 $< m_B <$ 5800 MeV/c^2 $p_T^B > 1200\ \mathrm{MeV}/c$ $\tau_B > 0.6$ ps $IP^B < 0.1$ mm OR $IP\chi^B < 12$

Table 5.4: Stripping 2011 Requirements for $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ candidates.

5.3.3.1 Kinematic Selection

The offline kinematic selection is designed to maximise the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal channel efficiencies compared with the combinatorial background. The kinematic selection comprises a series of simple variable cuts which are applied to the candidates in order to achieve this. Descriptions of these variables are given in Section 5.3.2. The difference in size between the 2010 and 2011 datasets, coupled with improvements in performance of the reconstruction software and PID performance that occurred between analyses, meant the offline selections for each analysis were optimised independently.

Kinematic Selection for 2010 dataset

The 2010 kinematic selection was designed to run on the relatively small $37\,\mathrm{pb}^{-1}$ of data collected during the data taking run. The cuts were optimised by using the *NeuroBayes* package [100] on $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ MC for the signal, and sideband data for the background. The final set of kinematic cuts used in the 2010 analysis are given in Table 5.5.

The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ invariant mass distribution from data both before and after the kinematic cuts are displayed in Figure 5.2.

An additional complication arises in this analysis via the 'swimming' method that is applied to all selected candidates, this is described in detail in Section 4.4. Unfortunately some of

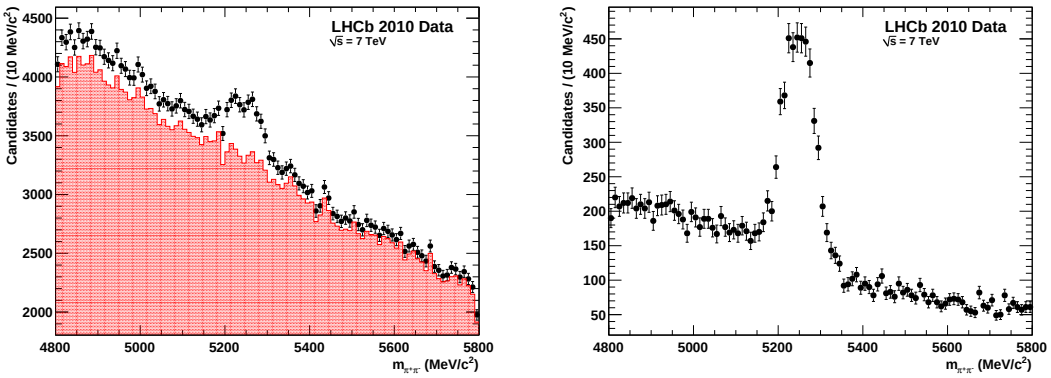
Cut Type	Cut
Daughter Cuts	$\min(p_T^{h^+}, p_T^{h'^-}) > 700$ MeV/c
	$\max(p_T^{h^+}, p_T^{h'^-}) > 2450$ MeV/c
	$\min(p^{h^+}, p^{h'^-}) > 13500$ MeV/c
	$\min(IP\ \chi^2_{h^+}, IP\ \chi^2_{h'^-}) > 80$
	$\max(IP\ \chi^2_{h^+}, IP\ \chi^2_{h'^-}) > 140$
	$\min(IP^{h^+}, IP^{h'^-}) > 0.11$ mm
	$\max(IP^{h^+}, IP^{h'^-}) > 0.25$ mm
	$\min((Track\chi^2/nDOF)^{h^+}, (Track\chi^2/nDOF)^{h'^-}) < 3$
	$\max((Track\chi^2/nDOF)^{h^+}, (Track\chi^2/nDOF)^{h'^-}) < 4$
Mother Cuts	$p_T^{B^0} > 900$ MeV/c
	$FD^{B^0} > 2.4$ mm
	$FD\ \chi^2_{B^0} > 50$
	$IP_{B^0} < 0.09$ mm
	$IP\ \chi^2_{B^0} < 16$

Table 5.5: Offline $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ selection cuts used for the 2010 lifetime measurement.

the 2010 data processed through early versions of the trigger software, are unavailable to the swimming code. The result of this, is that per-event acceptance functions could not be determined for the trigger selection for all 2010 data. The solution to this problem is that the offline selection cuts are altered, such that they cut harder on all decay time biasing variables than any of the trigger conditions used during the 2010 data taking period. It is then possible to swim only the offline selection, having assured that the trigger does not introduce any additional decay time bias.

Kinematic Selection for 2011 dataset

The re-optimisation of the kinematic cuts for the 2011 dataset is done via the ROOT package TMVA [101]. TMVA provides the ability to test multiple methods via MultiVariate Analysis (MVA) techniques, to determine an optimum set of cuts for the data. All MVA methods were trained so as to evaluate their performances. The training data used consisted of a 'signal' and 'background' dataset. The signal dataset was produced using the full LHCb $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ MC sample, which was passed through the stripping line cuts, provided in Table 5.4, to simulate the variable distributions seen in data (Figure 5.3). The background dataset comprised of a sub-sample of the full 2011 stripped data that was passed through



(a) Before offline selection.

(b) After offline selection.

Figure 5.2: The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ mass spectrum of the 2010 data under the $\pi^+\pi^-$ decay hypothesis. Before (left) and after (right) the offline kinematic selection. Candidates rejected by the offline selection are shown on the left in red [75].

the $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ specific stripping line, discussed in section 5.3.2, outside the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ signal peak range ($4800 < m_B < 5150$ and $5500 < m_B < 5800$).

The performances of all the MVA trained methods were similar, with a modest increase in the signal significance observed for the multivariate techniques compared to the cut-based selections. Due to the technical difficulties to determine the per-event acceptance functions for a multivariate selection, it was decided to implement the most optimal rectangular cuts based method instead. The output of this is a set of simple kinematic cuts, where the cuts are listed in Table 5.6. The re-optimisation of the selection increases the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal yield by $\sim 4\%$, compared with the 2010 selection. The resulting $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ mass spectrum is shown in Figure 5.4.

Cut Type	Cut
Daughter Cuts	$\min(p_T^{h^+}, p_T^{h'^-}) > 1100$ MeV/c
	$\max(p_T^{h^+}, p_T^{h'^-}) > 2500$ MeV/c
	$\min(IP\chi^{2\ h^+}, IP\chi^{2\ h'^-}) < 45$
	$\max(IP\chi^{2\ h^+}, IP\chi^{2\ h'^-}) < 70$
	$\max((Track\chi^2/nDOF)^{h^+}, (Track\chi^2/nDOF)^{h'^-}) < 3.3$
Mother Cuts	$IP\ \chi^2_{B^0} < 9$
	$DIRA\ Own\ PV_{B^0} > 0.999818$

Table 5.6: Offline $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ kinematic selection cuts used for the 2011 lifetime measurement.

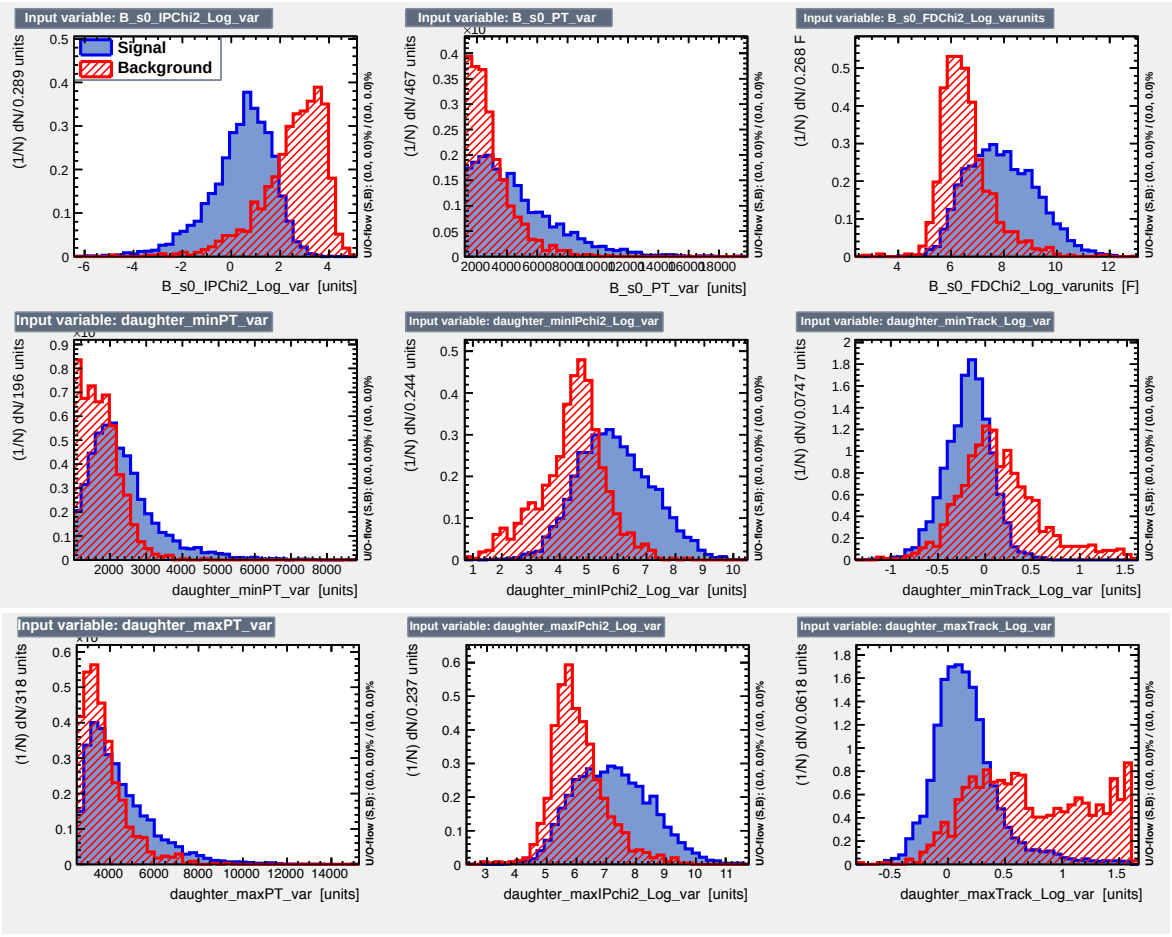


Figure 5.3: Signal (blue) and background (red) kinematic variables used in TMVA training.

5.3.3.2 Particle Identification Selection

After the offline kinematic selection is applied, the selected $B \to h^+h^-$ dataset is split into the different final states (KK, $K\pi$) using the particle identification abilities of the experiment. The relatively clean separation of the final states is made possible in LHCb through the use of the information from the two RICH detectors. In particular, we make use of the Difference in Log Likelihood (DLL) PID quantities $DLL_{K\pi}$, $DLL_{p\pi}$, $DLL_{e\pi}$ and $DLL_{\mu\pi}$ for final state pions and the quantities $DLL_{K\pi}$, DLL_{Kp}, DLL_{Ke} and $DLL_{K\mu}$ for final state kaons. The DLL distributions are derived from the difference in Log-likelihood distributions between the relevant particles.

Since PID distributions in LHCb Monte Carlo simulations do not accurately match those in data, a data-driven approach is applied in tuning the PID selections. However, a small contamination from specific final state decays is expected to remain.

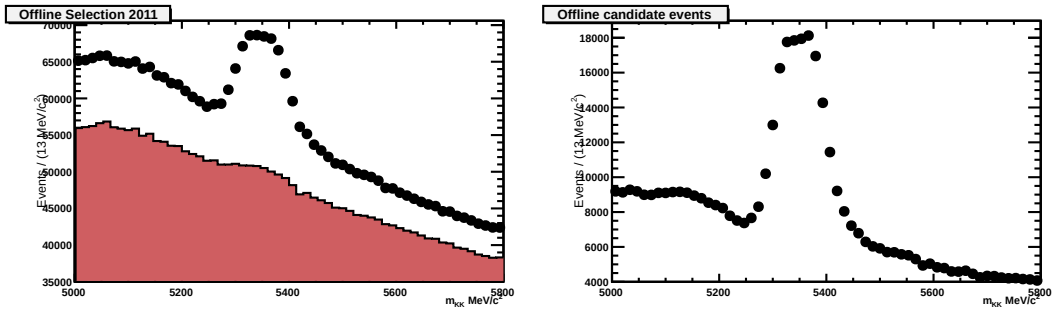


Figure 5.4: The $\mathrm{B} \to \mathrm{h^+h^-}$ mass spectrum of the 2011 data under the K^+K^- decay hypothesis. Before (left) and after (right) the offline kinematic selection. Candidates rejected by the offline selection are shown on the left in red.

KK and $\mathrm{K\pi}$ final state selection for 2010 dataset

The optimisation of the 2010 PID cuts for the KK final state involves fitting a Gaussian peak to the $\mathrm{B}_s^0 \to \mathrm{K^+K^-}$ mass peak. This is done to determine the amount of cross contamination in the sample. The presence of the largest contributing contamination background, $\mathrm{B}_d^0 \to \mathrm{K^+\pi^-}$, will tend to pull the fitted mean of the Gaussian lower owing to the lower invariant mass of the B_d^0 meson. By placing stronger requirements on the likelihood that both daughters are kaons, the fitted mass of the B_s^0 meson is seen to increase. Above a minimum PID cut of 7 the B_s^0 mass no longer increases, thus a cut at 8 is imposed that is well within the plateau region. In addition to the $K - \pi$ separation, cuts on the $K - p$, $K - e$ and $K - \mu$ variables are also made using the same method to further reject background. The kaon PID criteria, which are applied to both final state particles in the KK final state, are given in Table 5.7.

The $\mathrm{K\pi}$ final state optimisation is treated differently to the KK final state, due to the contribution to the signal from a secondary $\mathrm{K\pi}$ peak, the $\mathrm{B}_s^0 \to \mathrm{\pi^+K^-}$. For $\mathrm{K\pi}$, the metric chosen was the ratio of the yields of the decay modes $\mathrm{B}_s^0 \to \mathrm{\pi^+K^-}$ and $\mathrm{B}_d^0 \to \mathrm{K^+\pi^-}$. The $K\pi$ mass spectrum was fitted with two Gaussians to model these two signals. In the event of substantial contamination from either $\mathrm{B}_s^0 \to \mathrm{K^+K^-}$ or $\mathrm{B}_d^0 \to \mathrm{\pi^+\pi^-}$, an over-estimation of the amount of $\mathrm{B}_s^0 \to \mathrm{\pi^+K^-}$ signal would occur since these modes have a mean mass close to that of the $\mathrm{B}_s^0 \to \mathrm{\pi^+K^-}$ peak. One of the final state particles is required to satisfy the kaon PID conditions determined from the previous KK optimisation, which are provided in Table 5.7. The ratio stabilises when $\min(\mathrm{DLL}_{\mathrm{K\pi}}^{h^+}, \mathrm{DLL}_{\mathrm{K\pi}}^{h'^-}) < -5$. The full list of cuts for the π, to be used in parallel with the K cuts for the $\mathrm{K\pi}$ final state, are given in Table 5.7.

Particle	PID Cuts
K	$\mathrm{DLL}_{\mathrm{K}\pi} > +8$ & $\mathrm{DLL}_{\mathrm{K}p} > -5$ & $\mathrm{DLL}_{\mathrm{K}e} > 0$ & $\mathrm{DLL}_{\mathrm{K}\mu} > 0$
π	$\mathrm{DLL}_{\mathrm{K}\pi} < -5$ & $\mathrm{DLL}_{p\pi} < 0$ & $\mathrm{DLL}_{e\pi} < 0$ & $\mathrm{DLL}_{\mu\pi} < 0$

Table 5.7: PID criteria for K and π final state particles for 2010 data. These apply to both the KK and Kπ final states.

KK and Kπ final state selection for 2011 dataset

The re-optimisation of the PID values for the 2011 data uses an analogous method to that used for the 2010 data. The mean, sigma and peak significance of a Gaussian fitted to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ invariant mass peak are used as a metric to determine the amount of cross-contamination. As before, the presence of $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ contamination will tend to pull the mean of this Gaussian to lower values since the B^0_d meson has a smaller rest mass than the B^0_s meson. This contamination will also cause the sigma of the peak to be larger than it would be with just the KK final state present, due to the extra $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ modes broadening the peak. To maximise the signal yield from the data, the significance quantity is introduced. This is given by

$$\text{significance} = \frac{S}{\sqrt{S+B}}, \tag{5.4}$$

where S is the estimated signal yield and B the estimated background yield.
An iterative procedure is employed, where all PID cuts are fixed except the one being varied. The free PID variable is then iteratively changed between its minimum and maximum values, and a fit to the resulting invariant mass spectrum carried out. In Figure 5.5 the fitted mass of the B^0_s meson is seen to increase, until it becomes relatively stable as stronger requirements are made on the likelihood that both daughters are kaons. Once the mean becomes stable, the sigma distribution can be used to determine when the peak contains a high percentage of KK final state events. The significance is then used to optimise the yield. The significance is highly correlated with signal yield and the percentage of signal and background in the peak. Above a minimum PID cut of 3 the B^0_s mass no longer increases, however the sigma of the peak is still converging to its optimum value which occurs at ≈ 5. A cut at 5 is thus imposed, which is within the plateau region.
In addition to the $K-\pi$ separation cut, selections on the $K-p$, $K-e$ and $K-\mu$ variables are also made to further reject background. The kaon PID criteria, which are applied to both final state particles in the KK final state, are given in Table 5.8.
A similar technique was used to design the Kπ final state selection, with the same metrics

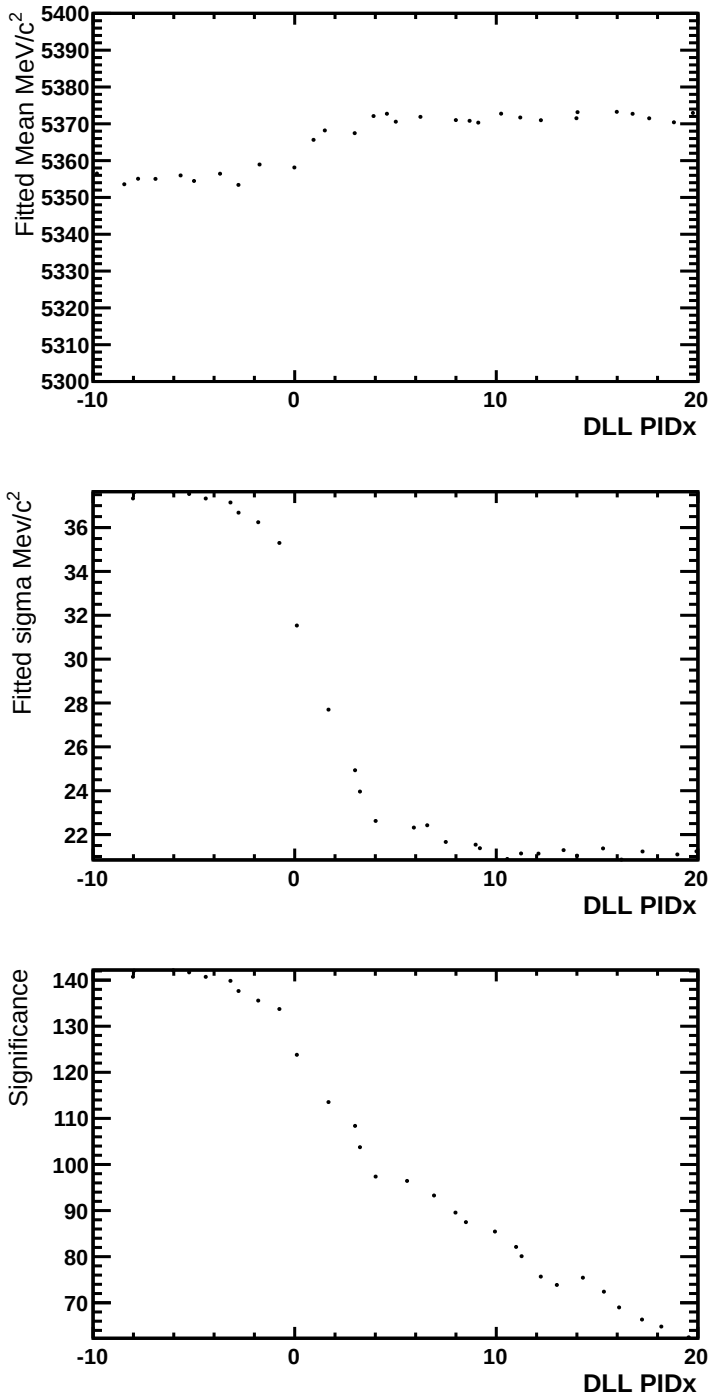


Figure 5.5: The fitted mass (top), sigma (middle) and significance (bottom) of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ peak as a function of a cut on min($\mathrm{DLL}^{h^+}_{\mathrm{K}\pi}, \mathrm{DLL}^{h'^-}_{\mathrm{K}\pi}$) using 2011 stripped data. The units DLLx represent the cut on min($\mathrm{DLL}^{h^+}_{\mathrm{K}\pi}, \mathrm{DLL}^{h'^-}_{\mathrm{K}\pi}$), where x can take the value of $x = \pi, K, p, \mu$ dependent on the PID variable studied.

chosen. The $\mathrm{K}\pi$ mass spectrum was fitted with a single Gaussian to model the B^0_d signal, with the sigma and significance checked to optimise the corresponding cuts. To simplify the iteration procedure, the re-optimised kaon cuts for the KK final state were fixed, and only the pion specific PID cuts varied. The full list of cuts for the π, to be used in parallel with the K cuts for the $\mathrm{K}\pi$ final state, are given in Table 5.8.

Particle	PID Cuts
K	$\mathrm{DLL}_{\mathrm{K}\pi} > +5$ & $\mathrm{DLL}_{\mathrm{K}p} > -4$ & $\mathrm{DLL}_{\mathrm{K}e} > +5$ & $\mathrm{DLL}_{\mathrm{K}\mu} > +5$
π	$\mathrm{DLL}_{\mathrm{K}\pi} < -4$ & $\mathrm{DLL}_{p\pi} < +3$ & $\mathrm{DLL}_{e\pi} < 0$ & $\mathrm{DLL}_{\mu\pi} < 0$

Table 5.8: PID criteria for K and π final state particles for 2011 data. These apply to both the KK and $\mathrm{K}\pi$ final states.

5.4 Mass Model Descriptions

The mass models used for the 2010 and 2011 analyses are important as they affect the calculated signal fractions and per-event weightings, detailed in Section 4.3.3. The 2010 and 2011 models differ primarily due to the statistics of each dataset available for the measurement, with the value of x in each function representing the invariant mass m. These will be discussed individually as follows.

5.4.1 Mass Models 2010

The limited statistics of the 2010 dataset allow some approximations for the models of the primary signal classes to be used. It is also expected [76] [75] that the dominant backgrounds for the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ measurement will be the specific $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ decay mode and the combinatorial background, while the primary background for the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ is combinatorial. The mass models used for the analysis of the 2010 data are described as follows.

5.4.1.1 Signal Model

For the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal modes, $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ in the KK and $\mathrm{K}\pi$ spectra respectively, a Gaussian model can be used as the approximate shape for the invariant mass distributions. This follows the standard definition

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{5.5}$$

where μ is the signal mean and σ is the width of the signal peak.

5.4.1.2 Misidentified Background Model

Misidentified decay channels are not expected to contribute heavily in this analysis. Thus no models are included.

5.4.1.3 Partially Reconstructed Background Model

The mass ranges selected for the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $5176 - 5800\,\mathrm{MeV}/c^2$ and $5272 - 5800\,\mathrm{MeV}/c^2$ respectively, are chosen to exclude the partially reconstructed background. Thus no models are included in this analysis.

5.4.1.4 Combinatorial Background Model

The combinatorial is fitted using a linear function given by

$$f(x) = ax + b \tag{5.6}$$

where a is the gradient of the combinatorial background and b is the intercept constant of the function.

5.4.2 Mass Models 2011

The increase in dataset size accumulated during the 2011 data taking run, results in the need to improve the definitions of the signal mass models and the addition of extra background models that become increasingly dominant with the increased event count. The two types of additional background incorporated into the 2011 analysis are first the partially reconstructed background, and second the misidentified signal channels. Partially reconstructed events consist of multi body events that are mis-reconstructed as two body decays. The mass distribution of these B^0_d and B^0_s decays lie predominantly in the mass range below the B^0_d and B^0_s mass peaks, however there is a small overlap which needs to be accounted for. The misidentified backgrounds are due to a misidentification of one, or both, of the final state particles due to PID inefficiencies. These backgrounds tend to overlap heavily with the main signal channel.

5.4.2.1 Signal Model

The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal classes, $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$, modelled with Gaussian distributions in the 2010 analysis are redefined using a Double Crystal Ball (DCB) function. Additionally, the $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ is also modelled using a DCB for this analysis. This function

is comprised of two separate single Crystal Ball (CB) functions [102], where there is a fractional component of each CB in the total PDF. One of the CB functions has a low mass tail component and the other a high mass tail component. The general expression for a single CB function is given by

$$f(x;\alpha,n,\mu,\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma^2}), & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\mu}{\sigma})^{-n} & \text{for } \frac{x-\mu}{\sigma} \leq -\alpha \end{cases} \tag{5.7}$$

where

$$\begin{aligned} A &= (\tfrac{n}{|\alpha|})^n \cdot \exp(-\tfrac{|\alpha|^2}{2}) \quad , \\ B &= \tfrac{n}{|\alpha|} - |\alpha| \quad . \end{aligned} \tag{5.8}$$

with μ the signal mean, σ the Gaussian component width, α the boundary value between the Gaussian and Exponential components and n the Exponential parameter. The tail parameters are not left free to fit but are fixed to the values fitted from LHCb MC, Figure 5.6. The values used for the $B_s^0 \to K^+K^-$, $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ models are provided in Appendix C.

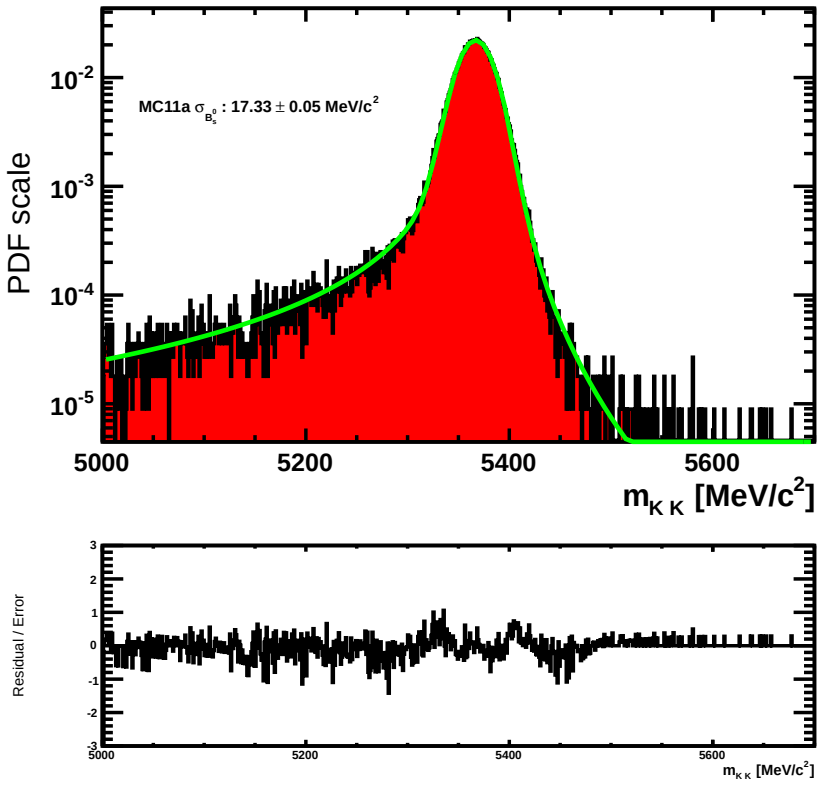


Figure 5.6: Fit to full LHCb $B_s^0 \to K^+K^-$ MC using a double crystal ball mass model. The tail parameters extracted from the fit are then fixed in the mass fit.

5.4.2.2 Misidentified Background Model

The misidentified specific backgrounds cannot easily be modelled parametrically, so a non parametric PDF is produced from MC and used as the fit model, Figure 5.7. Decay modes that are expected to contribute to each channel are included in the fits, these include: $\mathrm{B}^0_d \to \pi^+\pi^-$, $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ and $\Lambda_b \to p\pi^-$ for the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ channel and $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$, $\mathrm{B}^0_d \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ and $\Lambda_b \to p\mathrm{K}^-$ for the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ channel. Simulated datasets, using the full LHCb MC, of each background channel are produced and the full selection and PID cuts for either $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ or $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ applied. Due to the limited number of MC events passing these conditions, the remaining distribution is then smoothed before being used as the mass template for that background class.

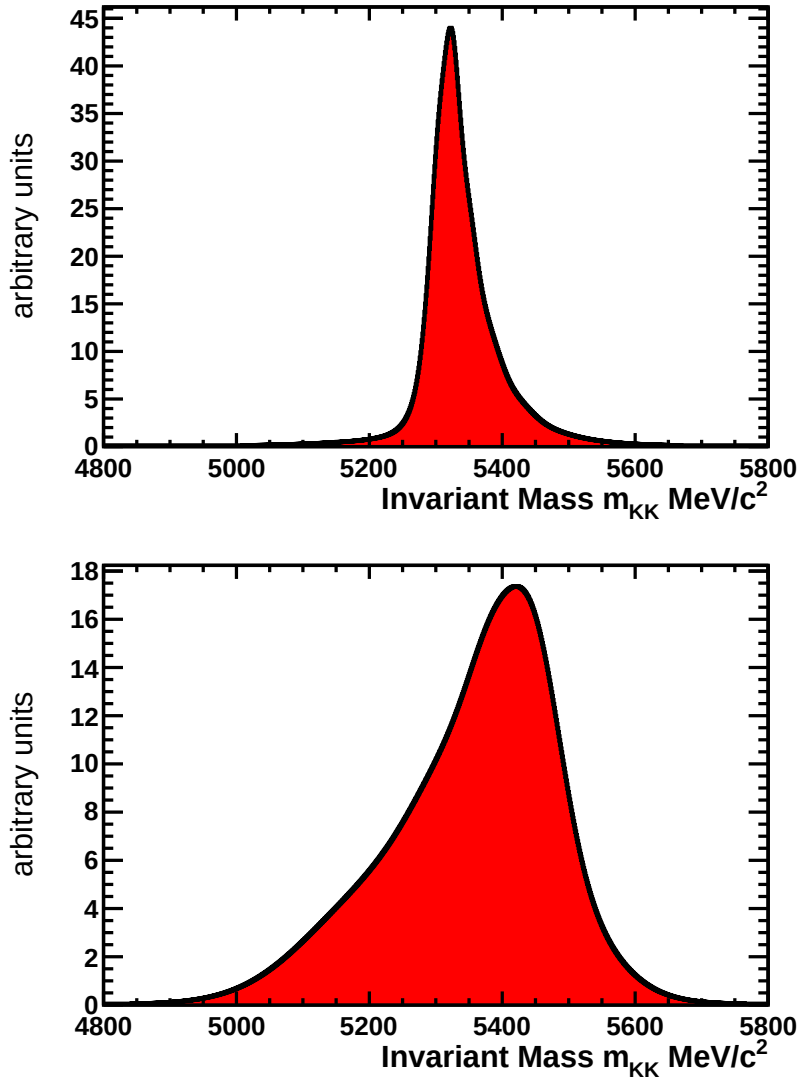


Figure 5.7: Misidentified background templates from full LHCb MC. Two contributing backgrounds to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ signal, $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ (top) and $\Lambda_b \to p\mathrm{K}^-$ (bottom).

5.4.2.3 Partially Reconstructed Background Model

The partially reconstructed background can be modelled using an Exponentially Modified Gaussian (EMG) function [103], given by

$$\mathrm{EMG}(x;\mu,\sigma,\lambda) = \frac{\lambda}{2} e^{\frac{\lambda}{2}(2x+\lambda\sigma^2-2\mu)} \cdot \mathrm{erfc}\left(\frac{x+\lambda\sigma^2-\mu}{\sqrt{2}\sigma}\right), \tag{5.9}$$

where the signs of the variable x and parameter μ are reversed compared to the standard definition. The distribution has different parameters for the $B_s^0 \to K^+K^-$ or $B_d^0 \to K^+\pi^-$ selected data, with the $B_d^0 \to K^+\pi^-$ background modelled using two EMG functions where the first describes the dominant $B_d^0 \to K^+\pi^-\pi^0$ decay, and the second describes the sum of the remaining contributing channels. The parameters for each model are extracted from a cocktail of MC channels (a full list is provided in Appendix B) with the parameters fitted (see Figure 5.8) provided in Appendix C. It is noted that the parameterised distribution for the $B_s^0 \to K^+K^-$ does not match the data as accurately as the $B_d^0 \to K^+\pi^-$ distribution. To resolve this issue, the μ and λ parameters are floated in the $B_s^0 \to K^+K^-$ fit.

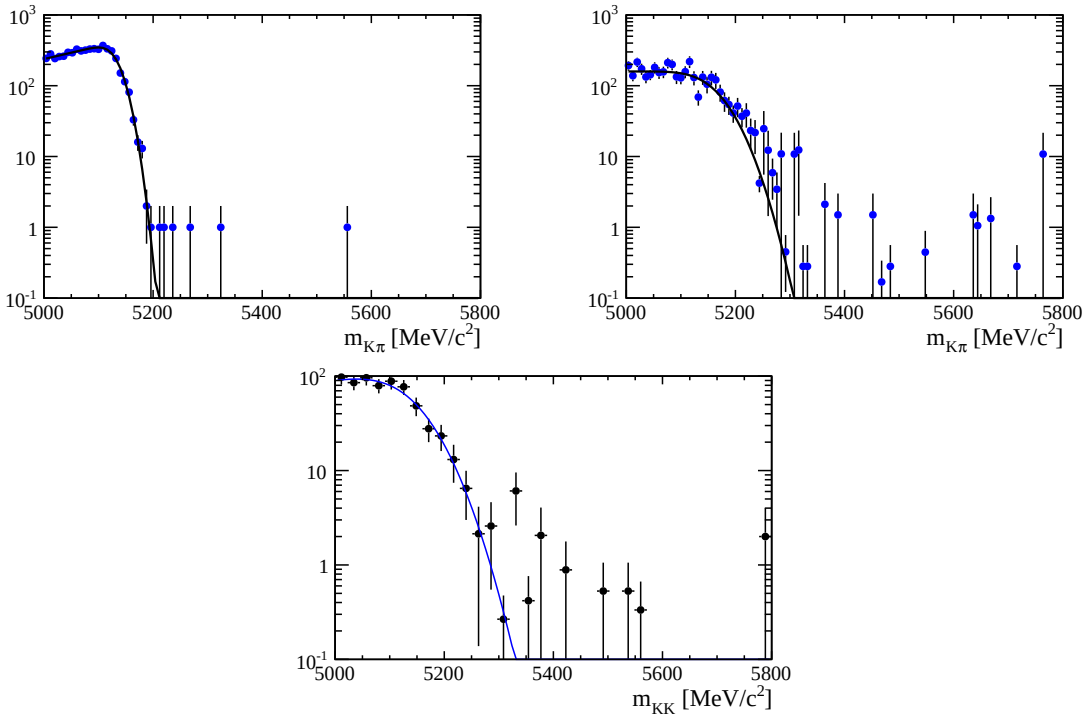


Figure 5.8: Partially reconstructed background, modelled using an exponentially modified gaussian (EMG). Weighted sum of the contributing decay modes from 2011 MC for the Kπ (top) and KK (bottom) mass spectra. The distributions in the Kπ spectrum is split into the dominant $B_d^0 \to K^+\pi^-\pi^0$ decay (top left) and the weighted sum of the other decays (top right) [89].

5.4.2.4 Combinatorial Background Model

The combinatorial background is modelled using the same description as detailed in Section 5.4.1.4.

5.5 Lifetime Model Descriptions

The models used to describe the signal channels present in each dataset are a mixture of parametric and non-parametric models. The parametric models use a single exponential function (Equation 4.1), with the non-parametric models determined via the method described in Section 4.5.2.

5.5.1 Lifetime Models 2010

The lifetime analysis of the 2010 dataset consists of two signal classes, the signal and combinatorial background. The signal is described with a single exponential function, Equation 4.1, with the combinatorial background being a non-parametric model determined by the method described in Section 4.5.2.

5.5.2 Lifetime Models 2011

The 2011 lifetime analysis adds additional models, with the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ datasets both composed of seven signal classes. The three primary $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal models, $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$, are described using the single exponential function as before, Section 5.5.1. Additionally there are the misidentified backgrounds and a partially reconstructed background that also contribute, as well as the standard combinatorial background. The misidentified backgrounds are modelled using a single exponential function, which is fixed using the world average lifetime for the respective decay mode. The partially reconstructed background is modelled using a non-parametric PDF, determined by the same method applied to the combinatorial background for reproducing the decay time distribution (Section 4.5.2).

5.6 Method Verification

Verification of the analysis method is done using a variety of techniques. The methods used between the 2010 and 2011 analyses are similar but with some subtle differences that will be

discussed. The methods primarily include using simplified and full Monte Carlo simulation, and also high statistics $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ decay data.

5.6.1 Verification of the Method with Simplified Simulations

The mass and lifetime components of the fitter are verified using a simplified and stand-alone Monte Carlo (MC) generator. The simplified MC generation, which henceforth will be referred to as 'toy MC', allows us to perform multiple studies. These are chiefly to evaluate the uncertainties of our fit results due to systematic effects, which are detailed in Section 6.3, and to verify that the fitter converges to the correct values of the fit parameters with correctly estimated errors. The toy MC simulation is produced using the ROOT [104] TRandom3 generator to generate the required distributions to match the data. The toy MC data samples generated for evaluation are a mixture of signal and background classes which are determined by the relative fractions that are given as parameters to the generator. For each of the analyses, a set of models are determined that best describe the data. A description of the models used for each analysis is provided as follows.

5.6.1.1 2010 Toy MC Simulation Models

The relative small size of the 2010 dataset, $37\,\mathrm{pb}^{-1}$, means various simplifications can be used in modelling the signal classes expected in the selected events. The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal channels, $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$, have their invariant mass distributions modelled approximately with a Gaussian distribution [76] [75]. The generated lifetimes of each signal class follow a single exponential function. Additionally a Gaussian distribution is used to generate the turning point, t_{min}, with the maximum, t_{max}, determined by the maximum lifetime cut in the fitter. The lifetimes generated are required to fall in the generated acceptance intervals. The signal and background models used for each dataset in the toy MC datasets are displayed in Figure 5.9 and detailed in Tables 5.9 and 5.10. The parameters used for the toy MC models are provided in Table 5.13.

Signal Class	Description	Mass Model	Lifetime Model
$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$	Signal	Gaussian	Exponential
$\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$	Specific background	Gaussian	Exponential
Combinatorial	Background	Linear	Exponential

Table 5.9: Signal channel models employed in $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ toy MC simulations for 2010 verification studies.

Signal Class	Description	Mass Model	Lifetime Model
$B_s^0 \to K^+K^-$	Signal	Gaussian	Exponential
Combinatorial	Background	Linear	Exponential

Table 5.10: Signal channel models employed in $B_s^0 \to K^+K^-$ toy MC simulations for 2010 verification studies.

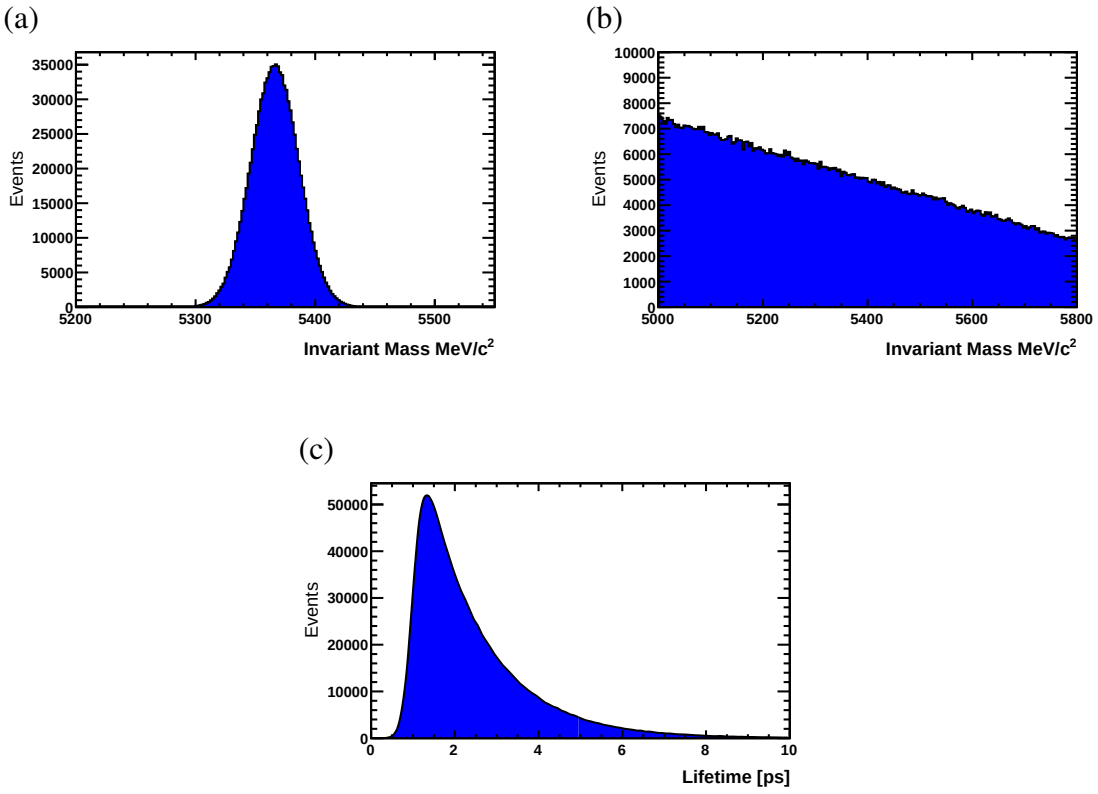


Figure 5.9: Toy MC generated mass distributions for 2010 verification. (a) Gaussian $B \to h^+h^-$ signal mass distribution, (b) Linear distribution for combinatorial background, (c) Exponential lifetime distribution smeared with Gaussian acceptance function.

5.6.1.2 2011 Toy MC Simulation Models

The 2011 dataset is substantially larger, by approximately a factor of 30, than the data collected during the 2010 data taking run. This increase in events also brings into play additional backgrounds and requires the redefinition of the previously implemented signal models. The mass models of the primary signal channels, $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$, are modelled using a double crystal ball function, as described in Section 5.4.2.1. Additional backgrounds that were not present in the 2010 simulation, such as partially reconstructed background and misidentified backgrounds are included. The partially reconstructed background is modelled with an EMG and has parameters similar to that of data. The misID backgrounds are modelled using the same templates as used in the data fits, Section 5.4.2.2. All decay time distributions are generated from an exponential distribution, with the turning point (t_{min}) distribution generated from a bifurcated Gaussian distribution with parameters from data. All acceptance intervals take the form of either a single step or top-hat function. As previously, the generated decay time is forced to fall within the generated acceptance interval. The signal and background models used for each dataset in the toy MC datasets are displayed in Figure 5.10 and detailed in Tables 5.11 and 5.12. The parameters used for the toy MC model are provided in Table 5.14.

Signal Class	Description	Mass Model	Lifetime Model
$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$	Signal	DCB	Exponential
$\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$	Signal	DCB	Exponential
$\mathrm{B}_d^0 \to \pi^+\pi^-$	MisID background	Template	Exponential
$\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$	MisID background	Template	Exponential
$\Lambda_b \to p\pi^-$	MisID background	Template	Exponential
Partially Reconstructed	Background	EMG	Exponential
Combinatorial	Background	Linear	Exponential

Table 5.11: Signal channel models employed in $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ toy MC simulations for 2011 verification studies.

Signal Class	Description	Mass Model	Lifetime Model
$\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$	Signal	DCB	Exponential
$\mathrm{B}_d^0 \to \mathrm{K}^+\mathrm{K}^-$	Specific background	Template	Exponential
$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$	MisID background	Template	Exponential
$\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$	MisID background	Template	Exponential
$\Lambda_b \to p\mathrm{K}^-$	MisID background	Template	Exponential
Partially Reconstructed	Background	EMG	Exponential
Combinatorial	Background	Linear	Exponential

Table 5.12: Signal channel models employed in $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC simulations for 2011 verification studies.

5.6.1.3 Event Generation

The events from the toy MC Generator are produced through the following sequence:

- The event class is determined from a number generated randomly from a rectangular distribution according to the required signal fractions in the parameter list.
- The decay time acceptance is modelled by generating a minimum acceptance time (τ_{min}) per event, from a Gaussian or Bifurcated Gaussian distribution, which is parameterised from data.
- The mass and decay time of the event are generated according to the distributions of that event class. The decay time generated is rejected if the decay time does not lie in the acceptance intervals, τ_{min} and τ_{max}, and a new decay time is then generated.

Hence the decay time distribution is identical to zero for $\tau < \tau_{min}$ and has a discontinuity at $\tau = \tau_{min}$. The default values of the parameters used to generate the 2010 $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ toy MC datasets are given in Table 5.13, with the additional parameters used to generate the 2011 toy MC datasets provided in Table 5.14.

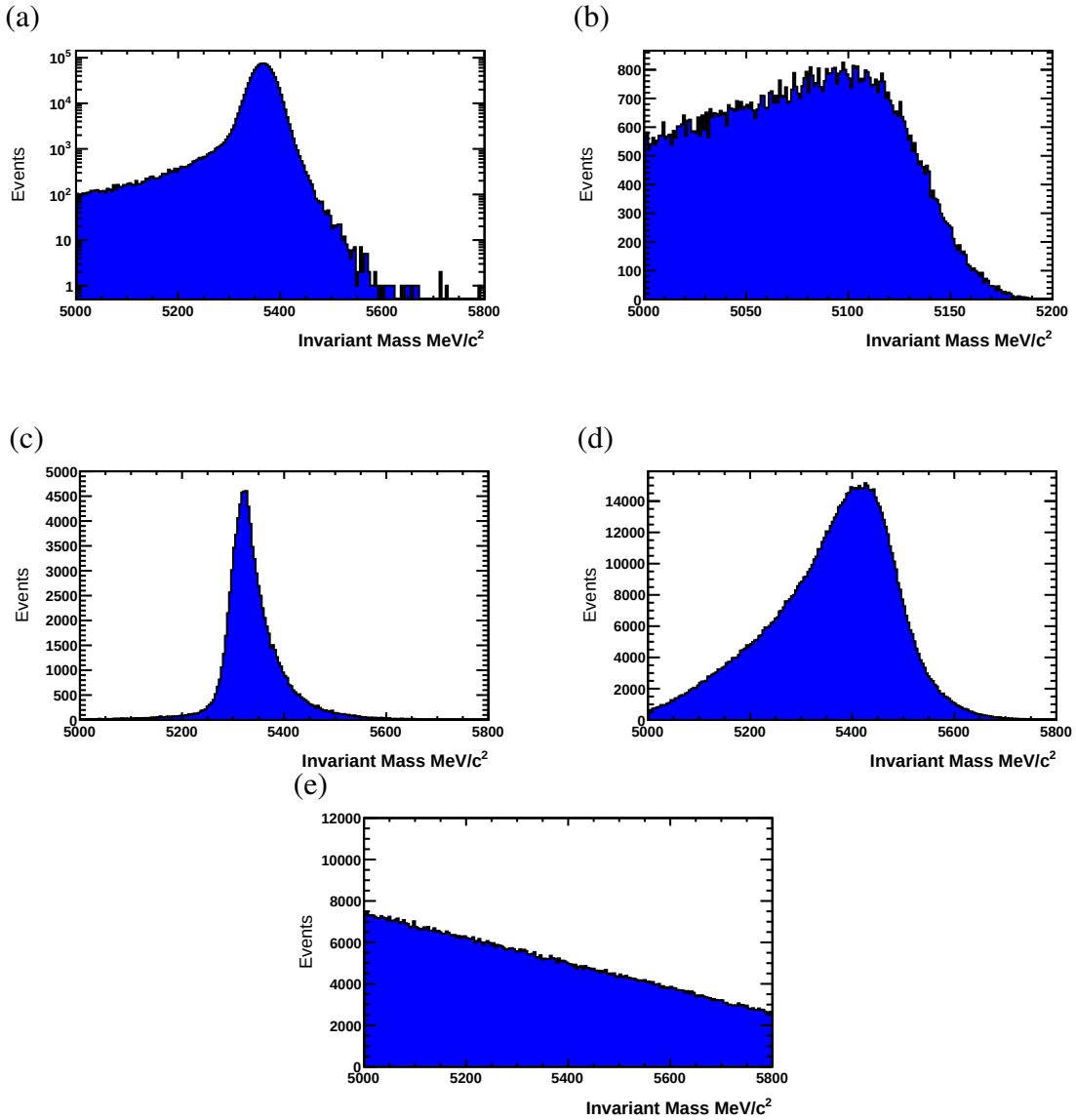


Figure 5.10: Toy MC Generated mass distributions for 2011 verification. Double Crystal Ball function for $B \to h^+h^-$ signal mass distribution (a), Exponentially Modified Gaussian function for Partially Reconstructed background (b), numerical PDF template for $B^0_d \to K^+\pi^-$ misidentified as $B^0_s \to K^+K^-$ (c), numerical PDF template for $\Lambda_b \to pK^-$ misidentified as $B^0_s \to K^+K^-$ (d), Linear function describing the combinatorial background (e).

Parameter	Default value	Comment
$\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$		
$m_{\mathrm{B}_s^0}$	5366.3 MeV/c^2	B_s^0 mass
$\sigma_{\mathrm{B}_s^0}$	25 MeV/c^2	B_s^0 Mass resolution
$\tau_{\mathrm{B}_s^0}$	1.477 ps	B_s^0 average lifetime
$\Delta\Gamma_\mathrm{s}$	0.1 ps^{-1}	Decay width difference, B_s^L vs. B_s^H
$\mathcal{A}_{\Delta\Gamma_s}$	-0.97, -0.8, -0.6	B_s^0 rate asymmetry
		(SM value, medium and large CP violation)
$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$		
$m_{\mathrm{B}_d^0}$	5279.5 MeV/c^2	B_d^0 mass
$\sigma_{\mathrm{B}_d^0}$	25 MeV/c^2	B_d^0 Mass resolution
τ_B	1.525 ps	B_d^0 lifetime
Combinatorial		
m_{min}^{bkg}	5176 MeV/c^2	Minimum background mass value ($\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$)
m_{min}^{bkg}	5272 MeV/c^2	Minimum background mass value ($\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$)
m_{max}^{bkg}	5800 MeV/c^2	Maximum background mass value
$\frac{S}{S+B}$	0.7	Signal fraction (similar to data)
∇^{bkg}	$-3 \cdot 10^{-6}$	Gradient of background mass PDF
τ_1^{bkg}	1 ps	Short lifetime component
τ_2^{bkg}	5 ps	Long lifetime component
R^{bkg}	0.8, 1.0	Fraction of short component
Lifetime		
σ_τ	50 fs	Proper time resolution
τ_{min}	1 ps	Average minimum acceptance time
$\sigma_{\tau_{min}}$	0.2, 0.3 ps	Spread in minimum acceptance

Table 5.13: Default values of the parameters used in the toy MC generator for verification of the fitter in preparation for the 2010 analysis. The values are from the PDG [22], or similar to those observed in LHCb data.

Parameter	Default value	Comment
$\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$		
$\alpha_{Low}^{\mathrm{B}_s^0}$	1.39	Low CB part boundary (B_s^0)
$\alpha_{High}^{\mathrm{B}_s^0}$	1.89	High CB part boundary (B_s^0)
$n_{Low}^{\mathrm{B}_s^0}$	1.65	Low CB part tail (B_s^0)
$n_{High}^{\mathrm{B}_s^0}$	6.78	High CB part tail (B_s^0)
$fraction_{Low}^{\mathrm{B}_s^0}$	0.30	MC11a fraction of low CB part (B_s^0)
KK Partial		
μ_{KK}^{EMG}	5179 ± 24 MeV/c^2	EMG peak
$\sigma_{\mathrm{KK}}^{EMG}$	53.80 MeV/c^2	EMG smearing
$\lambda_{\mathrm{KK}}^{EMG}$	$(4.5 \pm 0.7) \cdot 10^{-4}$ (MeV/c^2)$^{-1}$	EMG tail
$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$		
$\alpha_{Low}^{\mathrm{B}_d^0}$	1.16	Low CB part boundary (B_d^0)
$\alpha_{High}^{\mathrm{B}_d^0}$	1.81	High CB part boundary (B_d^0)
$n_{Low}^{\mathrm{B}_d^0}$	1.85	Low CB part tail (B_d^0)
$n_{High}^{\mathrm{B}_d^0}$	8.80	High CB part tail (B_d^0)
$fraction_{Low}^{\mathrm{B}_d^0}$	0.29	MC11a fraction of low CB part (B_d^0)
$\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$		
$\alpha_{Low}^{\mathrm{B}_s^0}$	1.00	Low CB part boundary (B_s^0)
$\alpha_{High}^{\mathrm{B}_s^0}$	1.96	High CB part boundary (B_s^0)
$n_{Low}^{\mathrm{B}_s^0}$	2.04	Low CB part tail (B_s^0)
$n_{High}^{\mathrm{B}_s^0}$	5.78	High CB part tail (B_s^0)
$fraction_{Low}^{\mathrm{B}_s^0}$	0.23	MC11a fraction of low CB part (B_s^0)
Kπ Partial		
$\mu_{\mathrm{K}\pi}^{EMG_1}$	5137.0 MeV/c^2	EMG peak
$\sigma_{\mathrm{K}\pi}^{EMG_1}$	20.0 MeV/c^2	EMG smearing
$\lambda_{\mathrm{K}\pi}^{EMG_1}$	$(4.16) \cdot 10^{-3}$ (MeV/c^2)$^{-1}$	EMG tail
$\mu_{\mathrm{K}\pi}^{EMG_2}$	5167.0 MeV/c^2	EMG peak
$\sigma_{\mathrm{K}\pi}^{EMG_2}$	43.60 MeV/c^2	EMG smearing
$\lambda_{\mathrm{K}\pi}^{EMG_2}$	$(7.3) \cdot 10^{-5}$ (MeV/c^2)$^{-1}$	EMG tail
$R_{\mathrm{K}\pi}^{EMG_{12}}$	11.70	Ratio of the two EMGs

Table 5.14: Additional parameters used in the toy MC generator for verification of the fitter in preparation for the 2011 analysis. The values are from the PDG [22], fitted from LHCb MC or similar to those observed in LHCb data.

5.6.1.4 Verification Fit Results

The fit method and fitter are verified using the toy MC generator detailed in Section 5.6.1. The toy MC generator generates multiple independent datasets with statistics comparable with that of the data, and includes all expected signal classes. Each of these datasets are then fitted individually, with the resultant fit values for each parameter and error value used to generate a pull distribution using the equation

$$\text{Pull} = \frac{(\mu_{fitted} - \mu_{true})}{\sigma_{fitted}}. \tag{5.10}$$

It is expected that the pull quantity be Gaussian distributed, with a mean of zero if the fitter is unbiased. If the errors are correctly estimated, the width should be one. The verification pull distributions of each lifetime for the 2010 and 2011 verification studies are provided.

2010 Verification Pull Plots

The free parameters of the 2010 toy MC models are fitted, and the pulls of these, based on Equation 5.10, are given in Table 5.15. The pull distribution of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime is shown in Figure 5.11, with the parameters fitted from the invariant mass spectrum provided in Appendix D. These verifications were done using 1000 independently generated toy MC datasets.

Parameter	Pull Mean	Pull Sigma
$\mu_{\mathrm{B}_s^0}$	0.02352 ± 0.03444	1.046 ± 0.025
$\sigma_{\mathrm{B}_s^0}$	-0.13660 ± 0.03240	1.004 ± 0.024
$f_{\mathrm{B}_s^0}$	-0.06038 ± 0.03282	1.017 ± 0.024
∇_{comb}	-0.05534 ± 0.03247	0.992 ± 0.023
$\tau_{\mathrm{B}_s^0}$	-0.05237 ± 0.03272	1.002 ± 0.025

Table 5.15: Pulls of the free parameters in the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC simulations corresponding to the 2010 lifetime measurement.

2011 Verification Pull Plots

All free parameters of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ 2011 toy MC models are fitted, with the pull distribution for each parameter determined from Equation 5.10. The pull values of all parameters from both the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ toy MC datasets, are given in Table 5.16 and 5.17 respectively. The lifetime pulls of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$

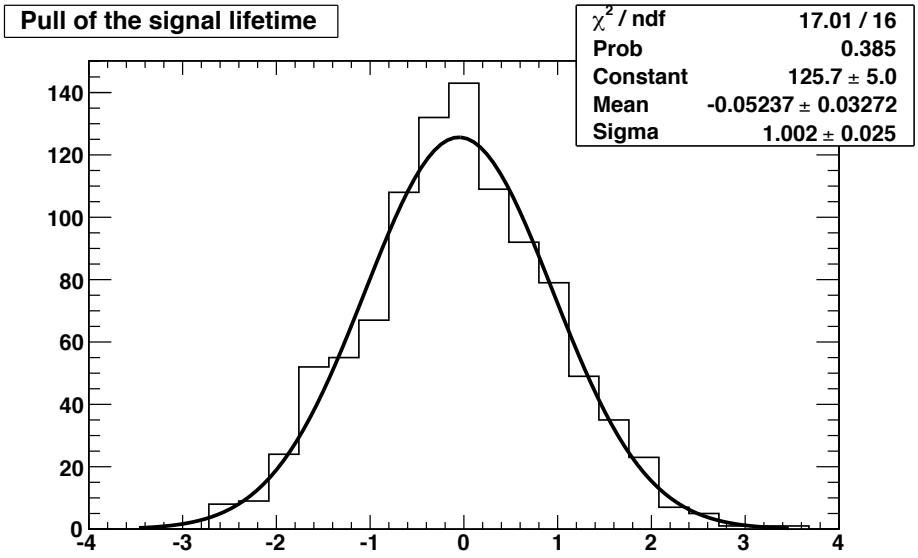


Figure 5.11: Pull on the lifetime from $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC studies using the 2010 fitter.

and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ are displayed in Figure 5.12. The remaining pulls to the mass distribution parameters are provided in Appendix D. These verifications were done using 5000 independently generated toy MC datasets.

Parameter	Pull Mean	Pull Sigma
$\mu_{\mathrm{B}_d^0}$	-0.01457 ± 0.01397	0.978 ± 0.010
$\sigma_{\mathrm{B}_d^0}$	0.00047 ± 0.01420	0.992 ± 0.011
$f_{\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-}$	-0.01375 ± 0.01418	0.995 ± 0.010
$f_{\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-}$	0.00439 ± 0.01366	0.957 ± 0.010
$f_{Partial}$	0.01806 ± 0.01423	0.998 ± 0.010
∇_{comb}	0.00720 ± 0.01391	0.974 ± 0.010
$\tau_{\mathrm{B}_d^0}$	0.00275 ± 0.01472	1.021 ± 0.010
$\tau_{\mathrm{B}_s^0}$	0.01492 ± 0.01539	1.073 ± 0.011

Table 5.16: Pulls of the free parameters in the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ toy MC simulations corresponding to the 2011 lifetime measurement.

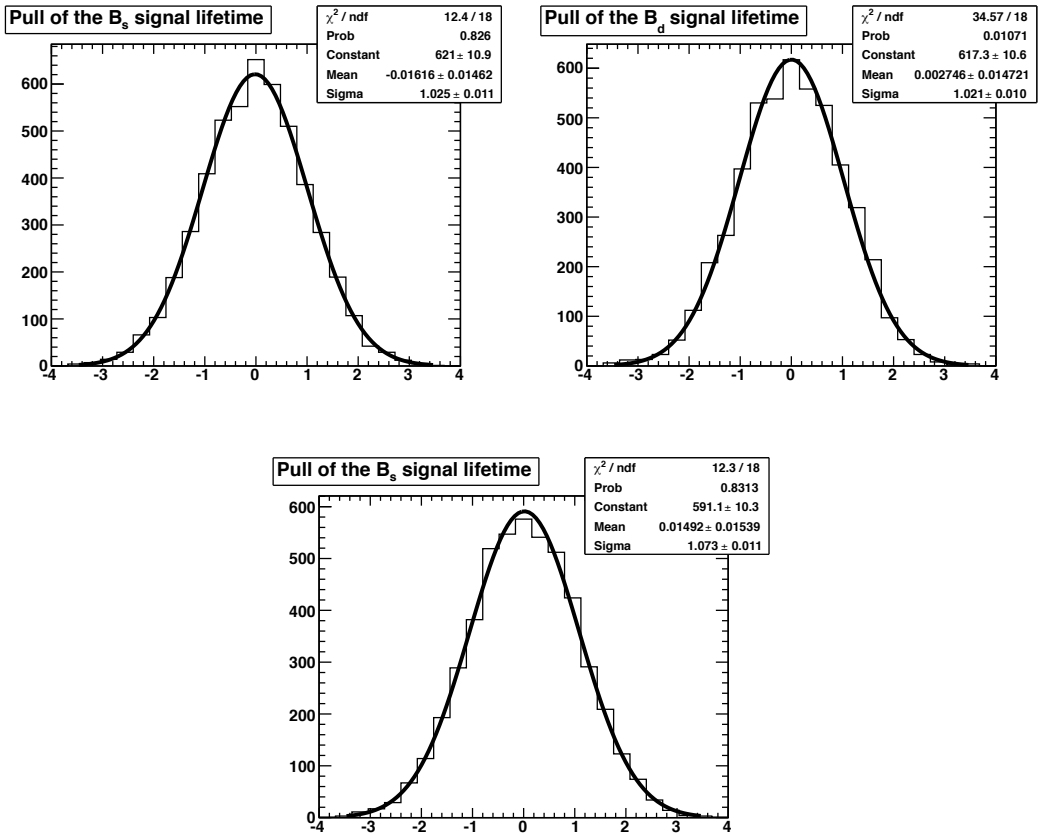


Figure 5.12: Pull on the lifetime from toy MC studies for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ (top left), $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ (top right) and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ (bottom centre) using the 2011 fitter.

Parameter	Pull Mean	Pull Sigma
$\mu_{\mathrm{B}_s^0}$	-0.00309 ± 0.01425	0.997 ± 0.010
$\sigma_{\mathrm{B}_s^0}$	-0.01367 ± 0.01417	0.993 ± 0.010
$f_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$	-0.00637 ± 0.01410	0.990 ± 0.010
$f_{Partial}$	0.01189 ± 0.01410	0.990 ± 0.010
∇_{comb}	-0.00549 ± 0.01408	0.975 ± 0.010
$\tau_{\mathrm{B}_s^0}$	-0.01616 ± 0.01462	1.025 ± 0.011

Table 5.17: Pulls of the free parameters in the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC simulations corresponding to the 2011 lifetime measurement.

5.6.2 Verification with Full LHCb Monte Carlo

The fitter performance is further verified by fitting and extracting parameters from full LHCb generated MC. The MC is selected using the full selection, which includes the trigger, stripping and offline selections, before being processed through the swimming procedure to determine the per-event acceptance functions. For the combinatorial background, there is an insufficient amount of minimum bias MC that passes the selection. As such, an equivalent sample is generated via the Toy MC generator, described in section 5.6.1, and then combined with the signal MC to replicate the data. Verification studies are performed with both the 2010 and updated 2011 fitter methods.

5.6.2.1 2010 MC verification

Using the LHCb selection software, 4000 $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ candidates are selected from the full MC data set. The true values of the mass mean, and lifetime, used for the generation of the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ MC were

$$m_{\mathrm{B}_d^0} = 5279.53\,\mathrm{MeV}/c^2 \;\;,\;\; \tau_{\mathrm{B}^0} = 1.525\,\mathrm{ps}. \tag{5.11}$$

The data set fitted contains a total of 6928 events, 3461 signal and 3467 background. The fit performed is shown in Figure 5.13 and gives

$$m_{\mathrm{B}_d^0} = 5280.7 \pm 0.37\ \mathrm{MeV}/c^2 \;\;,\;\; \tau_{\mathrm{B}^0} = 1.514 \pm 0.027\ \mathrm{fs}. \tag{5.12}$$

The lifetime fit result returned is in agreement with the value used in the MC generation. However the mass value deviates from the expected value by $\approx 3\sigma$. This deviation is likely due to the mass model used, as the single Gaussian mass model used for the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ signal peak is only an approximation as discussed in Section 5.4.1.

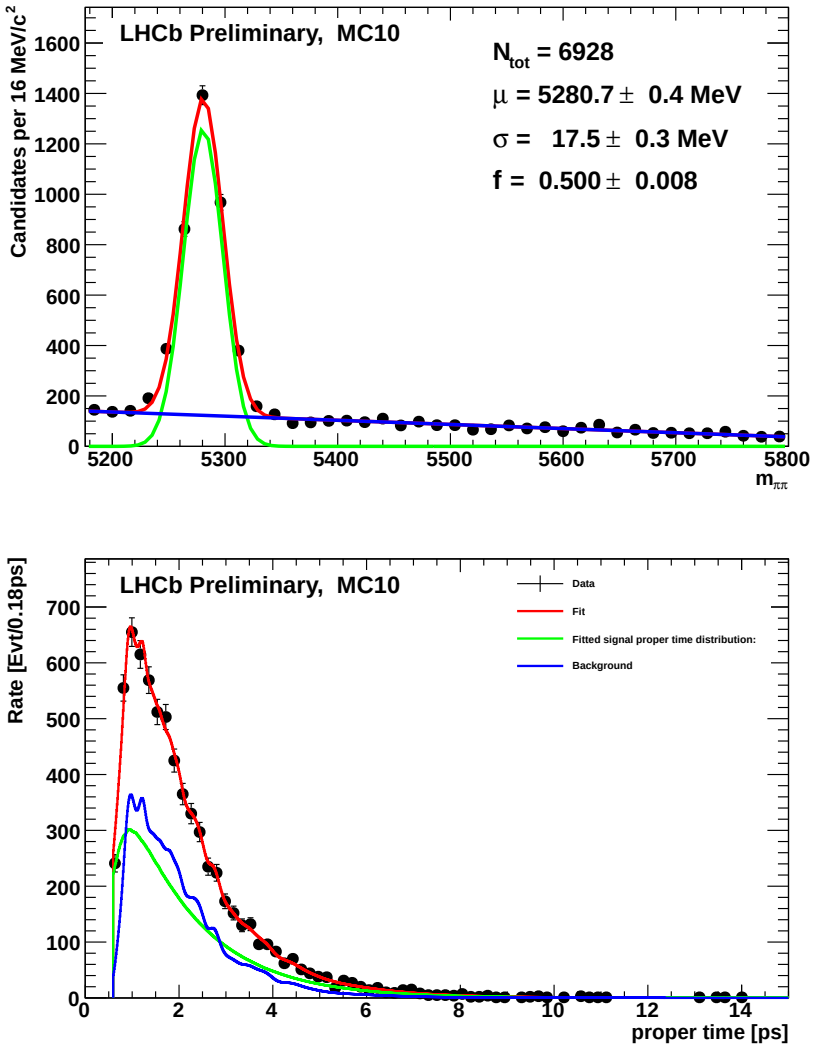


Figure 5.13: Mass (top) and lifetime fit (bottom) results for $MC10$ $B^0_d \to K^+\pi^-$.

5.6.2.2 2011 MC verification

The mass and lifetime fitter has also been verified using the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$, $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ and $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ datasets from the full 2011 updated LHCb MC. To simulate the way offline selected data looks, the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ MC is mixed in with a toy generated combinatorial sample (see Figure 5.14), with the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ combined only with toy combinatorial. The mass fits to these replicated data samples are shown in Figures 5.14 and 5.15 and the lifetime fits in Figures 5.16 and 5.17.
The results of the lifetime fit to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ MC11a signal data is presented in Table 5.18, these fits are all within the errors of the generated MC.

MC11a Signal Class	Generated $\tau_{B^0_{d/s}}$ (ps)	Fitted $\tau_{B^0_{d/s}}$ (ps)	Fitted $\tau_{B^0_{d/s}}$ error
$\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$	1.407	1.403	0.005
$\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$	1.525	1.503	0.019
$\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$	1.472	1.471	0.043

Table 5.18: MC11a lifetime fits results compared with generated lifetime

5.6.3 Verification with $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ Data

The fitter has also been verified using data. This was done specifically for the 2010 lifetime analysis as the event statistics for the measurement were relatively low. The topologically similar $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ family have been used in an independent analysis [105] and can be used in this instance as a high statistics data sample for testing. This measurement is also suitable as a verification as it uses the same method, and fitting framework, as the B^0_s and B^0_d lifetime measurements made in this thesis. The main difference between the $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ and $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ data is the need to distinguish between promptly produced D mesons from those produced from the decay of long-lived particles. This separation is performed using the χ^2 of the impact parameter of the D with respect to the primary vertex.
Reference [105] reports lifetime measurements in tagged and untagged samples for $\mathrm{D} \to \mathrm{K}\pi$ and $\mathrm{D} \to \mathrm{KK}$ for the determination of y_{CP} and A_Γ. The lifetimes for the $\mathrm{D} \to \mathrm{K}\pi$ modes are measured to be

$$\tau^{eff}(\mathrm{D}^0) = 410.6 \pm 1.3 \text{ fs} \quad (5.13)$$

$$\tau^{eff}(\overline{\mathrm{D}}^0) = 409.9 \pm 1.3 \text{ fs}, \quad (5.14)$$

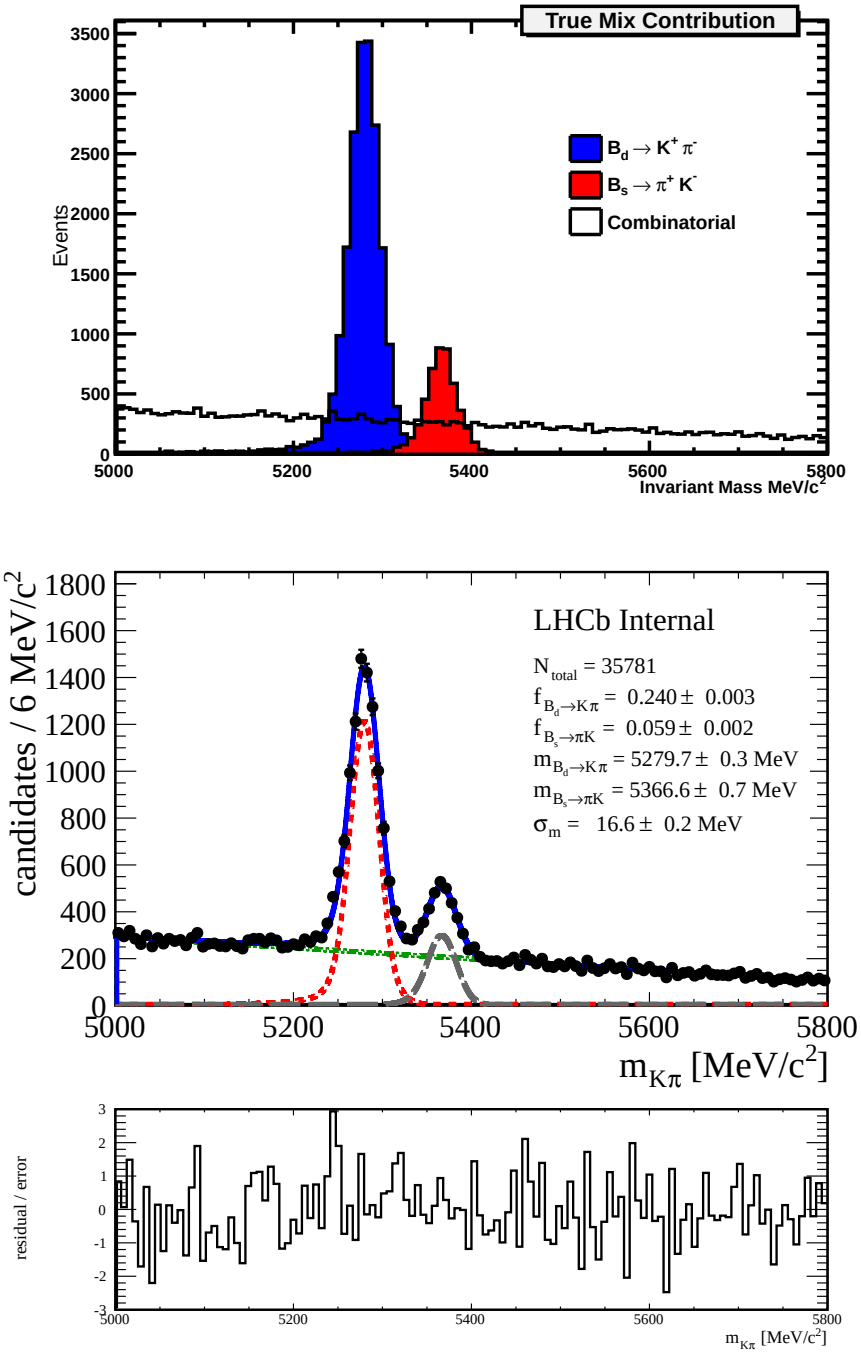


Figure 5.14: Mix of MC11a $B^0_d \to K^+\pi^-$, $B^0_s \to \pi^+K^-$ and Toy MC combinatorial (top), MC11a $B^0_d \to K^+\pi^-$, $B^0_s \to \pi^+K^-$ and Toy MC combinatorial mass fit (middle) and the mass fit residual (bottom).

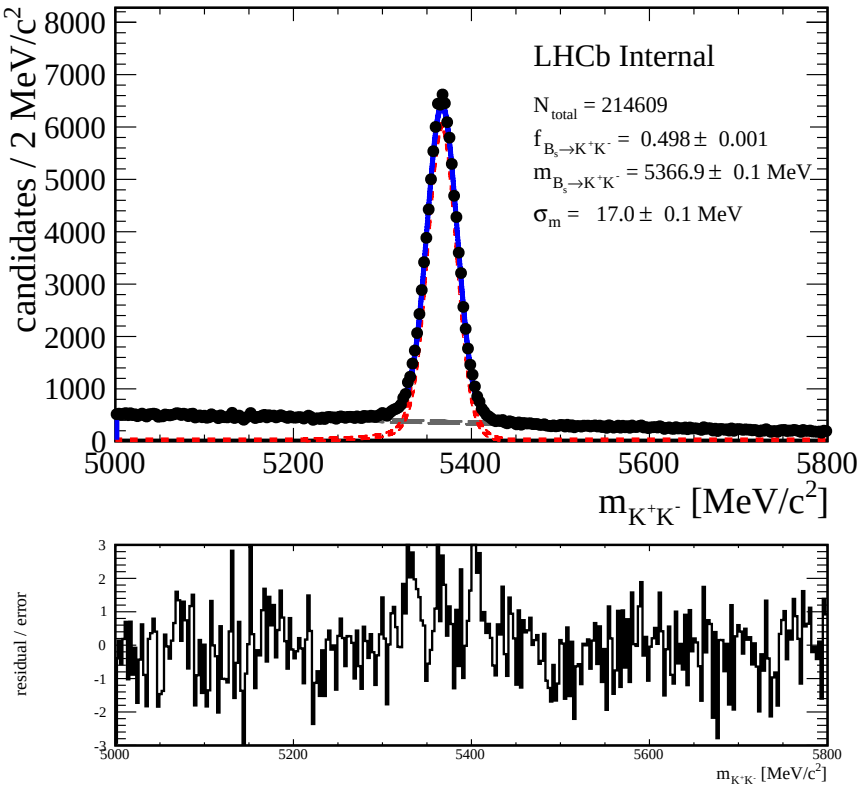


Figure 5.15: $B_s^0 \to K^+K^-$ MC11a signal and Toy MC combinatorial mass fit (top) and mass fit residual (bottom).

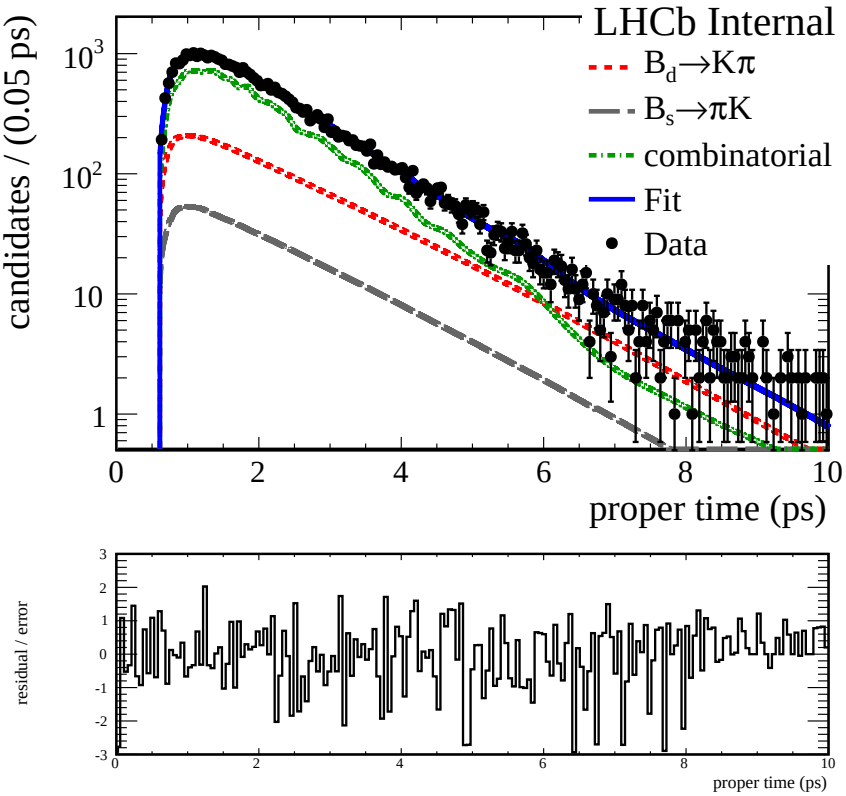


Figure 5.16: MC11a $B^0_d \to K^+\pi^-$, $B^0_s \to \pi^+K^-$ and toy MC combinatorial lifetime fit (top) and lifetime fit residual (bottom)

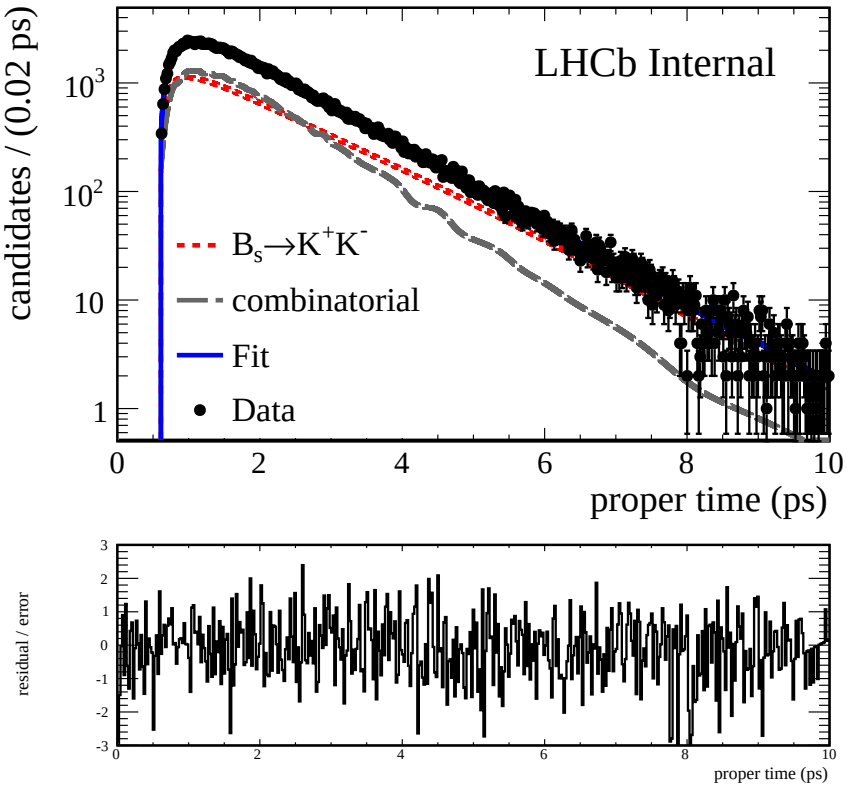


Figure 5.17: $B_s^0 \to K^+K^-$ MC11a signal and Toy MC combinatorial lifetime fit (top) and lifetime fit residual (bottom).

where the quoted errors being statistical only. The average of these lifetimes ($410.3 \pm 0.9\ (\mathrm{stat})\ \mathrm{fs}$) should be compared with the current world best measurements of $410.1 \pm 1.5\ \mathrm{fs}$ [22].

5.7 Summary

Preparation studies for the analysis of 2010 and 2011 data presented in Chapter 6 are provided in this chapter, beginning with an overview in Section 5.2 of the experimental status for the key $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime measurement. Details of the selection applied on each of the analyses are discussed in Section 5.3, including the methods used to optimise these cuts for the kinematic and PID variables. The mass models used to describe the data evolve between the analyses presented, as such, two sets of models are provided with the descriptions and motives for their use given in Section 5.4. The lifetime model descriptions follow, although particular handling of the non-parametric backgrounds is detailed in the previous chapter. Verification studies of the fitter to confirm its reliability use toy MC simulation data, full LHCb MC and high statistics data from a topologically similar data source, $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$. All studies have shown the fitter performs reliably and converges to the correct values and errors, with these shown in Section 5.6.

Chapter 6

$\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ Systematic Studies and Results.

6.1 Introduction

The previous two chapters, Chapters 4 and 5, have introduced the methodology behind the lifetime measurement followed by descriptions of the data selection, the models used to define the signals and verifications performed on the full fitter. This chapter primarily discusses the evaluation of the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes mentioned throughout this thesis, both for the 2010 and 2011 data taking periods. This begins in Section 6.2 by discussing the extracted signal yields for each dataset. The extracted signal yields are applicable to other analyses, which will be detailed.
Following on from this, Section 6.3 discusses the sources of systematic uncertainty on the lifetime fit and the methods used to determine them. The final fit results of the lifetime analyses are then presented in Section 6.4. A discussion of the physics implications of the results is then presented in Section 6.5.

6.2 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ Event Yields

Prior to the fit to the decay time distribution, a fit to the invariant mass is performed. The resultant signal fractions and event weightings (Section 4.3.3), are passed to the lifetime fitter for evaluation of the reconstructed decay time distribution. The event yields for each signal channel are determined via these signal fractions. The signal yields of the relevant $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ channels are provided for each analysis in Sections 6.2.1 and 6.2.2. A second analysis, which also utilises the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ event yields, is detailed in Section 6.2.3.

6.2.1 Mass Fits to the 2010 Dataset

The 2010 LHCb dataset comprises a total integrated luminosity of $\approx 37\,\mathrm{pb}^{-1}$. After applying the full selection for both the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ decay channels, given in Section 5.3, the number of events remaining in each data sample is 965 and 376 respectively. The $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ mass spectrum is fitted in the invariant mass range $5176 - 5800\ \mathrm{MeV}/c^2$ and provided in Figure 6.1. For this analysis, the mass range selected is imposed to try to remove the majority of the partially reconstructed three-body background in the low mass region, with systematic studies showing that this is indeed the case [75].

The mass fit for the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ is performed both with, and without, the inclusion of the $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ specific background. This was done to ensure the models used for both the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ were the same for this analysis. Both these fits are displayed in Figure 6.1 (top and bottom).

The $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ measurement utilises an analogous mass model to the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ previously detailed. The range of the B^0_s mass spectrum selected for fitting, differs from the range for the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ as the daughter mass hypothesis assumes both daughters are reconstructed as kaons rather than one kaon and one pion. The mass range is selected to remove the partially reconstructed 3-body background with the region selected being $5272 - 5800$ MeV/c^2. The mass fit component of the full lifetime fit is shown in Figure 6.2.

The mean fitted mass of the B^0_s peak is 0.04% lower than the value from PDG (5366.3 MeV/c^2), which is compatible with the known error in the mass scale of LHCb. The mass resolution is compatible with the mass resolution of LHCb measured in other decays. To determine the event yield per channel, the signal fractions provided from the invariant mass fit, Figures 6.1 (top) and 6.2, can be used. These values are provided in Table 6.1.

Signal Channel	Fitted Signal Fraction	Total events in data sample	Signal Yield
$\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$	0.667 ± 0.037	376 ± 19	251 ± 19
$\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$	0.516 ± 0.027	965 ± 31	498 ± 31

Table 6.1: Event Yield for 2010 datasets. The signal fraction is calculated in the mass range $5176 - 5800\,\mathrm{MeV}/c^2$ for $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $5272 - 5800\,\mathrm{MeV}/c^2$ for $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$.

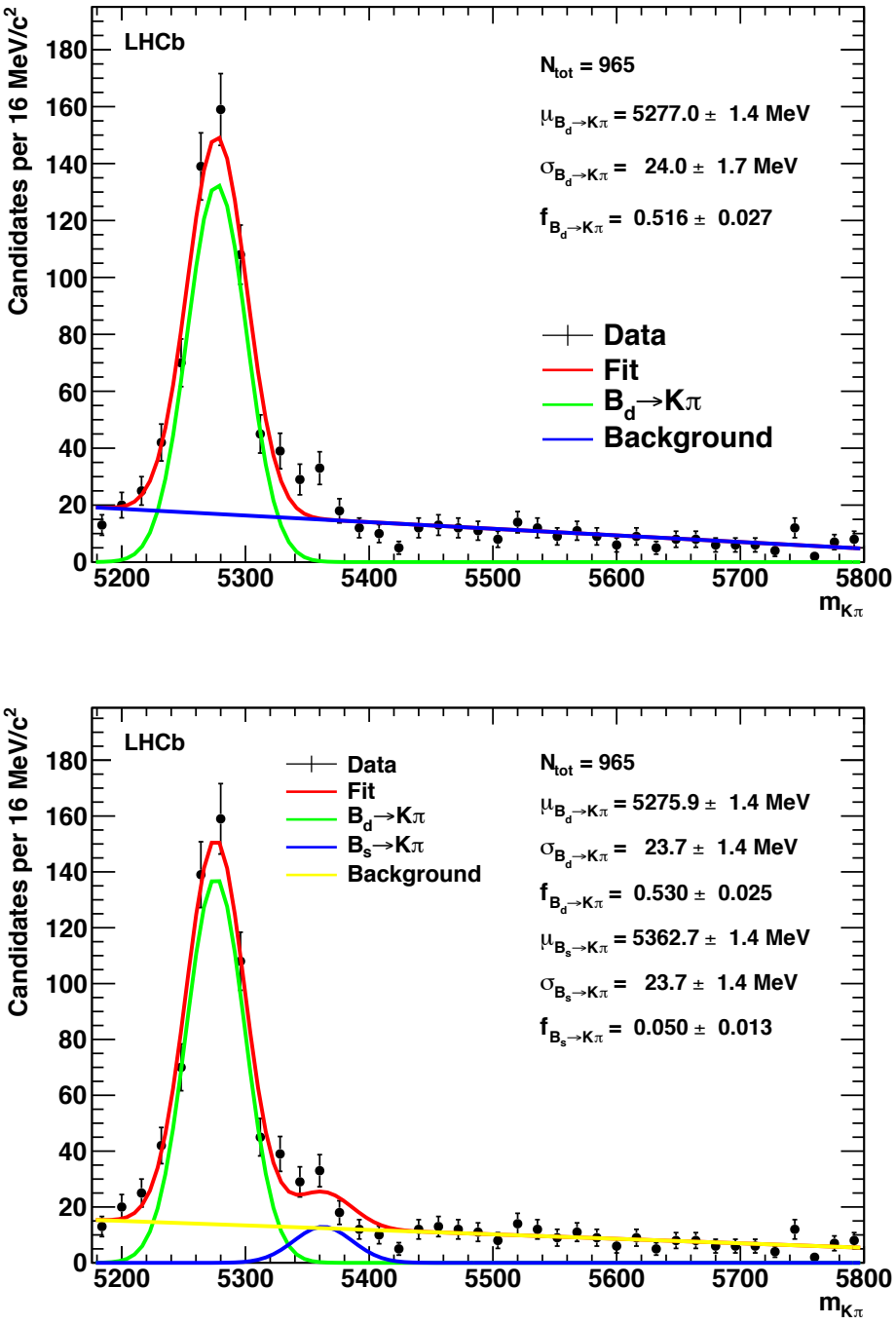


Figure 6.1: $B^0_d \to K^+\pi^-$ Mass distribution and mass fit result using the 2010 fitter. The upper plot shows the mass taking into account only the $B^0_d \to K^+\pi^-$ signal class, which is used in the default fitter. The lower plot also includes the $B^0_s \to \pi^+K^-$ signal class, which is used in the cross-check fit [75, 76]

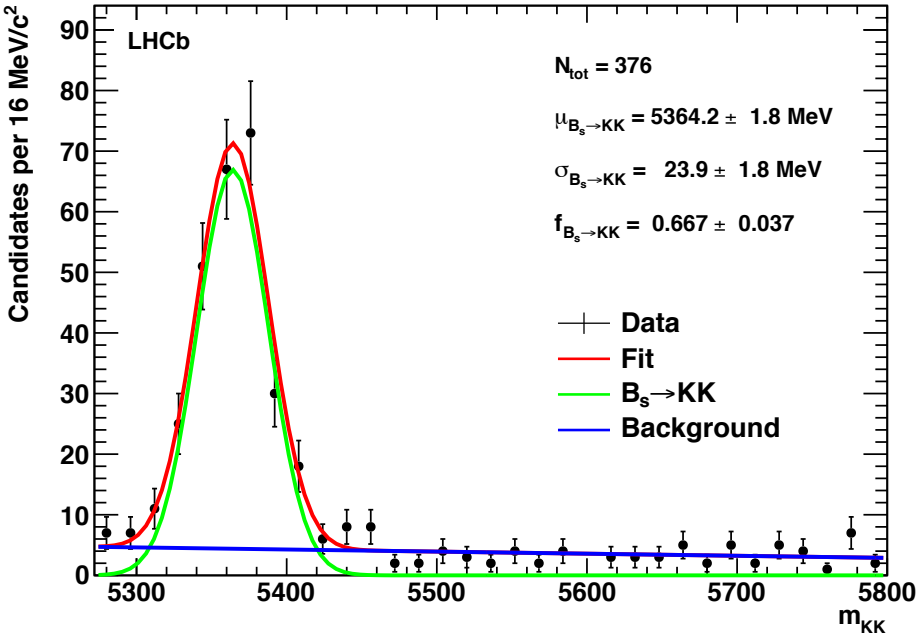


Figure 6.2: Mass distribution and mass fit of $B_s^0 \to K^+K^-$ using the 2010 fitter [75, 76].

6.2.2 Mass Fits to the 2011 Dataset

The 2011 LHCb dataset is significantly larger than that for the 2010 data set, with the collected data having a total integrated luminosity of $\approx 1.0\,\text{fb}^{-1}$. The increase in statistics means that both the selection, Section 5.3, and signal mass models, Section 5.4, were required to be redefined for the analysis of this dataset. After applying the full selection the number of events remaining for the $B_s^0 \to K^+K^-$ and $B_d^0 \to K^+\pi^-$ datasets were 22498 and 60596 respectively.

The composition of the 2011 datasets includes significantly more signal channels than the previously studied 2010 data. As the precision of the measurements are due to increase with the available statistics, subtle effects from these additional backgrounds are required to be accounted for. These have been discussed in Sections 5.4.2.2 and 5.4.2.3.

The $B_d^0 \to K^+\pi^-$ mass spectrum is fitted in the invariant mass range $5000 - 5800\ \text{MeV}/c^2$, provided in Figure 6.3. For this analysis, the mass range includes the partially reconstructed three-body background in the low mass region. The profile of the partially reconstructed background is determined from MC distributions, detailed in Section 5.4.2.3, is observed to overlap with the low mass tail of the primary $B_d^0 \to K^+\pi^-$ signal peak. Thus, its inclusion is required to model the contamination expected from this background channel.

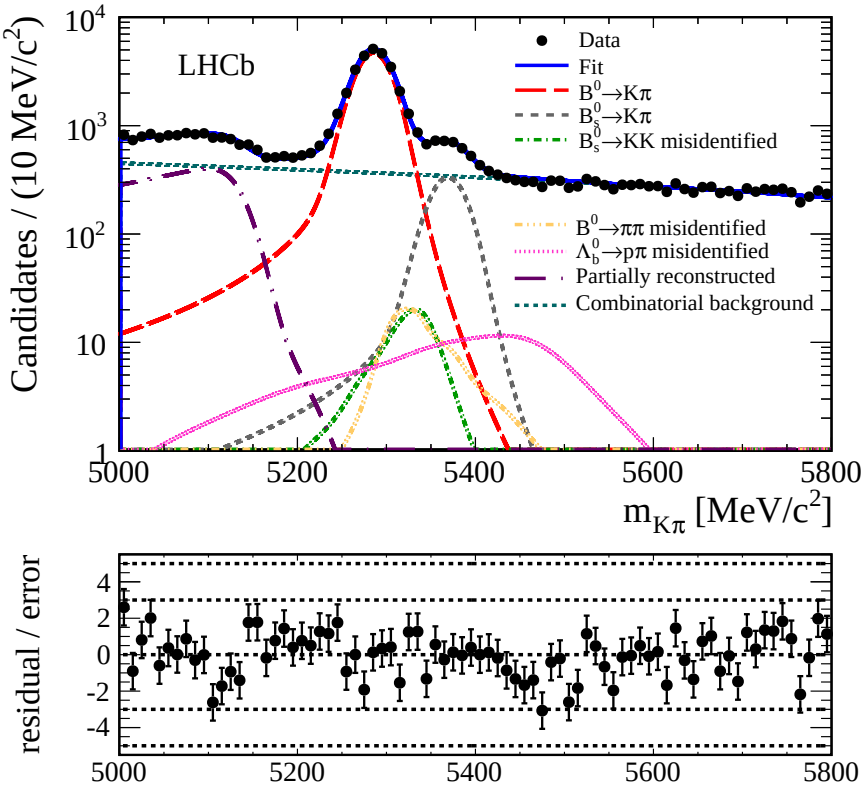


Figure 6.3: Mass distribution and mass fit of $B_d^0 \to K^+\pi^-$ using the 2011 fitter. The upper plot shows the fit to the mass spectrum in a log scale. The lower plot shows the binned residuals of the fit.

The fit to the $B_s^0 \to K^+K^-$ invariant mass spectrum covers the same range as the $B_d^0 \to K^+\pi^-$ fit, due to the reasons detailed previously with respect to the partially reconstructed background. As discussed in Section 5.4.2, the fit includes the dominant misidentified backgrounds, one specific background, partially reconstructed three body decays and the combinatorial. The result of the fit is given in Figure 6.4 and the resultant signal fractions and yields are provided in Table 6.2.

Further information regarding the parameters from the invariant mass fits to the $B_d^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ datasets, can be found in Appendix C.

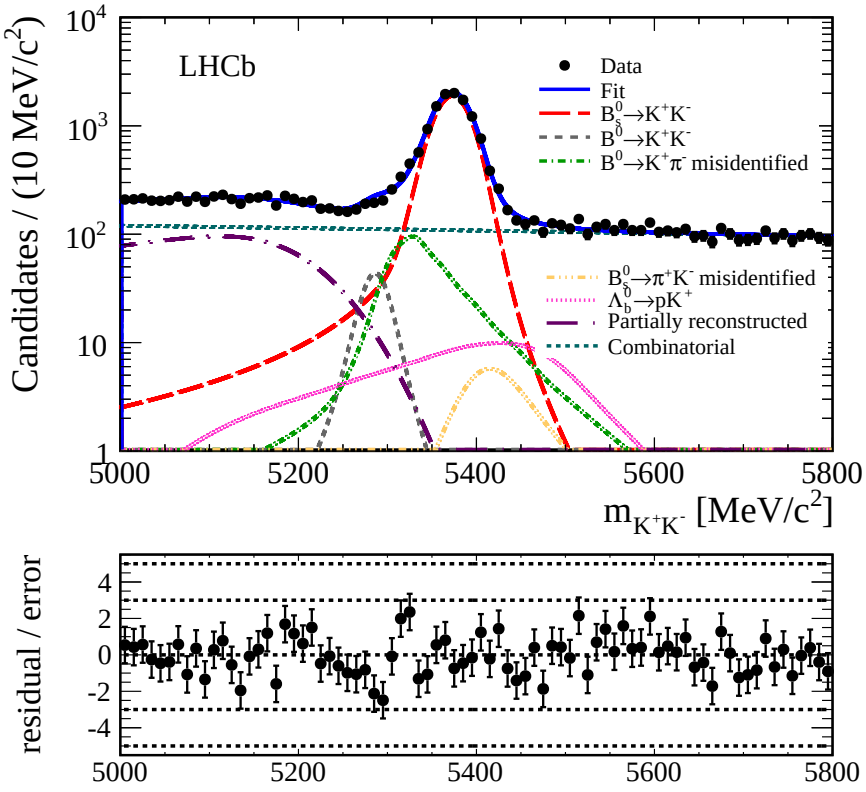


Figure 6.4: Mass distribution and mass fit of $B_s^0 \to K^+K^-$ using the 2011 fitter. The upper plot shows the fit to the mass spectrum in a log scale. The lower plot shows the binned residuals.

Signal Channel	Fitted Signal Fraction	Total events in data sample	Signal Yield
$\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$	0.4654 ± 0.0044	22498 ± 150	10471 ± 121
$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$	0.4327 ± 0.0028	60596 ± 246	26220 ± 200
$\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$	0.0312 ± 0.0014	60596 ± 246	1891 ± 85

Table 6.2: Event Yield for 2011 datasets. The signal fraction is calculated in the full mass range of $5000 - 5800\,\mathrm{MeV}/c^2$.

6.2.3 Search for $\mathrm{B}^0_{\mathrm{d/s}} \to \mathrm{p\bar{p}}$ 2011

The selection of $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ events also lends itself to another analysis, the search for the rare baryonic B-decay, $\mathrm{B}_d^0 \to \mathrm{p\overline{p}}$ [106]. The primary objectives for this analysis were the search for the $\mathrm{B}^0_{\mathrm{d/s}} \to \mathrm{p\bar{p}}$ signal and either make a first observation or set new, improved, upper limits for them. The $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ is a suitable choice for the normalisation channel, which allows an accurate extraction of the $\mathrm{B}_d^0 \to \mathrm{p\overline{p}}$ branching faction through

$$\mathcal{B}(\mathrm{B}_d^0 \to \mathrm{p\overline{p}}) = \frac{N(\mathrm{B}_d^0 \to \mathrm{p\overline{p}})}{N(\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-)} \cdot \frac{\epsilon_{\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-}}{\epsilon_{\mathrm{B}_d^0 \to \mathrm{p\overline{p}}}} \cdot \mathcal{B}(\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-), \tag{6.1}$$

where $\mathcal{B}(\mathrm{B}_d^0 \to \mathrm{p\overline{p}})$ and $\mathcal{B}(\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-)$ are the respective branching fractions, $N(\mathrm{B}_d^0 \to \mathrm{p\overline{p}})$ and $N(\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-)$ are the respective yields and $\epsilon_{\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-}$ and $\epsilon_{\mathrm{B}_d^0 \to \mathrm{p\overline{p}}}$ are the respective total efficiencies. The $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ channel is chosen as this mode provides large statistics, and is topologically identical and kinematically similar. The branching fraction is also very precisely known (Appendix B). The data is processed through the same selection criteria as that for the 2011 lifetime analysis, with the main difference being the slight alteration of the trigger conditions applied. The trigger conditions for the $\mathrm{B}_d^0 \to \mathrm{p\overline{p}}$ includes an additional trigger, which accepts events from an inclusive 2-body topological trigger.

An independent mass fitter was developed and verified for the $\mathrm{B}_d^0 \to \mathrm{p\overline{p}}$ analysis, with a fit to the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ selected invariant mass spectrum given in Figure 6.5. The full table of parameters for the fit can be found in Appendix C, with the signal fraction yields provided in Table 6.3.

The signal yield of the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ normalisation channel has contributed to the first evidence for the $\mathrm{B}_d^0 \to \mathrm{p\overline{p}}$, with a statistical significance of 3.3σ [107]. No evidence was found for the $\mathrm{B}_s^0 \to \mathrm{p\overline{p}}$, but the limit on its branching ratio was improved by three orders of magnitude. Measurements of the branching fractions from data for both the $\mathrm{B}_d^0 \to \mathrm{p\overline{p}}$ and $\mathrm{B}_s^0 \to \mathrm{p\overline{p}}$ are

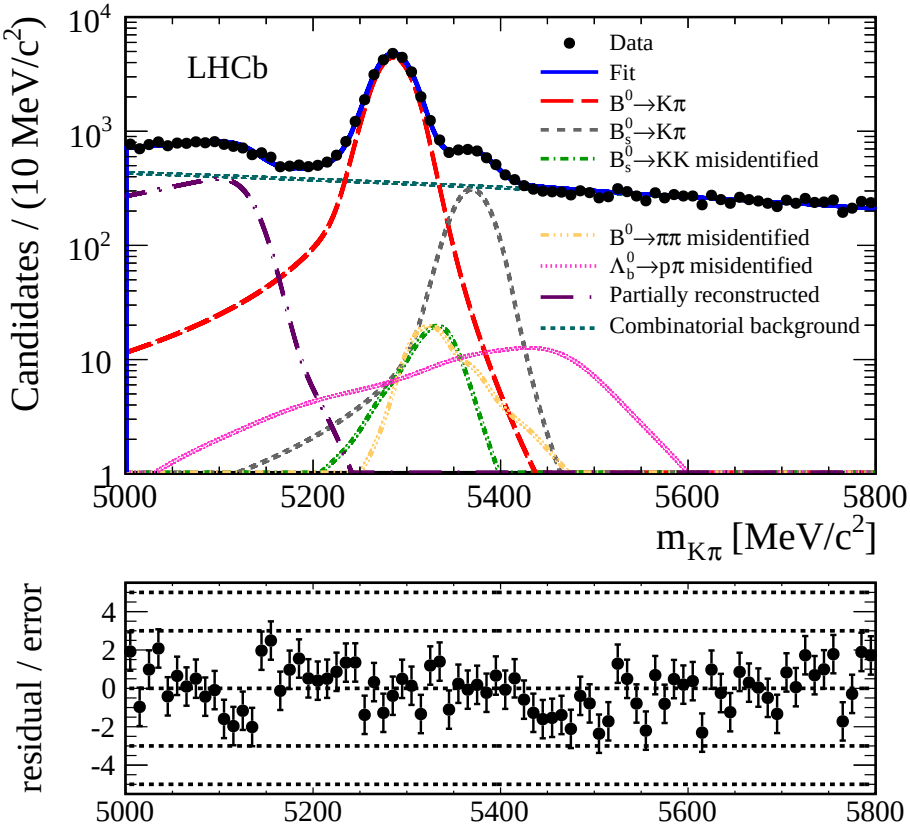


Figure 6.5: Mass fit to the $B^0_d \to K^+\pi^-$ normalisation channel displayed in the upper plot using a log scale [107]. The lower plot shows the binned residuals of the fit divided by the fit error. The parameters are extracted from an unbinned maximum likelihood fit, the binning of the data is only for illustration.

Signal Channel	Fitted Signal Fraction	Total events in data sample	Signal Yield
$\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$	0.4304 ± 0.0029	58009 ± 241	24967 ± 198
$\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$	0.0305 ± 0.0014	58009 ± 241	1771 ± 82

Table 6.3: Event Yield for $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ control channel using 2011 LHCb data in the $\mathrm{B}^0_{\mathrm{d/s}} \to \mathrm{p\bar{p}}$ analysis.

$$\mathcal{B}(\mathrm{B}^0_d \to \mathrm{p\overline{p}}) = (1.47\,^{+0.62}_{-0.51}\,^{+0.35}_{-0.14}) \times 10^{-8}, \tag{6.2}$$

$$\mathcal{B}(\mathrm{B}^0_s \to \mathrm{p\overline{p}}) = (2.84\,^{+2.03}_{-1.68}\,^{+0.85}_{-0.18}) \times 10^{-8}. \tag{6.3}$$

6.3 Studies of Systematic Uncertainties

As well as the statistical errors on the fitted lifetime, systematic uncertainties must also be accounted for in each measurement. In general, the systematic uncertainties arise from a variety of sources, these can include incomplete calibration of measurement instruments, changes in the environment that interfere with the measurement process and imperfect methods of observation. In LHCb analyses, the potential sources of systematics are many, however these can be constrained and are dependent on the analysis being performed. As both the 2010 and 2011 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analyses are similar, a large number of these systematics will apply to both analyses. Due to improvements in the reconstruction software, detector performance, etc, between data taking periods, many of these systematics can be simply rerun and updated. However, differences between the datasets also require new systematic studies to be designed, these include effects from the detectors, designated models and the fit method. The systematic studies are generally studied using two methods: toy MC simulation (Section 5.6.1) and data. The toy MC datasets are composed of different signals and statistics dependent on the analysis. Unless otherwise specified, the following conditions are used. The 2010 toy MC $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ datasets contain one signal and the combinatorial background. These are generated with 400 events consisting of 70% signal and 30% background. In the 2011 studies, there are more signal channels to incorporate. The $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ toy MC datasets are composed of: 45% $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, 2% $\mathrm{B}^0_d \to \mathrm{K}^+\mathrm{K}^-$, 2% $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$, 2% $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$, 2% $\Lambda_b \to p\mathrm{K}^-$, 5% partially reconstructed background and 42% combinatorial background. Each independently generated $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ data set contains a total of 25,000

events similar to that observed in data. The $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ toy MC datasets are comprised of: 42% $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$, 5% $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$, 2% $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, 2% $\mathrm{B}^0_d \to \pi^+\pi^-$, 2% $\Lambda_b \to p\pi^-$, 5% partially reconstructed background and 42% combinatorial background. Each independently generated $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ dataset contains a total of 50,000 events similar to that observed in data. The relevant systematics are discussed in detail in the following subsections, with some additional consistency checks provided in Appendix E. A full summary of the uncertainties from each systematic study, for each analysis, is provided in Section 6.3.9.

6.3.1 Systematic Uncertainties from the Fitter

The uncertainty on the lifetime from the fitter is determined from the precision to which it was verified, Section 5.6.1.4. The total number of statistically independent toy MC datasets used for the 2010 measurement is 1000, hence the statistical uncertainty of the mean and σ of the pull distributions is approximately 0.033. For the 2011 measurements, the precision on this statistical uncertainty is improved upon by using 5000 toy MC datasets, where this uncertainty is found to be approximately 0.011. The conclusion is made that the verifications are done to the level of 3.3% of the statistical uncertainty for the 2010 measurement, and 1.1% for the 2011 measurements. In addition, the verification fits to the $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$, shown in Figure 5.12 and provided in Table 5.16, are observed to have underestimated statistical uncertainty. As such, for $\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-}$ the uncertainty from the fitter, and hence the final fitted statistical uncertainty, is scaled up by 7% to account for this effect. Thus, the uncertainty on the 2010 $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime analysis is $3.2\,\mathrm{fs}$, and the uncertainties for the 2011 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime measurements are $0.17\,\mathrm{fs}$ for $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $0.11\,\mathrm{fs}$ for $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $0.59\,\mathrm{fs}$ for $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$.

6.3.2 Systematic Uncertainties from Mass Models

The choice of models used to describe the invariant mass have an impact on the final value of the fitted lifetime. The uncertainty based on these models differs between analyses due primarily to the different models implemented between datasets.

For the 2010 lifetime measurement, the models used are detailed in Sections 5.4.1 and 5.5.1. The mass model applied to the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal channels is a Gaussian, this is a suitable model to use due to the low number of events selected in the 2010 dataset. This model does not incorporate radiative effects which reveal themselves as a tail component in the signal profile. To determine the uncertainty on the lifetime fit due to this, the mass model is changed to a single crystal ball function and refitted on the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ control channel data. The difference between the lifetimes fitted using each model is then assigned as the uncertainty

from this source, with the uncertainty for the 2010 $B_s^0 \to K^+K^-$ lifetime found to be $1.9\,\text{fs}$. The method used to determine the uncertainties from the mass models in the 2010 analysis is altered for the 2011 analysis to deal with the updated, and additional, mass models implemented for the $B_s^0 \to K^+K^-$ and combined $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ datasets. The mass models assigned to the signal, misidentified and partially reconstructed backgrounds are described by profiles extracted from full LHCb MC, described in Section 5.4.2. Uncertainties associated with the mass models are determined by varying each signal mass model independently. For the $B \to h^+h^-$ signal and partially reconstructed background models, described by a double crystal ball and exponentially modified Gaussian respectively, the models are tested by varying the fixed parameters within 5σ of their uncertainties. The misidentified backgrounds are described by non-parametric template distributions (Figure 5.7) with their locations in the invariant mass spectrum determined by the fitted mean of the largest signal peak. The uncertainty from these models is determined by offsetting the position of the background by the error on the fitted $\mu_{B_s^0}$ or $\mu_{B_d^0}$, as the misidentified background is located with respect to the signal mean. The combinatorial background is treated slightly differently, as the model used is unaltered between the 2010 and 2011 analyses. The combinatorial mass is modelled using a linear function, see Section 5.4.1.4, however it could also be modelled using an exponential distribution. This uncertainty is determined by fitting the combinatorial background on each dataset with both linear and exponential functions, then using the difference in fitted lifetimes as the uncertainty from this source. The total uncertainties from these sources for the 2011 lifetime measurements are $1.10\,\text{fs}$ for $B_s^0 \to K^+K^-$, $2.51\,\text{fs}$ for $B_d^0 \to K^+\pi^-$ and $6.74\,\text{fs}$ for $B_s^0 \to \pi^+K^-$.

6.3.3 Systematic Uncertainties from Lifetime Models

The methods used to separate and reconstruct the lifetime distributions for each signal channel spawn a number of possible sources of uncertainty. Primarily these are related to non-parametric models that are used to describe both partially reconstructed and combinatorial backgrounds. However, the discussion will also include sources of uncertainty from parametric models, used to describe the $B \to h^+h^-$ signal channels.

6.3.3.1 Parametric Lifetime Model Uncertainties

In the 2010 and 2011 analysis of $\tau_{B_s^0 \to K^+K^-}$, the parametric model used to describe the lifetime distribution is a single exponential function described in Section 5.5. This is a suitable description for the $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$. However, a small non-negligible amount of CP-violation is predicted for the $B_s^0 \to K^+K^-$ system, caused by the finite decay width

difference $\Delta\Gamma_\mathrm{s}$ that is described in Section 1.4. Thus, there are technically two exponential contributions to the total lifetime distribution of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$. The relative contribution of each exponential to the total lifetime is determined via the parameter $\mathcal{A}_{\Delta\Gamma}$, Equation 1.43, which quantifies the level of CP-violation predicted. If there was no CP-violation predicted, then the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ would be purely from the $|B_L\rangle$ mass eigenstate. However, there is a small contribution from the $|B_H\rangle$ mass eigenstate giving the lifetime distribution a mixture of these two states. This source of uncertainty is also called the 'effective lifetime bias'.

To determine how accurately the fit to the reconstructed lifetime spectrum extracts the 'effective' lifetime of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ signal channel, an additional model that includes contributions from both $|B_L\rangle$ and $|B_H\rangle$ dependent on the $\mathcal{A}_{\Delta\Gamma}$ parameter, is added in the toy MC simulation code (Section 5.6.1). The $\mathcal{A}_{\Delta\Gamma}$ parameter is then set to the values: -0.6, -0.8, -0.97, -1.0 and +1.0. The lifetime fits are then run to determine if the extracted 'effective' lifetime is biased, due to the inclusion of the double exponential model with different rate contributions. The uncertainty determined from the studies was $2.8\,\mathrm{fs}$ for the 2010 analysis and $1.20\,\mathrm{fs}$ for the 2011 analysis. It should be noted that the modelling of the effect was improved between analyses, which contributes to the decrease in uncertainty.

The statistics accumulated in the 2011 data, moves the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ analysis into the regime where contributions from B_c^+ decays become relevant. Studies of the $\mathrm{B}_c^+ \to \mathrm{B}_s^0\pi^+$ system have shown that up to 1% of all B_s^0 decays originate from B_c^+ decays [108]. This additional effect can cause the measured lifetime to shift higher, if not properly accounted for and understood. The uncertainty from the B_c^+ contribution is modelled in the toy MC generator by adding an additional lifetime contribution to represent the B_c^+ decays. The decay time for the 1% of decays that are assumed to originate from B_c^+ decays is modelled as the sum of two exponential distributions. First an exponential distribution with the average B_c^+ lifetime, set to the PDG [22] value of $0.453\,\mathrm{ps}$, is sampled. This is followed by a second exponential that generates a decay time for the B_s^0. The uncertainty is determined by comparing the lifetime fits with and without this B_c^+ lifetime contribution, and is found to be $1.12\,\mathrm{fs}$.

A parametric model uncertainty arises from the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ misidentified signals that are only found in the 2011 analysis. Fixed models are assigned for the misidentified $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ channels, as detailed in Section 5.4.2.2. The parameters for the exponential models employed for these backgrounds, are fixed to the world averages from the PDG [22]. The fixed parameters of the models are varied within the uncertainty on their values, to determine the magnitude of any lifetime shift. The results of the systematic study from these misidentified sources yield the uncertainties: $0.70\,\mathrm{fs}$ for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $0.40\,\mathrm{fs}$ for $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $3.60\,\mathrm{fs}$ for $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$.

The total systematic uncertainties for the 2011 analysis associated to the parametric uncer-

tainties discussed are: $1.78\,\text{fs}$ for $B_s^0 \to K^+K^-$, $0.40\,\text{fs}$ for $B_d^0 \to K^+\pi^-$ and $3.60\,\text{fs}$ for $B_s^0 \to \pi^+K^-$.

6.3.3.2 Non-Parametric Lifetime Model Uncertainties

As with the parametric models, non-parametric models also include systematic effects that are equally relevant to both the 2010 and 2011 lifetime analyses. However, further studies are included for the 2011 analyses due to the inclusion of additional non-parametrically described backgrounds. The discussion will begin describing the systematics studies required for both analyses, followed by more specific studies of the 2011 lifetime measurements.

There are two relatively simple studies that can be performed for all lifetime analyses in relation to the method for reconstructing the background lifetime distributions. The first systematic study relates to the kernel width, while the second study determines the effect of varying a parameter, called the ϵ parameter, which is used to remove negative values of the background PDFs that can occur.

The kernel method, detailed briefly in Section 4.5.2 and further in [88], effectively smoothes the reconstructed decay time distribution of the backgrounds with the width of the smoothing function defined by Equation 4.24. The systematic uncertainty from this source is determined by altering the width of the distribution to double and half of its standard value. The uncertainty from this source is determined to be negligible for the 2010 lifetime analysis, with the values for the 2011 analysis found to be $0.78\,\text{fs}$ for $B_s^0 \to K^+K^-$, $0.63\,\text{fs}$ for $B_d^0 \to K^+\pi^-$ and $6.36\,\text{fs}$ for $B_s^0 \to \pi^+K^-$.

The ϵ parameter is a small value that is included in the fit method to correct for negative values of the background PDF that can occur. These negative values can arise due to statistical fluctuations or incorrect fit values for the signal lifetime, and can contribute to stability issues in the fit. The value of ϵ is tuned to a small positive value to remove these effects. The systematic for this is determined by setting the value of ϵ to an order of magnitude higher and an order of magnitude lower than the default value of 0.001 and observing any shift in the fitted lifetime. For both the 2010 and 2011 lifetime measurements the observed uncertainty is deemed to be negligible.

The *sWeights* method, Section 4.3.3.2, used to differentiate the signal channels in the 2011 analysis explicitly assumes that there is no correlation between mass and decay time. This also has to be accounted for in the 2010 analysis, although Bayesian probabilities, Section 4.3.3.1, are used instead of *sWeights*. This correlation is tested in the combinatorial background using two distinct techniques. The first, applied in the 2010 analysis, removes the $200\,\text{MeV}/c^2$ region in the range $5600 - 5800\,\text{MeV}/c^2$ from the $B_s^0 \to K^+K^-$ dataset. The fit is then performed with and without the selected region, with the difference found to be

1.5 fs and assigned as the uncertainty.

The 2011 analysis takes a slightly different approach. The invariant mass range $5480 - 5880\,\mathrm{MeV}/c^2$ is selected, Figure 6.6 (top), and then divided into three bins of lifetime, shown in Figure 6.6 (bottom). Each bin has approximately the same number of events, with any mass-decay time correlation being observed if the distributions noticeably differ. The decay time distributions in the three mass regions are observed to be identical, hence no systematic uncertainty is assigned.

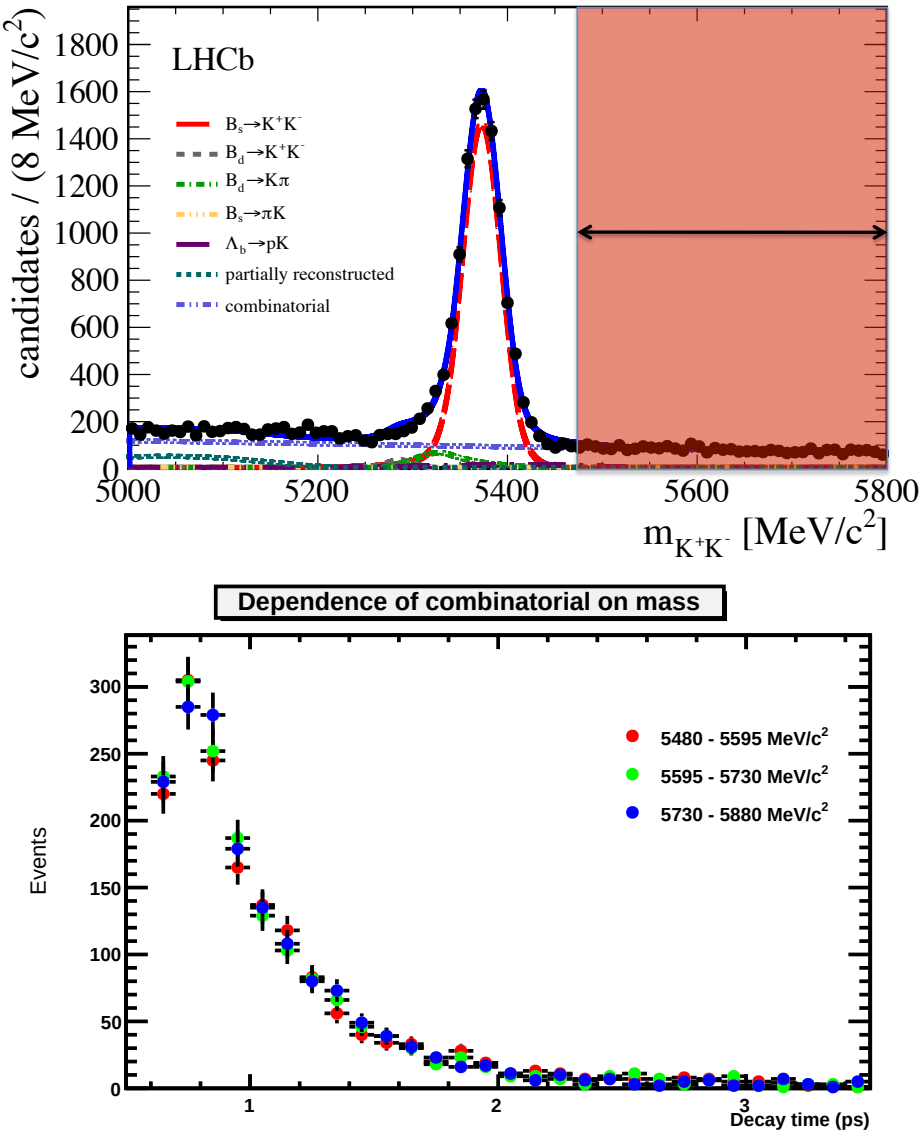


Figure 6.6: Top: Upper side band of the combinatorial background from the 2011 $B_s^0 \to K^+K^-$ dataset. Bottom: Decay time distribution from $B_s^0 \to K^+K^-$ combinatorial background in the invariant mass range $5480 - 5880\,\mathrm{MeV}/c^2$. The decay time has been split into 3 bins containing 1650 events in each.

The remaining backgrounds include the partially reconstructed and misidentified $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ backgrounds. These need to be studied to determine if they have a mass-decay time correlation. Although the misidentified backgrounds are modelled parametrically, for reasons of simplicity they are included in this particular systematic study. To determine if there is an effect and its probable magnitude, the correlation is modelled in the toy MC generator, Section 5.6.1. To add the correlation effect to the generated datasets, the decay time smearing

$$\tau_{corr} = \tau_{gen} \times \frac{mass_{gen}}{mass_{true}} \tag{6.4}$$

is added, where τ_{corr} is the new correlated decay time, τ_{gen} the original generated decay time determined from an exponential distribution, $mass_{gen}$ which is the mass generated from the mass model description and $mass_{true}$ the true mean value of the signal mass. One toy MC dataset is generated with no correlation. A second is then reproduced from this original dataset, with the correlation given by Equation 6.4 applied to the original decay time per event dependent on the event signal class. The difference between the fitted lifetimes of the two datasets provides a 'worst case' uncertainty for this systematic source. The results of the studies show the uncertainties to be $0.25\,\mathrm{fs}$ for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $1.44\,\mathrm{fs}$ for $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $2.10\,\mathrm{fs}$ for $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$.

The total systematic uncertainties for the 2011 analysis associated to the non-parametric uncertainties discussed are: $0.82\,\mathrm{fs}$ for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $1.57\,\mathrm{fs}$ for $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $6.70\,\mathrm{fs}$ for $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$.

6.3.4 Uncertainties related to the VELO

The VELO makes a vital contribution to lifetime measurements, due to its excellent position resolution. The primary method for accounting for the acceptance introduced through the lifetime biasing cuts is via the swimming method, detailed in Section 4.4. However, another source of acceptance effects is the VELO. There are a number of ways the VELO can contribute to limit the acceptance, these will be individually discussed in the following subsections.

6.3.4.1 Finite length of the VELO

One effect that can lead to a reduced acceptance for events at high decay times, is due to the finite length of the VELO. The daughters of B events that have long flight distances may not register the required number of 'hits' in the VELO needed for reconstructed (three hits for the offline reconstruction and five hits for the trigger reconstruction). This effect is not accounted for in the swimming as it is the PV that is moved along the momentum vector of

each candidate, so the tracks are not re-reconstructed at any point. However, the number of hits can be determined analytically by computing the number of VELO stations each track passes through, based on the position of the B vertex for a given swimming step and track direction. Together with the per-event acceptance function determined from the selection and trigger, this method also produces a per-event VELO acceptance function, which succeeds only if the B daughters pass through three VELO stations (offline selection requirement) and five VELO stations (trigger requirement). It was found that the VELO acceptance effect is negligible for both 2010 and 2011 analyses.

6.3.4.2 Radial dependence on reconstruction efficiency

The effects of a high decay time reconstruction efficiency is studied by introducing a radial flight distance cut around the beam axis. As the inner aperture of the VELO is $5.5\,\mathrm{mm}$, this automatically introduces a limit where we do not accept events that have traversed the RF-foil. This study was developed primarily for the 2011 analysis as there were almost no B_s^0 mesons found in the affected region using the 2010 dataset [75], hence the effect was assumed to be negligible for that analysis.

To determine the significance of this effect, the radial cut is implemented in the fitter to remove events which fail these cut criteria. This is tested using both a fully swum $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ dataset and a toy MC simulation. Each event in the dataset is assigned a pass/fail binary value and an additional turning point, $\tau_{radialcut}$, as this cut limits the acceptance. The datasets are simplified by fitting only for one signal and the combinatorial background, with the cut varied between 0.5 and $6.0\,\mathrm{mm}$. The distributions of the fitted lifetime after each radial cut for data and simulation are provided in Figure 6.7.

The fitted lifetimes for the toy MC show that with the additional turning point the lifetime is constant, however, the data displays an effect below $2.0\,\mathrm{mm}$. This effect observed in the data is not fully understood, however since we apply a cut at $4.0\,\mathrm{mm}$ on data this effect should be avoided. Rejecting events with a large radial flight distance should reduce the decay time dependent track reconstruction efficiency. Since we observe no lifetime shift at a radial cut of $4.0\,\mathrm{mm}$, no uncertainty is given to this source. The actual number of events rejected from each dataset due to this cut are 59 for the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and 174 for the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$.

6.3.4.3 Tracking efficiency acceptance

In addition to the previous two effects just described, it can also be shown that the VELO track reconstruction imposes an additional decay-time acceptance unconnected to the VELO length. Essentially the track finding algorithms prefer tracks which come from the beamline

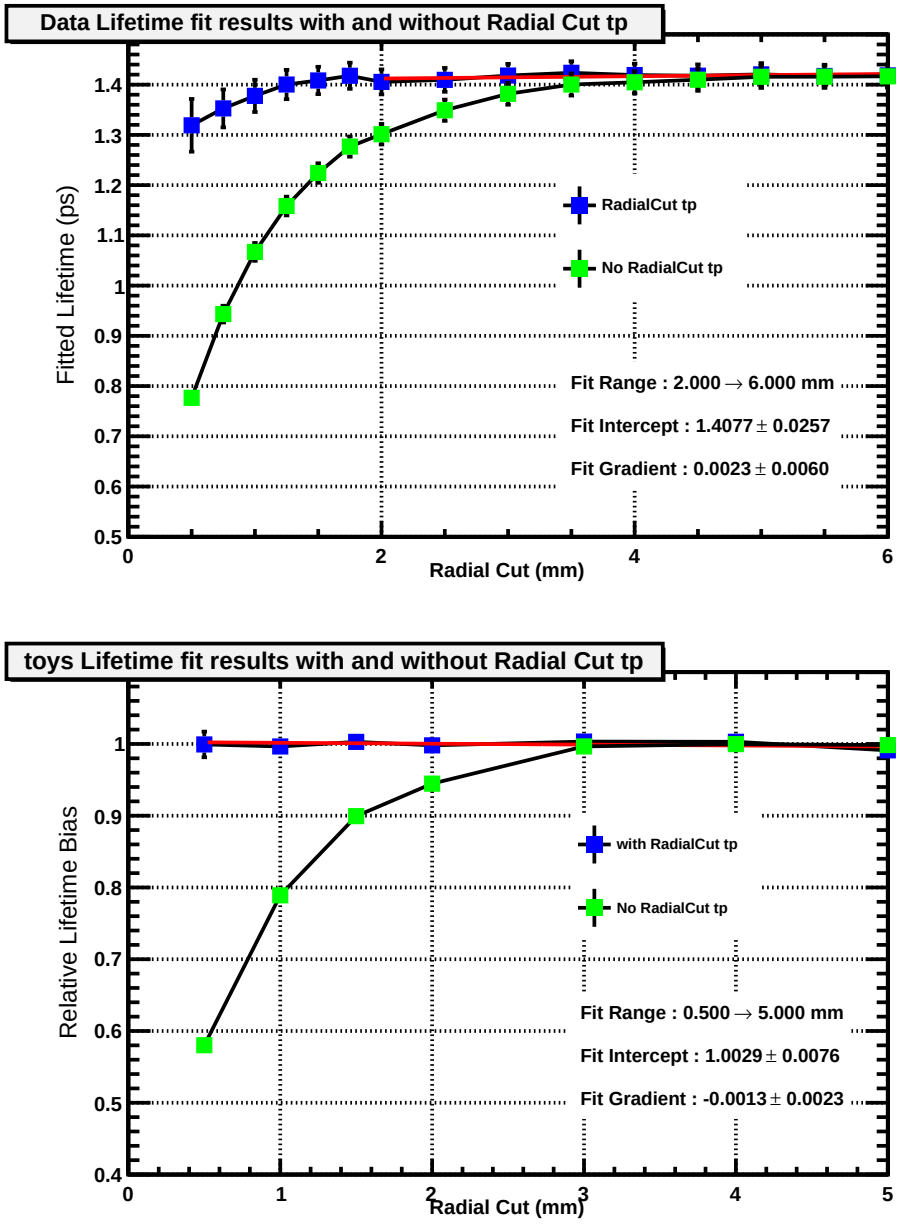


Figure 6.7: Effects of application of the radial cut using 2011 data and toy MC. Top: Fitted lifetime values, using a relative scaling factor (blinded), from each radial cut on $B_s^0 \rightarrow K^+K^-$ 2011 data. Bottom: Ratio of fitted lifetime to true lifetime using toy MC simulated radial cut. Additional turning points derived from the radial cut are applied (blue markers), and not applied (green markers), to both data and toy MC simulation.

("on-axis") and, as B mesons are produced with transverse momentum and fly radially before decaying, tracks originating from long-lived B decays point away from the beamline ("off-axis"). These tracks may have a lower efficiency to be reconstructed. This effect leads to a reduction in the reconstruction efficiency at high decay times: an upper decay time acceptance. This is another effect that can limit the acceptance that is is not accounted for in the swimming method, and therefore needs to be studied independently. Similar to the radial dependence study, this study was developed primarily for the 2011 analysis as there were almost no B_s^0 mesons found in the high lifetime region where this effect would be seen [75], hence the effect was assumed to be negligible for the 2010 analysis. The track reconstruction efficiency is parameterised by a quadratic formula

$$\epsilon(\mathrm{DOCAZ}) = \alpha + \beta * \mathrm{DOCAZ}^2, \tag{6.5}$$

where DOCAZ is the distance of closest approach to the z-axis for a track. The parameters α and β are determined from data and are valid for all low multiplicity events. Equation 6.5 can also be parameterised with respect to the decay time using the relationship given by

$$\mathrm{DOCAZ} = \gamma + \delta \cdot t, \tag{6.6}$$

with the parameters γ and δ determined for each event. It then follows that, for each event, the dependence of the efficiency on the decay time is

$$\epsilon(t) = \alpha + \beta \cdot (\gamma + \delta \cdot t)^2. \tag{6.7}$$

The decay time distribution including the efficiency $\epsilon(t)$ is shown in Figure 6.8, over the decay time range accepted by the fitter ($0.61\,\mathrm{ps}$ to $10\,\mathrm{ps}$) in the 2011 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analysis.

The distribution of $\epsilon(t)$ shows a decrease over the decay time range of the fitter, Figure 6.8. To determine the magnitude of this effect, the efficiency distribution is simulated in the toy MC generator. The efficiencies are generated per-event from Equation 6.7, with a failure to reconstruct the variable assigned by generating a random number between 0 and 1. The event is rejected if the efficiency is found to fall below this randomly generated value. A fit is performed on 1000 toy MC datasets containing 100,000 events each, comprising a mix of 50% signal and 50% combinatorial background.

Since the parameters describing the $\epsilon(t)$ distribution are extracted from data, the result of the fits can be used to account for this effect. The value determined from the toy MC gives a $5.65\,\mathrm{fs}$ bias on the fitted lifetime, as such, this is added to the final fit value from data. A value of half the observed bias, $2.83\,\mathrm{fs}$, is assigned as the uncertainty.

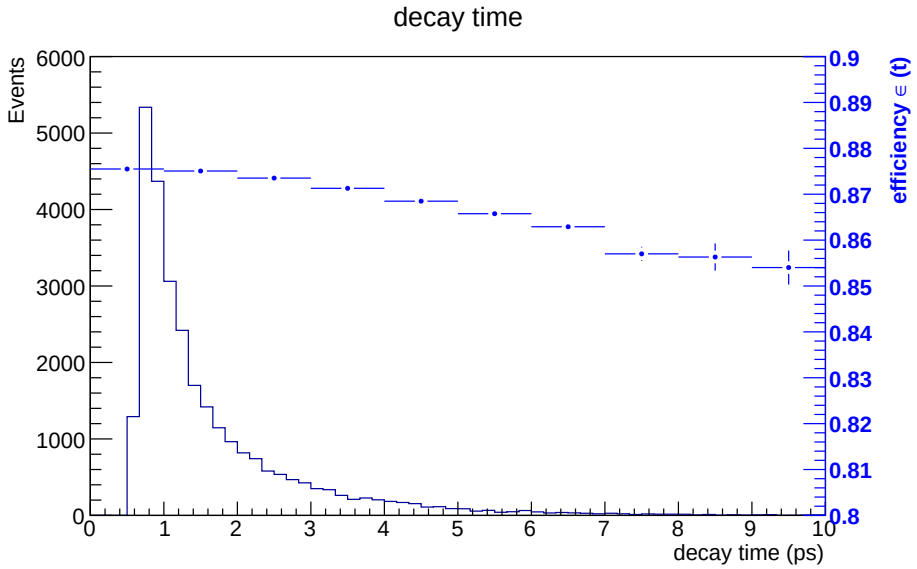


Figure 6.8: Decay-time distribution of $B_s^0 \to K^+K^-$ 2011 data (histogram distribution) with the corresponding $\epsilon(t)$ efficiency (data points with errors).

6.3.4.4 Uncertainties from Acceptance Correction

Due to the limited statistics in the 2010 dataset a reduced set of specific systematics was evaluated. Instead, a global systematic on the acceptance correction was derived from D decays. To determine the potential uncertainty associated with the acceptance correction, $D \to h^+h^-$ data from the same accumulated dataset is used. These data are used as it has very high statistics compared with the $B \to h^+h^-$ dataset for the same period, and is topologically similar. An average of the measured $D^0 \to K^-\pi^+$ and $\overline{D}^0 \to K^+\pi^-$ lifetimes [109] is then compared with the world average lifetime and found to be compatible [22]. The statistical uncertainty of the two measurements is then added in quadrature, before multiplying by the ratio of the two lifetimes ($\frac{\tau_{B_s^0}}{\tau_{D^0}}$) to provide an estimate of the maximum uncertainty on the 2010 $B_s^0 \to K^+K^-$ lifetime. This uncertainty for the 2010 lifetime measurement is determined to be $6.3\,\text{fs}$.

6.3.4.5 Uncertainties from Primary Vertex Assignment

Following on from Section 6.3.4.4, it is noted that the acceptance correction relies on the accurate reconstruction of the primary vertex. As the number of collisions at the interaction point increases, so does the event multiplicity. This can cause events to have multiple reconstructed primary vertices, with incorrectly assigned ones likely to occur at an unknown rate. For the smaller dataset size of the 2010 analysis, it is difficult to examine this effect so again $D \to h^+h^-$ data is used for the reasons given in Section 6.3.4.4. A fit to the lifetime of the

D^0 and $\overline{D}^0$ is performed on a sub-set with one, two and three primary vertices. Any observed difference can then be scaled with $\frac{m_B}{<p_B>} / \frac{m_D}{<p_D>}$ to show the effect for the B mesons. The uncertainty for the 2010 lifetime analysis is found to be $1.2\,\mathrm{fs}$.
For the 2011 analysis, the dataset is large enough to study this effect on the $B_s^0 \to K^+K^-$, $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ lifetimes independently. The KK and Kπ datasets are fitted using a blinded fit and binned with respect to four categories. The four categories chosen are: decays with one, two, greater than two but less than ten, and smaller than ten primary vertices (the default cut used in the analysis). The resultant fits of the three lifetimes, for each category, are shown in Figure 6.9. As the fitted lifetimes for each category, for each decay, are consistent within the uncertainties, no systematic values are assigned to this source.

6.3.4.6 Absolute decay time scale uncertainty

The final VELO uncertainty to be discussed, is the absolute decay time scale. The uncertainty on the absolute decay time scale has more than one source. The dominant source is related to the knowledge of the absolute value of the decay time. This has two main contributors. First, the imperfect knowledge of the longitudinal (z) scale of the detector contributes to the systematic uncertainty. It is obtained by comparing the track-based alignment and survey data and evaluating the track distribution in the vertex detector. The second contribution comes from the knowledge of the overall momentum scale. This has been evaluated using mass measurements of well-known resonances. Deviations from the reference values [22] are then measured. Since both the measured invariant mass and momentum enter the calculation of the decay time, this effect is effectively cancelled. The resulting systematic on the decay time scale for the 2010 analysis is $1.5\,\mathrm{fs}$, with the 2011 analysis being $0.41\,\mathrm{fs}$ for $B_s^0 \to K^+K^-$ and 0.43 for $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$.

6.3.5 Cross Contamination due to Particle Misidentification

As well as significant sources of uncertainties originating from the VELO subdetector, the PID performance of the RICH subdetector also contributes to the total systematic uncertainty. The PID performance of the RICH is important as the $H_b \to h^+h^-$ stripping selection, Section 5.3.2, contains multiple signal channels that are kinematically similar to the $B \to h^+h^-$ signal channels being studied, but have different final states. Inefficiencies in the PID performance can cause contamination from specific $B \to h^+h^-$ channels to enter the event selection. The level of contamination expected cannot be determined by studying full LHCb MC alone, as the PID performance is seen to differ significantly between this and data. As such, a data driven method is used to determine a re-weighting scheme that can be applied

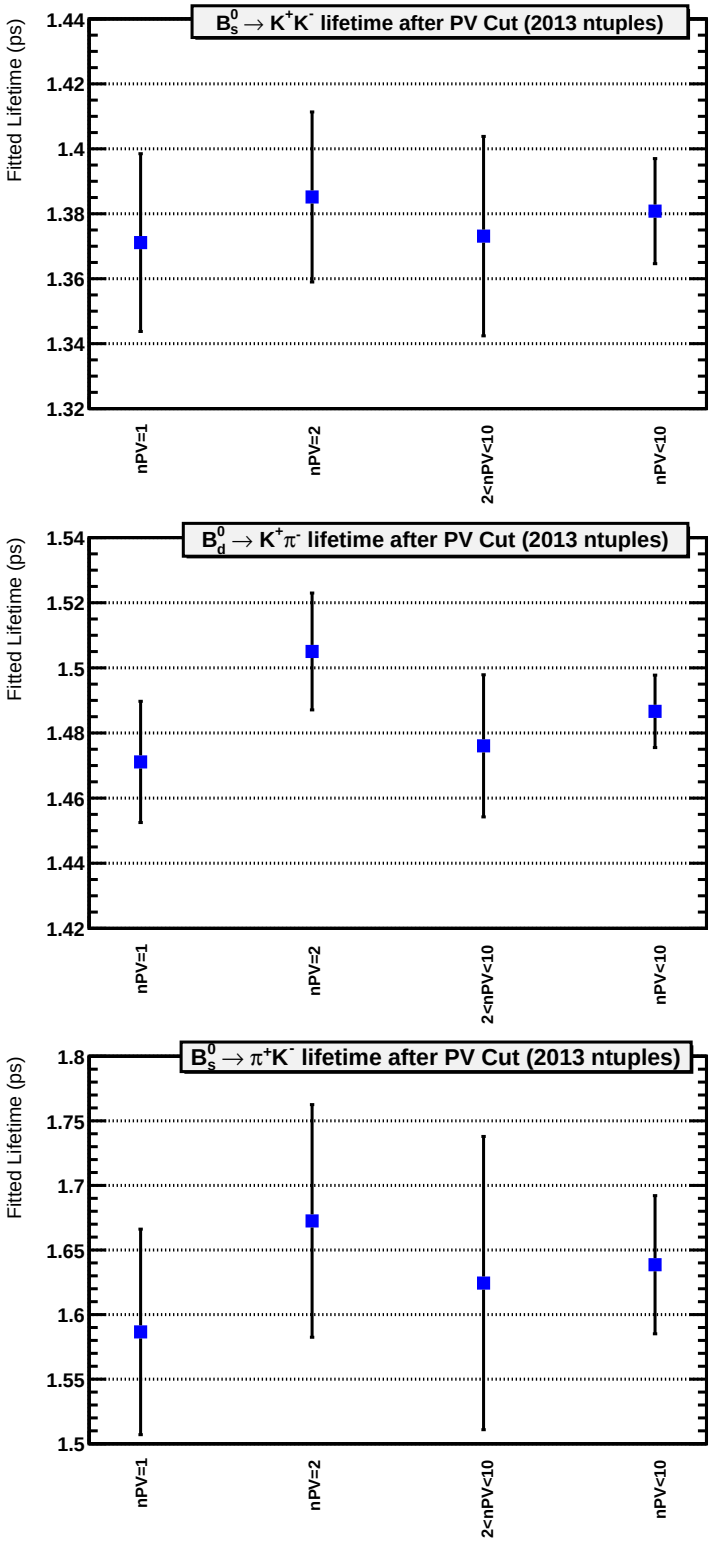


Figure 6.9: The fitted lifetime of $B_s^0 \to K^+K^-$ (top), $B_d^0 \to K^+\pi^-$ (middle) and $B_s^0 \to \pi^+K^-$ (bottom) as a function of number of primary vertices (nPV). The lifetime fitted using nPV < 10 is performed using the default fit configuration. The data fits are all performed blinded, with the same blinding factor, using the 2011 dataset.

to the MC PID distributions, to allow more accurate determinations of the misidentification rates of final state hadronic particles. This is used for both the 2010 and 2011 analyses. A reweighting method is performed by using a sample of calibration tracks for the RICH where the particle identification can be determined independently of the RICH information: pions are obtained from K^0_S, D and Λ decays; kaons are obtained from D and ϕ decays; protons from Λ decays. The calibration tracks are then binned into a set of variables on which the PID is dependent. These variables are the momentum, p, transverse momentum, p_T, and the number of primary vertices, n_{PV}, of an event. All tracks residing in a given bin will have approximately consistent RICH PID decisions that are determined from the calibration track sample. The PID mis-identification rates for any channel can then be determined by considering the event distribution across these characteristic variables, and reweighting the identification rates accordingly. The kinematic distributions of the two final state tracks in all $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays are similar, and hence the process does not need to be repeated for each signal or background. The percentage mis-identification of a final state hadronic particle is obtained by integrating the normalised DLL distribution of the reweighted RICH calibration data, for a given PID cut.

For the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ family of decays studied, a detailed list of decay channels that could contaminate is compiled and provided in Appendix B. The main contributors are expected to be the other $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ channels, that are collected in the stripping selection.

In the 2010 analysis, the mass model includes only two contributing signal modes: the signal and combinatorial background (Section 5.4.1), with the mass range chosen to remove partially reconstructed background contributions. It was found that the largest $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ contaminants to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ were the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ (2%) and $\mathrm{B}^0_d \to \mathrm{K}^+\mathrm{K}^-$ (1%) channels. These models were included into the toy MC generator to determine the value of the lifetime uncertainty from not including the models in the full fit, with the uncertainty found to be $1.9\,\mathrm{fs}$. A similar study was conducted for the partially reconstructed background distribution in this analysis. As this background model is not included due to expectation that this is excluded through the selected mass range, the potential signal overlap region needs to be studied for uncertainties. To estimate the proportion of partially reconstructed decays required in the toy MC dataset, the combinatorial background of the full $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ dataset was fitted in the upper mass region and extrapolated to the lower mass region, Figure 6.10 (left). The excess of data events in the lower mass region is then attributed to partially reconstructed decays. A profile for the partially reconstructed background events is then produced from MC, using a similar method to that used in Section 5.4.2.3, and used as the distribution in the toy MC simulations, Figure 6.10 (right). The uncertainty from this is deduced by including this model in the fit and determining the difference between fits. The

result of this was an uncertainty of $1.9\,\mathrm{fs}$.

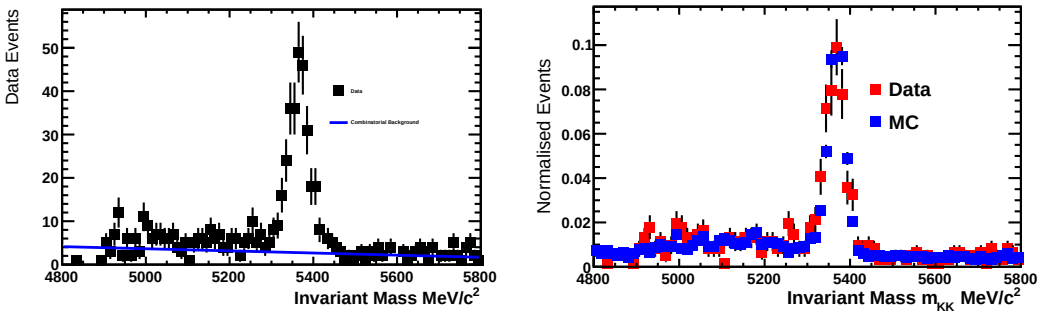


Figure 6.10: Estimation of partially reconstructed background from the 2010 $B_s^0 \to K^+K^-$ dataset. This is done by fitting to the mass band $5500 - 5800\,\mathrm{MeV}/c^2$ and extrapolating to the lower mass region (left). The normalised $B_s^0 \to K^+K^-$ 2010 data distribution compared with the constructed MC10 distribution with signal, combinatorial background and partially reconstructed events (right).

For the 2011 analysis, the contamination uncertainties for both the $B_s^0 \to K^+K^-$ and $B_d^0 \to K^+\pi^-$ datasets are minimised via the inclusion of misidentified signal models in the fit. The relative contamination, and their uncertainties, can be constrained using two distinct methods, which are be discussed.

The increase in dataset size inherently increases the yield of each potential contamination background, therefore any channel that could contribute at $> 0.5\%$ to the dataset is included with a full mass (Section 5.4.2.2) and lifetime model (Section 5.5.2). The misidentification rates from other $B \to h^+h^-$ modes can be estimated from the misidentification rates determined by the PID reweighting method described earlier, the branching ratios of the decays and the relative hadronisation fractions. The yields of the cross-contamination of the $B \to h^+h^-$ is calculated relative to the yield in the main signal peak in the two spectra and is given by

$$k^{B\to h^+h^-}_{B^0_d\to K^+\pi^-} = \frac{f_x}{f_d} \times \frac{\mathcal{B}(B \to h^+h^-)}{\mathcal{B}(B^0_d \to K^+\pi^-)} \times \frac{\omega^{B\to h^+h^-}_{K\pi}}{\epsilon^{B^0_d\to K^+\pi^-}_{K\pi}} \tag{6.8}$$

and

$$k^{B\to h^+h^-}_{B^0_s\to K^+K^-} = \frac{f_x}{f_s} \times \frac{\mathcal{B}(B \to h^+h^-)}{\mathcal{B}(B^0_s \to K^+K^-)} \times \frac{\omega^{B\to h^+h^-}_{KK}}{\epsilon^{B^0_s\to K^+K^-}_{KK}} \tag{6.9}$$

for the $K\pi$ and KK fits respectively. The relative hadronisation fraction is given by $\frac{f_x}{f_{d,s}}$, the misidentification rate is given by $\omega^{B\to h^+h^-}_{K\pi,KK}$ and the PID efficiency is given by $\epsilon^{B^0_d\to K^+\pi^-}_{K\pi}$ and $\epsilon^{B^0_s\to K^+K^-}_{KK}$ respectively. The misidentification rates to the $K\pi$ and KK final states, and the

relative yields, are given in Appendix B.

Applying this method to determine the relative contamination and its respective uncertainty from the $\Lambda_b \to p\mathrm{K}^-$ contribution to the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, causes this to become the dominant systematic. This is due to the predicted central value of the relative contribution from Equation 6.9, combined with the large uncertainties on the measured branching ratio and proton misidentification rate (Tables B.1 and B.2 respectively). Thus, a second method is employed to determine the contamination rate directly from data. This method is performed by applying a set of identical kinematic cuts, Table 5.6, in selecting both $\Lambda_b \to p\mathrm{K}^-$ and $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ candidates simultaneously from a subsample of the 2011 data comprising 865 pb^{-1}. A set of specific PID cuts are applied on the $\Lambda_b \to p\mathrm{K}^-$ candidates to select $p\mathrm{K}$ final state particles, these are given in Table 6.4.

Table 6.4: PID criteria for K and p final state particles. These cuts are used to select $\Lambda_b \to p\mathrm{K}^-$ candidates to determine their relative contribution to the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ channel using 2011 data.

Particle	PID Cuts
K	$\mathrm{DLL}_{\mathrm{K}\pi} > +5$ & $\mathrm{DLL}_{\mathrm{K}p} > -4$ & $\mathrm{DLL}_{\mathrm{K}e} > 5$ & $\mathrm{DLL}_{\mathrm{K}\mu} > 5$
p	$\mathrm{DLL}_{p\mathrm{K}} > 10$ & $\mathrm{DLL}_{p\pi} > 10$ & $\mathrm{DLL}_{pe} > 0$ & $\mathrm{DLL}_{p\mu} > 0$

The yield $\Lambda_b \to p\mathrm{K}^-$ in the KK spectrum relative to the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ can be expressed as

$$\frac{N^{\mathrm{KK}}_{\Lambda_b \to p\mathrm{K}^-}}{N^{\mathrm{KK}}_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}} = \frac{N^{pk}_{\Lambda_b \to p\mathrm{K}^-}}{N^{\mathrm{KK}}_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}} \times \frac{\omega^{\Lambda_b \to p\mathrm{K}^-}_{\mathrm{KK}}}{\epsilon^{\Lambda_b \to p\mathrm{K}^-}_{pK}}, \tag{6.10}$$

where $N^{\mathrm{KK},pK}_{\Lambda_b \to p\mathrm{K}^-,\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$ is the number of events of the respective decay reconstructed with KK or $p\mathrm{K}$ PID selection, and ϵ and ω are the PID efficiencies and misidentification rates. The yields from the two fits on this dataset are

$$N^{pK}_{\Lambda_b \to p\mathrm{K}^-} = 1\,665 \pm 55 \tag{6.11}$$

and

$$N^{\mathrm{KK}}_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-} = 9\,047 \pm 111. \tag{6.12}$$

These yields, together with the PID efficiencies and misidentification rates given in Table B.2, are used to calculate the ratio given in Equation 6.10. The ratio of yields of the two decays in the KK spectrum calculated from these numbers is

$$\frac{N^{\mathrm{KK}}_{\Lambda_b \to p\mathrm{K}^-}}{N^{\mathrm{KK}}_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}} = 2.52 \pm 0.93 \times 10^{-2}. \tag{6.13}$$

This second method is not exclusive to the $\Lambda_b \to p\mathrm{K}^-$ background. Where it is possible, the method can be applied analogously to the other background channels, potentially reducing the uncertainties determined from the first method. The results of these additional fits are provided in Table B.3.

These two methods determine the relative cross contamination from the dominant misidentified backgrounds to each decay. However, the uncertainties on each of the relative contributions must then be converted into an uncertainty on the fitted lifetime. For the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ channel, the contributing backgrounds are the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ and $\Lambda_b \to p\mathrm{K}^-$, and for the $\mathrm{K}\pi$ channels the backgrounds are $\mathrm{B}_d^0 \to \pi^+\pi^-$, $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and $\Lambda_b \to p\pi^-$. To determine the respective uncertainties from these backgrounds, the relative fraction of each background is varied within $\pm 1\sigma$ of the uncertainties given in Table B.3, with the observed shift in fitted lifetime taken as the systematic uncertainty for each background. The uncertainties related to the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ misidentified backgrounds are found to be 1.03 fs for $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$, 3.40 fs for $\mathrm{B}_d^0 \to \mathrm{K}^+\mathrm{K}^-$, 0.16 fs for $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ and 3.16 fs for $\Lambda_b \to p\mathrm{K}^-$. The uncertainties for the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ from the $\mathrm{K}\pi$ misidentified backgrounds are 0.02 fs for $\mathrm{B}_d^0 \to \pi^+\pi^-$, 0.03 fs for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and 1.30 fs for $\Lambda_b \to p\pi^-$, with the uncertainties for the $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ being 0.13 fs for $\mathrm{B}_d^0 \to \pi^+\pi^-$, 0.72 fs for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and 5.89 fs for $\Lambda_b \to p\pi^-$.

The total uncertainties from these sources for the 2011 lifetime measurements are 4.76 fs for $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, 1.30 fs for $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and 5.94 fs for $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$.

6.3.6 Detector Resolution Effects

As discussed in Section 4.5.1, the decay time resolution of the VELO is ≈ 50 fs. The physical model of the decay time used in the fitter is defined as a single Gaussian distribution with a width of 50 fs. The sensitivity to this assumption is investigated by re-running the fitter with 40 fs and 60 fs decay time resolution. The largest variation observed from changing the value of the resolution is taken as the uncertainty. This is found to be negligible for both the 2010 and 2011 lifetime analysis.

6.3.7 Minimum Accepted Lifetime

This cut is required by the background estimator, described in section 4.5.2, and studies the sensitivity to the minimum accepted lifetime cut. The study was only performed for the 2010 analysis due to the limited statistics available in the dataset. Toy MC datasets are generated with a minimum lifetime cut set at the start of the acceptance, which is at the lower end of the turning point distribution. The datasets have the cut varied between 0.6 ps and 0.8 ps, with

the observed lifetime shift in the fitted lifetime found to be $1.1\,\mathrm{fs}$. The uncertainty assigned to this source is thus $1.1\,\mathrm{fs}$.

6.3.8 Production Asymmetry

Initial assumptions in Equation 1.43 suggest that B_s^0 and $\overline{\mathrm{B}}_s^0$ mesons are produced in equal quantities at LHCb. However, since the LHC collides protons the production of particles versus antiparticles is asymmetric. The production asymmetry can be defined as

$$\mathcal{A}_p = \frac{R_{\mathrm{B}_s^0} - R_{\overline{\mathrm{B}}_s^0}}{R_{\mathrm{B}_s^0} + R_{\overline{\mathrm{B}}_s^0}}, \tag{6.14}$$

where $R_{\mathrm{B}_s^0}$ and $R_{\overline{\mathrm{B}}_s^0}$ are the rates of B_s^0 and $\overline{\mathrm{B}}_s^0$ production at LHCb. The presence of such an asymmetry can alter the measured proper decay time distribution of $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and have an effect on the measurement of the effective lifetime $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$. Ignoring the effects of decay time resolution and acceptance the effective lifetime, $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$, is simply

$$\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-} = \frac{\int_0^\infty t \cdot \Gamma[KK, t]\mathrm{d}t}{\int_0^\infty \Gamma[KK, t]\mathrm{d}t}. \tag{6.15}$$

The effect of $\mathcal{A}_p$ is determined from the following equations, derived in [89],

$$\begin{aligned} \int_0^\infty t \cdot \Gamma[f, t]\mathrm{d}t = \; & N' \; \left\{ \frac{1 - \mathcal{A}_{\Delta\Gamma}}{\Gamma_L^2} + \frac{1 + \mathcal{A}_{\Delta\Gamma}}{\Gamma_H^2} \right. \\ & + \; \left. 2\mathcal{A}_p \left[\mathcal{A}_{CP}^{dir} \frac{\Gamma_s^2 - \Delta m^2}{(\Delta m^2 + \Gamma_s^2)^2} + \mathcal{A}_{CP}^{mix} \frac{2\Gamma_s \Delta m}{(\Delta m^2 + \Gamma_s^2)^2} \right] \right\} \end{aligned} \tag{6.16}$$

and

$$\begin{aligned} \int_0^\infty \Gamma[f, t]\mathrm{d}t = \; & N' \; \left\{ \frac{1 - \mathcal{A}_{\Delta\Gamma}}{\Gamma_L} + \frac{1 + \mathcal{A}_{\Delta\Gamma}}{\Gamma_H} \right. \\ & + \; \left. 2\mathcal{A}_p \left[\mathcal{A}_{CP}^{dir} \frac{\Gamma_s}{\Delta m^2 + \Gamma_s^2} + \mathcal{A}_{CP}^{mix} \frac{\Delta m}{\Delta m^2 + \Gamma_s^2} \right] \right\}. \end{aligned} \tag{6.17}$$

Equations 6.16 and 6.17 allow us to calculate the bias on the effective lifetime originating from the production asymmetry, but require numerical values for the parameters involved. The oscillation parameters are taken from HFAG averages [36]: Γ_s = 0.663 ps^{-1}, $\Delta\Gamma_s$ = 0.095 ps^{-1} and Δm = 17.719 ps^{-1}. The asymmetry parameters are taken from the LHCb measurement of the time dependent CP asymmetries in $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ [110]. The values are chosen as the central values plus the total uncertainty for the two parameters: $\mathcal{A}_{dir} = 0.254$ and $\mathcal{A}_{mix} = 0.427$. Since $(\mathcal{A}_{\Delta\Gamma})^2 + (\mathcal{A}_{CP}^{dir})^2 + (\mathcal{A}_{CP}^{mix})^2 = 1$ this choice requires that $\mathcal{A}_{\Delta\Gamma} = 0.868$.

The bias in the effective lifetime as a function of the production asymmetry, $\mathcal{A}_p$, is shown in Figure 6.11. It is defined as $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}(\mathcal{A}_p) - \tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}(\mathcal{A}_p = 0)$, where $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$ is the ratio of Equations 6.16 and 6.17. The value determined is assigned as the systematic uncertainty.

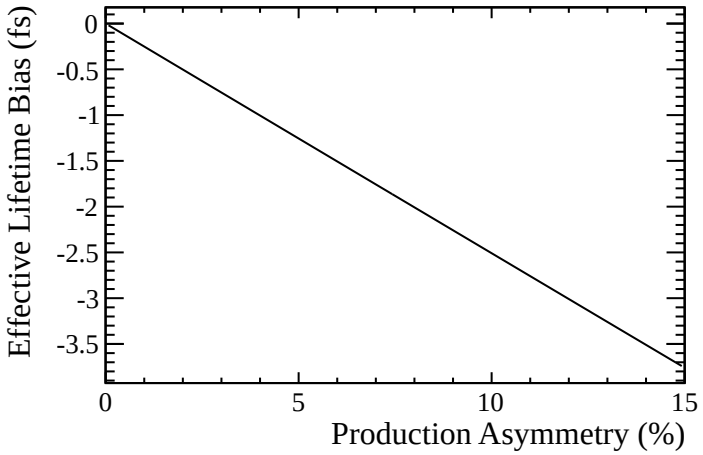


Figure 6.11: The bias in effective lifetime as a function of production asymmetry $\mathcal{A}_p$ using the parmeters in the text. The function is symmetric around $\mathcal{A}_p = 0$ [89].

For the 2010 analysis, the value of $\mathcal{A}_p$ for the B_d^0 system is taken from the LHCb measurement in [109], which was measured as $0.024 \pm 0.013\,(\mathrm{stat})\,0.010\,(\mathrm{syst})$. For the B_s^0 system it is conservatively assumed that the production asymmetry is the same. Thus to determine an upper limit on the size of such an effect, the value of $\mathcal{A}_p = 0.040$ (1σ larger than the LHCb measured value) is used. From this, the uncertainty assigned for the 2010 analysis is $1.4\,\mathrm{fs}$. The 2011 analysis uses a measurement of the B_s^0 production asymmetry, which is found to be $7 \pm 5\%$ [111]. An upper limit on the uncertainty of the production asymmetry is determined to be $3.00\,\mathrm{fs}$.

6.3.9 Summary of Systematic Uncertainties

The list of systematic uncertainties for the 2010 and 2011 lifetime measurements are provided in Tables 6.5 and 6.6 respectively. The total systematic error for $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$ using the 2010 dataset is $8.3\,\mathrm{fs}$. For the lifetime measurements performed using the 2011 dataset, the total systematic errors for $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$ is $6.70\,\mathrm{fs}$, $\tau_{\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-}$ is $4.34\,\mathrm{fs}$ and $\tau_{\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-}$ is $12.13\,\mathrm{fs}$.

Source	Uncertainty on $B_s^0 \to K^+K^-$ (fs)
Method and verification	3.2
Mass Model	1.9
Parametric Decay Time	-
Non-Parametric Decay Time	1.5
VELO related uncertainties	
Finite Length of VELO	-
Radial Dependence	-
Tracking Efficiency	-
Acceptance Correction	6.3
PV assignment	1.2
Absolute Decay Time Bias	1.5
Cross Contamination	2.7
Decay Time Resolution	-
Minimum Lifetime	1.1
Production Asymmetry	1.4
Total	8.3
Effective Lifetime Bias (*)	2.8

Table 6.5: Contributions to the systematic uncertainty on the $B_s^0 \to K^+K^-$ 2010 lifetime measurement. (*) This uncertainty is described in Section 6.3.3.1 and classified as a parametric uncertainty.

Source	Uncertainty (fs)		
	$B_s^0 \to K^+K^-$	$B_d^0 \to K^+\pi^-$	$B_s^0 \to \pi^+K^-$
Method and verification	0.17	0.11	0.59
Mass Model	1.10	2.51	6.74
Parametric Decay-Time	1.78	0.40	3.60
Non-Parametric Decay Time	0.82	1.57	6.70
VELO related uncertainties			
Finite VELO Length	-	-	-
Radial Dependence	-	-	-
Tracking Efficiency	2.83	2.83	2.83
Acceptance Correction	-	-	-
PV assignment	-	-	-
Absolute Decay Time Bias	0.41	0.43	0.43
Cross Contamination	4.76	1.30	5.94
Decay Time Resolution	-	-	-
Minimum Lifetime	-	-	-
Production Asymmetry	3.00	-	-
Total	6.70	4.34	12.13

Table 6.6: Contributions to the systematic uncertainty on the $B_s^0 \to K^+K^-$, $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ 2011 lifetime measurements.

6.4 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ Lifetime Measurement Results

The lifetime measurement results presented in this section are split into two. The first set of results detail the measurement made with the 2010 dataset. The second details the measurements with the much larger, and statistically independent, 2011 dataset. Each analysis is then split into a measurement of particular $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ signal modes, which are described in each section.

6.4.1 2010 Lifetime Measurement Results

The fit results using the full 2010 LHCb sample of $\approx 37\,\mathrm{pb}^{-1}$ incorporate both the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and control channel of $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$. The $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ is chosen as the control channel as its lifetime has been well measured and understood (0.6% uncertainty in PDG [22]). In addition, the combination of hadronisation and branching fraction parameters mean that the channel is also more abundant than that of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$. An added simplification of the fit is that the selection criteria is identical to that of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ with the exception of the PID cuts.

6.4.1.1 $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ Results

As discussed in Section 4.6, the resultant parameters from the mass fit are used in the lifetime fit. Candidate events with a lifetime less than the minimum value for the acceptance function, $0.6\,\mathrm{ps}$, are rejected. Events with lifetimes greater than 10 times the actual lifetime of the B_d^0, $15\,\mathrm{ps}$, are similarly removed. The fit to the total lifetime distribution is shown in Figure 6.12. The figure shows the total distribution, the background estimation, the signal fit and the total fitted PDF.
The result of the lifetime fit of the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ is

$$\tau_{\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-} = 1.535 \pm 0.074 \ \mathrm{ps\ (stat)} \tag{6.18}$$

The result can be compared to the average B_d^0 lifetime value [22]:

$$\tau_{\mathrm{B}} = 1.519 \pm 0.007 \ \mathrm{ps}. \tag{6.19}$$

For the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ mass fit and subsequent lifetime measurement, the $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ (bottom plot Figure 6.1) is omitted so the models used are identical to those employed in the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime measurement. To cross check the control measurement an additional model including the secondary $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ peak is added and fitted with the results being:

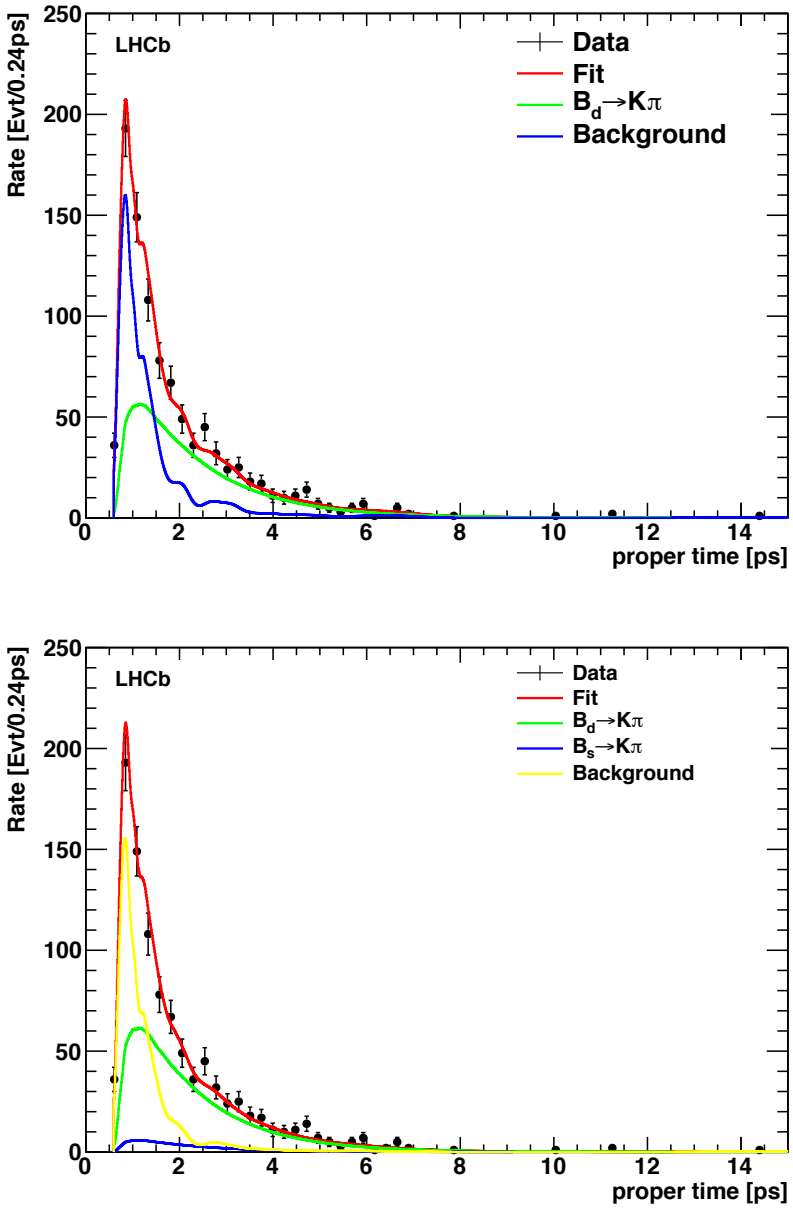


Figure 6.12: The histograms show the total time distribution of the events for the $B^0_d \to K^+\pi^-$ (and $B^0_s \to \pi^+K^-$) datasets. The graphs show the background estimation, the fitted signal distributions and the total fitted PDF. Top: the fit with only $B^0_d \to K^+\pi^-$ and combinatorial background. Bottom: fit including also the $B^0_s \to \pi^+K^-$.

$$\tau_{B^0_d \to K^+\pi^-} = 1.544 \pm 0.068 \text{ ps (stat)}. \tag{6.20}$$

The current $B^0_s \to \pi^+K^-$, or flavour specific, lifetime [36] is given by:

$$\tau_{B^0_s} = 1.463 \pm 0.032 \text{ ps}. \tag{6.21}$$

6.4.1.2 $B^0_s \to K^+K^-$ Results

As with the measurement in section 6.4.1.1, $B^0_s \to K^+K^-$ candidates that have a lifetime lower than the minimum of the acceptance function (0.6 ps) are rejected along with ones that have a lifetime greater than 1.5 ps. The fit to the total lifetime distribution is shown in Figure 6.13. The figure shows the total distribution, the background estimation, the signal fit and the total fitted PDF.

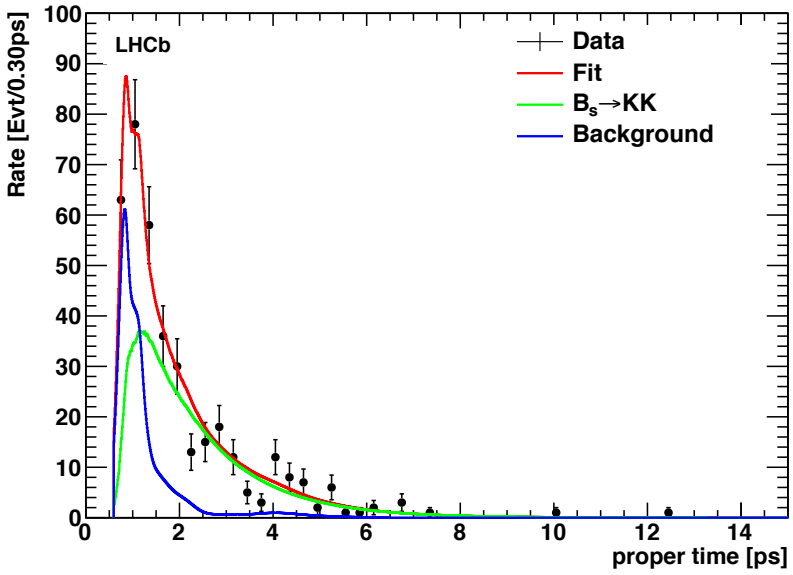


Figure 6.13: The histogram shows the total time distribution of the events for the $B^0_s \to K^+K^-$ dataset. The plot also shows the background estimation, the fitted signal distribution and the total fitted PDF.

The result of the lifetime fit of the $B^0_s \to K^+K^-$ is

$$\tau_{B^0_s \to K^+K^-} = 1.440 \pm 0.096 \text{ ps (stat)} \pm 0.008 \text{ ps (syst)}. \tag{6.22}$$

6.4.2 2011 Lifetime Measurement Results

The fit results for the 2011 LHCb dataset with total integrated luminosity of $\approx 1.0\,\mathrm{fb}^{-1}$, incorporate additional signal modes compared with the measurement conducted with the smaller 2010 dataset. The lifetimes being measured with this data include: $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$, $\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-}$ and $\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-}$. Full verification of the updated fitter was performed with simplified MC, Section 5.6.1.4, and full LHCb MC, Section 5.6.2.2. As with the 2010 analysis, there are a set of pre-fit lifetime cuts applied to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ selected datasets. The lower lifetime cut, $0.61\,\mathrm{ps}$ is applied to remove events with a lifetime lower than the minimum of the acceptance function. The higher lifetime cut, of $10.0\,\mathrm{ps}$, is applied to stabilise the acceptance function at high lifetimes where there is low statistics.

6.4.2.1 $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ Results

The $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ lifetime measurements are performed simultaneously on the 2011 selected data, with both incorporating fully evaluated systematics. Analysis of both lifetimes is performed using the blinding method described in Section 4.7 to avoid analyst bias. The unblinded fit to the total lifetime distribution, with separation into the individual signal components by the sWeights described in Section 4.3.3.2, is shown in Figure 6.14.
The result of the lifetime fit of the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ is

$$\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-} = 1.524 \pm 0.011\ \mathrm{ps\,(stat)}\ \pm\ 0.004\ \mathrm{ps\,(syst)}, \tag{6.23}$$

with the result of the fit of the $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ being

$$\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-} = 1.597 \pm 0.056\ \mathrm{ps\,(stat)}\ \pm\ 0.012\ \mathrm{ps\,(syst)}. \tag{6.24}$$

The lifetime of the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ is within 1σ of the world average, Equation 6.19, with the lifetime of the $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ within 2.5σ of the flavour specific B^0_s lifetime, Equation 6.21.

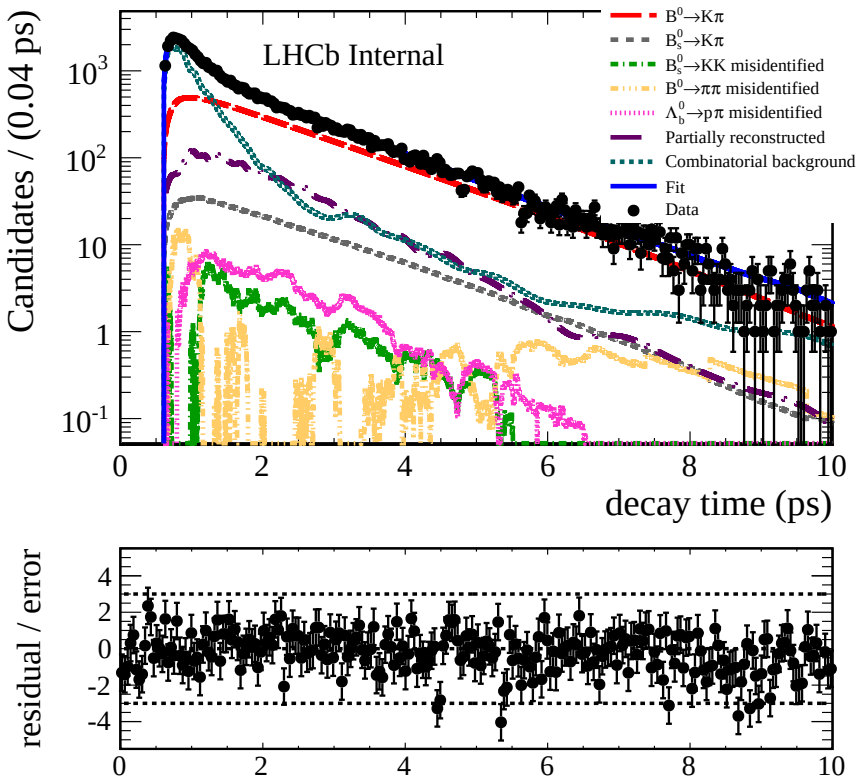


Figure 6.14: The histograms show the total time distribution of the events from the 2011 data selected as $B_d^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$. The graphs show the background estimation, the fitted signal distributions and the total fitted PDF.

6.4.2.2 $B_s^0 \to K^+K^-$ Results

The primary measurement of interest is the $B_s^0 \to K^+K^-$ lifetime that provides access to the interesting SM quantities discussed in detail in Section 1.4. As with the $B_d^0 \to K^+\pi^-$ and the $B_s^0 \to \pi^+K^-$ lifetimes, the fit is performed blinded until the fit process is fully verified as in Section 5.6.2.2. The unblinded fit to the total lifetime distribution, with separation into the individual signal components by the sWeights described in Section 4.3.3.2, is shown in Figure 6.15.

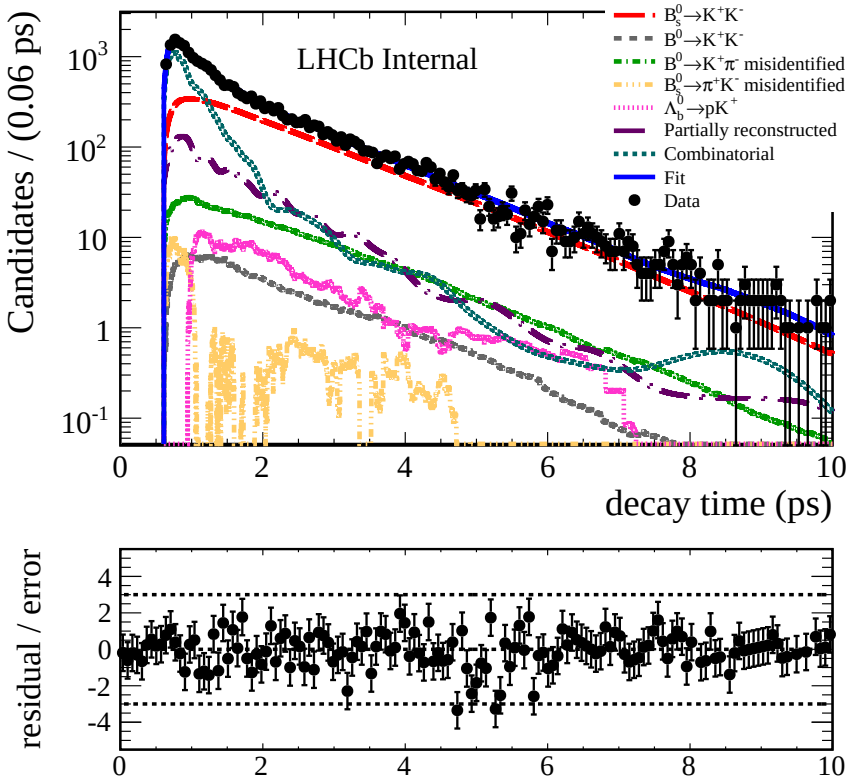


Figure 6.15: The histogram shows the total time distribution of the events from the 2011 dataset selected as $B_s^0 \to K^+K^-$. The plot also shows the background estimation, the fitted signal distribution and the total fitted PDF.

The result of the lifetime fit to the $B_s^0 \to K^+K^-$ is

$$\tau_{B_s^0 \to K^+K^-} = 1.407 \pm 0.016 \text{ ps (stat)} \pm 0.007 \text{ ps (syst)}. \quad (6.25)$$

The uncertainty on the $B_s^0 \to K^+K^-$ lifetime fit is approximately five times smaller than the 2010 measurement, and approximately 3 times smaller than the previous LHCb measurement using the same dataset [99]. One of the key advantages of the method in this thesis, compared to previous measurements, is the increased statistics that can be utilised, since more events

are rejected in the unbiased selection [99].

6.5 Implications for $\tau_{B_s^0 \to K^+K^-}$ and the Asymmetry Parameter $\mathcal{A}_{\Delta\Gamma}$

The three measurements of $\tau_{B_s^0 \to K^+K^-}$ made at LHCb, along with the first measurement made by CDF, are displayed graphically in Figure 6.16. The precision of the 2011 $\tau_{B_s^0 \to K^+K^-}$ measurement is a significant improvement on the current world average value [36].

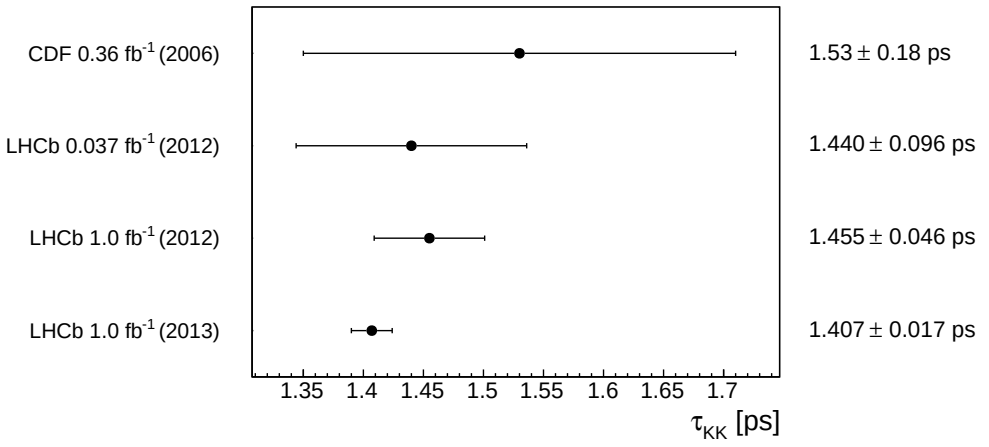


Figure 6.16: Graphical comparison of the four available measurements of the specific $B_s^0 \to K^+K^-$ lifetime measurement from [32], [76], [99] and [89].

As well as making two world's best measurements of the value of $\tau_{B_s^0 \to K^+K^-}$ in this thesis, the 2011 measurement given in Equation 6.25 now has the smallest uncertainties of all current measurements. This measurement of the lifetime from the *CP*-even $B_s^0 \to K^+K^-$ decay, combined with the lifetime from the *CP*-odd $B_s^0 \to J/\Psi\, f_0(980)$ decay [112],

$$\tau_{B_s^0 \to J/\Psi f_0(980)} = 1.700 \pm 0.040 \text{ ps (stat)} \pm 0.026 \text{ ps (syst)}, \tag{6.26}$$

provides independent verification that the decay rate difference $\Delta\Gamma_s$ is positive and measurable, with a greater than 5σ precision. The value of the decay rate asymmetry, $\mathcal{A}_{\Delta\Gamma}$, can be determined from the 2011 measurement of $\tau_{B_s^0 \to K^+K^-}$ by combining Equations 1.47 and 1.48 to first order to give

$$\mathcal{A}_{\Delta\Gamma} = \frac{2\Gamma_s^2}{\Delta\Gamma_s}\tau_{B_s^0 \to K^+K^-} - \frac{2\Gamma_s}{\Delta\Gamma_s}. \tag{6.27}$$

This requires further input from recent LHCb measurements of Γ_s and $\Delta\Gamma_s$ [113]. These values are given as

$$\begin{aligned} \Gamma_s &= 0.661 \pm 0.004 \text{ ps}^{-1} \text{ (stat)} \pm 0.006 \text{ ps}^{-1} \text{ (syst)}, \\ \Delta\Gamma_s &= 0.106 \pm 0.011 \text{ ps}^{-1} \text{ (stat)} \pm 0.007 \text{ ps}^{-1} \text{ (syst)}, \end{aligned}$$

The value of $\mathcal{A}_{\Delta\Gamma}$ evaluated using the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime measurement, Equation 6.25, is found to be

$$\begin{aligned} \mathcal{A}_{\Delta\Gamma} = & -0.87 \pm 0.13 \, (\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}(\text{stat})) \pm 0.06 \, (\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}(\text{syst})) \\ & \pm 0.06 \, (\Gamma_s(\text{stat})) \pm 0.10 \, (\Gamma_s(\text{syst})) \\ & \pm 0.09 \, (\Delta\Gamma_s(\text{stat})) \pm 0.06 \, (\Delta\Gamma_s(\text{syst})), \end{aligned} \tag{6.28}$$

where the dominant sources of uncertainty can be seen by splitting the final value into the constituent components from each input measurement. The largest source of uncertainty originates from the statistical precision on $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$, although other significant contributors are the systematic uncertainty of Γ_s and the statistical uncertainty of $\Delta\Gamma_s$. The value of $\mathcal{A}_{\Delta\Gamma}$, with the total statistical and systematic uncertainties added in quadrature, is

$$\mathcal{A}_{\Delta\Gamma} = -0.87 \pm 0.17 \text{ (stat)} \pm 0.13 \text{ (syst)}. \tag{6.29}$$

This is currently consistent with the standard model prediction, given by $\mathcal{A}_{\Delta\Gamma}(\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-) = -0.972^{+0.014}_{-0.009}$ [35], discussed in Section 1.4. This result can be interpreted using the results detailed in [110], to determine further evidence for *CP*-violation in the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ system. The measurements[1] are

$$\mathcal{A}_{dir}^{\mathrm{KK}} = -0.14 \pm 0.11 \text{ (stat)} \pm 0.03 \text{ (syst)}, \tag{6.30}$$

$$\mathcal{A}_{mix}^{\mathrm{KK}} = 0.30 \pm 0.12 \text{ (stat)} \pm 0.04 \text{ (syst)}, \tag{6.31}$$

$$\rho(\mathcal{A}_{dir}^{\mathrm{KK}}, \mathcal{A}_{mix}^{\mathrm{KK}}) = 0.02, \tag{6.32}$$

where $\mathcal{A}_{dir}^{\mathrm{KK}}$ and $\mathcal{A}_{mix}^{\mathrm{KK}}$ are the direct and mixing amplitudes described in Section 1.3.5 and 1.4, and $\rho(\mathcal{A}_{dir}^{\mathrm{KK}}, \mathcal{A}_{mix}^{\mathrm{KK}})$ is the correlation coefficient. Combining the results from Equations 6.29, 6.30 and 6.31, and neglecting the small correlation, we observe

[1] In the paper, the parameters provided are referred to as $C_{\mathrm{KK}} = -\mathcal{A}_{dir}^{\mathrm{KK}}$ and $S_{\mathrm{KK}} = \mathcal{A}_{mix}^{\mathrm{KK}}$.

$$|\mathcal{A}_{dir}^{\mathrm{KK}}|^2 + |\mathcal{A}_{mix}^{\mathrm{KK}}|^2 + |\mathcal{A}_{\Delta\Gamma}|^2 = 0.87 \pm 0.31\ (\mathrm{stat})\ \pm\ 0.23\ (\mathrm{syst}), \tag{6.33}$$

which is consistent with 1. The value of $\mathcal{A}_{\Delta\Gamma}$ measured in this analysis is consistent with evidence of *CP*-violation in mixing in the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ system, which is shown by an independent analysis [110].

6.6 Summary

This chapter has continued the subject of lifetime measurements from the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decay family that was first begun with the lifetime fit method, described in Chapter 4, and then continued with the preparation of the data and fit before verification of the full fitter, Chapter 5. The results of the fits to the invariant mass spectrum have first been provided, these detail the individual signal yields for both the 2010 and 2011 analyses and also provide details of a secondary analysis, the search for $\mathrm{B}^0_{\mathrm{d/s}} \to \mathrm{p\bar{p}}$, where the fit to the 2011 $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ dataset is both complimentary and necessary for the measurement. A discussion is then given on the systematic studies that are performed to determine the potential biases. The sources of bias can be observed to originate from areas such as the models used to describe the signal mass distributions, the method for ascertaining the non-parametric lifetime distributions and detector performance. As there are two analyses performed using datasets from different data taking periods, there are two sets of final results. The value of $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$ using the full $37\,\mathrm{pb}^{-1}$ from 2010 is found to be

$$\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-} = 1.440 \pm 0.096\ \mathrm{ps}\ (\mathrm{stat})\ \pm\ 0.008\ \mathrm{ps}\ (\mathrm{syst}), \tag{6.34}$$

The $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime values evaluated from the full $1.0\,\mathrm{fb}^{-1}$ and with full statistical and systematic uncertainties are

$$\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-} = 1.407 \pm 0.016\ \mathrm{ps}\ (\mathrm{stat})\ \pm\ 0.007\ \mathrm{ps}\ (\mathrm{syst}), \tag{6.35}$$

$$\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-} = 1.524 \pm 0.011\ \mathrm{ps}\ (\mathrm{stat})\ \pm\ 0.004\ \mathrm{ps}\ (\mathrm{syst}), \tag{6.36}$$

$$\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-} = 1.597 \pm 0.056\ \mathrm{ps}\ (\mathrm{stat})\ \pm\ 0.012\ \mathrm{ps}\ (\mathrm{syst}). \tag{6.37}$$

As of the time of measurement, the 2010 value of $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$ that was published in [76] had the lowest uncertainties of any previous measurement. Since this value was made available, a measurement using an unbiased selection on $1.0\,\mathrm{fb}^{-1}$ of the 2011 dataset has been

published [99] which reduced the uncertainties down to $0.046\,\mathrm{ps}$. The second set of measurements from the full 2011 dataset presented here, supercede this, by reducing the uncertainties further to $0.017\,\mathrm{ps}$. In particular, the value of $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$ is the world's most accurate measurement of this lifetime and will dominate the world average determination. Specific interest is given to the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime, which allows the SM parameter $\mathcal{A}_{\Delta\Gamma}$ to be constrained. The value of $\mathcal{A}_{\Delta\Gamma}$ determined from the 2011 value of $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$ is

$$\mathcal{A}_{\Delta\Gamma} = -0.87 \pm 0.17\,(\mathrm{stat}) \;\pm\; 0.13\,(\mathrm{syst}), \tag{6.38}$$

which is the world's first determination of this parameter from data, and is currently consistent with the SM prediction.

Chapter 7

Conclusions and Outlook

7.1 Summary

This thesis has presented lifetime measurements of the family of decays known as $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$, with specific interest shown to the decay mode $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$. Lifetime measurements in heavy flavour physics are interesting as they can provide access to SM model parameters, in the case of $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ this occurs through the decay rate difference parameter $\Delta\Gamma_\mathrm{s}$ and the decay rate asymmetry parameter $\mathcal{A}_{\Delta\Gamma}$. The lifetime value $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$, obtained through the analysis described through Chapters 4 to 6, originates from a decay that is largely CP even but also dominated by loop processes. This means the channel is sensitive to NP particles that can alter the channel from SM expected values. The possible contribution of NP effects to the channel can be studied by using the extracted value of $\tau_{\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-}$ to determine values of $\Delta\Gamma_\mathrm{s}$ and $\mathcal{A}_{\Delta\Gamma}$, where $\mathcal{A}_{\Delta\Gamma}$ is calculated in this thesis.
Monitoring studies of the RICH, in particular RICH 2, were also discussed in detail. The first study used a system developed to monitor the stability of the RICH mirrors. Observations from data displayed actual movement within the mirror structure, where further investigation showed a strong correlation with temperature changes within RICH 2. However, it is shown that these mirror movements are negligible, compared to the RICH 2 angular resolution. A second study was proposed to use environmental data from the monitoring system, to determine a method for performing continuous calibration of the absolute gas mass, refractive index (n) and Cherenkov angle (θ_c) as a function of time. The calibration was in excellent agreement with previous measurements that are used currently and performed by the RICH group, although the measurements are taken at larger time intervals.

An overview of particle physics and the theory behind the SM is provided in Chapter 1. Particular focus is given to flavour physics including: quark mixing, the CKM matrix, CP-

violation and the phenomenon of neutral meson mixing. The impact of these effects are then discussed in terms of the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ family of decays, with specific focus given to the physics case behind the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime measurement.

The second chapter presents a detailed description of the LHCb experiment that is part of the LHC. An overview of the LHC complex is presented that includes details of the main ongoing experiments, this is then followed by a detailed description of the LHCb experiment. The sub-detectors are introduced in the order of their relative position to the particle interaction point, with their performance then discussed. In addition, the reconstruction and data processing software used in the lifetime analyses, which is an integral part of the experiment, are introduced.

The Ring Imaging Cherenkov (RICH) detectors of the LHCb experiment provide excellent Particle IDentification (PID) performance for charged hadrons, as documented in the description of the detectors in Chapter 2. Their consistent performance is aided by constant monitoring of the system. Two forms of this monitoring are studied in depth in Chapter 3. A study to determine the stability of the spherical and flat mirrors using the Laser Alignment Monitoring System (LAMS) has shown there is some physical movement within the system. Before analysing the data fully, each mirror had to be calibrated independently to determine an angular movement from that observed by the monitoring cameras. The movements observed over the 2011 and 2012 monitoring period are in the order of $10 - 70\ \mu\mathrm{rad/K}$, which agrees with previous estimates of the magnitude of the potential movement. Further investigation into the origins of this movement focussed on the effects of environmental changes in the pressure, temperature, and the shift in magnet polarity. Of these three, a direct correlation with temperature was found that could approximately account for the size of movement observed. A second study was performed that proposed an alternative method to the currently employed RICH procedure that extracts the value of the refractive index of the RICH 2 gas. The alternative method is purely data driven, with the only prior assumptions being the ideal gas law and the relationship between gas density and refractive index. In addition, the method has the potential to make measurements for any point in time where there is relevant environmental data, this cannot be done with the current methodology. The method can also make measurements of the expected Cherenkov angle (θ_c) and absolute molecular gas mass ($< M >$). After calibration with data samples, the results showed excellent agreement between the new method and the current method, with the differences between the value of n being of the order of 1% and the difference between θ_c being of the order of $< 1\%$. The absolute gas mass was also in very good agreement with the measurements, with fluctuations in the value between current measurements easier to track.

The performance of the RICH detectors is crucial for the flavour physics analyses performed

at LHCb. This is particularly important for the study of $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays, as different final states can only be distinguished if the final state particles are correctly identified. The $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ decay is part of this family, with a lifetime measurement of this being of particular interest.

The details of the lifetime measurements made for the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and other members of the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ family commence in Chapter 4, and are then discussed further in Chapters 5 and 6.

The methodology behind the fit to extract the effective lifetime of a specific decay is presented in Chapter 4. Biases in a lifetime measurement, through event selection criteria from cuts on lifetime correlated variables, are discussed initially. This can occur through multiple stages of the LHCb event selection process. The fit method described is able to remove the lifetime bias incurred by using a purely data driven method. This is done by determining a per-event lifetime acceptance function that is accounted for in the fitting. The advantage of this approach means fewer assumptions are made in determining the acceptance function per decay.

The chapters that follow describe the analyses themselves, with both Chapter 5 and Chapter 6 providing the details for these two analyses. The first uses the full 2010 LHCb dataset, and makes a measurement of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime only. The second uses the full 2011 LHCb dataset and incorporates two more channels as well as the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, these are the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ decays.

Preparation for the analysis of the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes is discussed in Chapter 5. An Initial overview of the current experimental status of the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime is provided, this includes two of the previous measurements made at LHCb (including the 2010 analysis discussed in this thesis). The optimisation of the data selection for each analysis is presented, with the mass and lifetime models ascribed to each signal discussed. The fitter is then subjected to a range of verification studies to ensure its performance, these include using simplified simulation, full LHCb MC and high statistics $\mathrm{D} \to \mathrm{h}^+\mathrm{h}^-$ data. The quality of these fits are then assessed and found to be in excellent agreement with the expected values.

The final chapter, Chapter 6, presents fit results from the mass and lifetime fits, and the evaluation of systematic uncertainties that can affect the precision of the lifetime measurement. The signal yields for the relevant $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ channels fitted in each analysis, are given.

The fit to the invariant mass spectrum stage of the full lifetime fit is also used in a separate analysis, to search for the rare baryonic B decays $B_{d/s}^0 \to p\bar{p}$. The signal yield of a $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ dataset is used for the critical role of the normalisation and yielded the first evidence for the $\mathrm{B}_d^0 \to \mathrm{p}\overline{\mathrm{p}}$, with a statistical significance of 3.3σ [107]. Estimates of the branching fractions from data for both the $\mathrm{B}_d^0 \to \mathrm{p}\overline{\mathrm{p}}$ and $\mathrm{B}_s^0 \to \mathrm{p}\overline{\mathrm{p}}$ are

$$\mathcal{B}(\mathrm{B}^0_d \to \mathrm{p\overline{p}}) = (1.47\,^{+0.62\ +0.35}_{-0.51\ -0.14}) \times 10^{-8} \tag{7.1}$$

$$\mathcal{B}(\mathrm{B}^0_s \to \mathrm{p\overline{p}}) = (2.84\,^{+2.03\ +0.85}_{-1.68\ -0.18}) \times 10^{-8} \tag{7.2}$$

Before the full lifetime fit is conducted, a range of systematic studies are performed. These incorporate a large number of potential sources of uncertainty, with the dominant being from the acceptance correction and reconstruction effects for the 2010 $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime analysis, and cross contamination for the 2011 $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetime analysis. These result in a value for the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime using the 2010 dataset of

$$\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-} = 1.440 \pm 0.096\ \mathrm{ps\ (stat)}\ \pm\ 0.008\ \mathrm{ps\ (syst)}, \tag{7.3}$$

with the $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ lifetimes measured from the 2011 dataset being

$$\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-} = 1.407 \pm 0.016\ \mathrm{ps\ (stat)}\ \pm\ 0.007\ \mathrm{ps\ (syst)}, \tag{7.4}$$

$$\tau_{\mathrm{B}^0_d \to \mathrm{K}^+\pi^-} = 1.524 \pm 0.011\ \mathrm{ps\ (stat)}\ \pm\ 0.004\ \mathrm{ps\ (syst)}, \tag{7.5}$$

$$\tau_{\mathrm{B}^0_s \to \pi^+\mathrm{K}^-} = 1.597 \pm 0.056\ \mathrm{ps\ (stat)}\ \pm\ 0.012\ \mathrm{ps\ (syst)}. \tag{7.6}$$

These measurements are consistent with previous measurements, with the value of $\tau_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$ measured from the 2011 dataset having a greater precision than the current world average. From this, we are able to calculate the value of the quantity $\mathcal{A}_{\Delta\Gamma}$ and perform a full measurement of this quantity for the first time. The evaluated value is found to be

$$\mathcal{A}_{\Delta\Gamma} = -0.87 \pm 0.17\ \mathrm{(stat)}\ \pm\ 0.13\ \mathrm{(syst)}, \tag{7.7}$$

which is currently consistent with the SM predicted value.

7.2 Outlook

The integrated luminosity accumulated at LHCb has already increased dramatically from $\approx 37\,\mathrm{pb}^{-1}$ to $\approx 1.0\,\mathrm{fb}^{-1}$ between the data taking periods of 2010 and 2011. This is a factor ≈ 30 increase in data size, which has already seen a decrease in statistical uncertainty from $0.096\,\mathrm{ps}$ to $0.017\,\mathrm{ps}$ on the measured value of the effective $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime. The data taking run of 2012 has since increased the centre of mass beam energy from $\sqrt{s} = 7\ \mathrm{TeV}$ to $\sqrt{s} = 8\ \mathrm{TeV}$, and has accumulated an additional $2.0\,\mathrm{fb}^{-1}$ of data.

In current measurements of the effective $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ lifetime the statistical uncertainty is still the largest error, with the largest systematic uncertainties from the latest LHCb measurements originating from cross contamination from other $\mathrm{H_b} \to \mathrm{h}^+\mathrm{h}^-$ backgrounds. A future measurement using the 2012 dataset would reduce the statistical uncertainty to approximately the level of the current systematic uncertainty. However with the increase in centre of mass beam energy, increase in accumulated data and updates to the methods employed, it is expected that we would reduce the systematic uncertainties even further.

The main updates to the method would involve reoptimisation of the event selection, this would be due to improvements in detector calibration and reconstruction performance. The signal models would also require updating, with additional models for the mass being included as the precision on the measurement increases. The misidentified signal lifetime distributions are modelled with a fixed parametric model, these can be altered to a non-parametric model with the statistics increase associated with the larger integrated luminosity. As cross contamination has a significant input to the uncertainty of the 2011 measurements, the PID cuts would be required to be tightened to reduce the contributions of these backgrounds.

As more data is collected, the uncertainty on the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ will decrease such that the measurement will be systematics dominated. This improvement will allow us to probe the SM with greater accuracy, in particular the value of the $\mathcal{A}_{\Delta\Gamma}$ parameter calculated in this thesis. Future measurements of this would also require better knowledge of the values of $\Delta\Gamma_\mathrm{s}$ and Γ_s, which will come from updated LHCb measurements with full $3\,\mathrm{fb}^{-1}$ of the 2012 dataset. Deviations from the SM predicted values of these quantities could point in the direction of new sources of CP-violation, this would open up new areas of study in high energy physics and improve our understanding of the universe.

Appendix A

LAMS Positions and Vectors

This appendix details the values incorporated into the ray tracing code employed to calibrate the LAMS mirror system. The tables in this appendix provide the values of: the central positions of the beam splitters in Table A.1, the vectors from the central beam splitter position to the respective mirror in Table A.2, the central positions of the CMOS cameras in Table A.3, the vector from each CMOS camera in Table A.4 centre to each mirror (pointing direction) and the rotation of the camera planes used in Table A.5.

Table A.1: RICH beam splitter central position coordinates.

RICH Side	Camera	x	y	z
A-Side	00	726.71	-3033	10528.26
	01	826.16	-3033	10245.22
	02	960.16	-3033	10457.88
	03	1333.70	-3033	10567.67
	04	1647.36	-3033	10993.03
	05	1601.37	-3033	10747.29
	06	2071.71	-3033	11015.35
	07	2025.72	-3033	10769.62
C-Side	08	-2025.72	-3033	10769.62
	09	-2071.71	-3033	11015.35
	10	-1601.38	-3033	10747.29
	11	-1647.36	-3033	10993.03
	12	-1333.70	-3033	10567.67
	13	-960.16	-3033	10457.88
	14	-826.16	-3033	11245.22
	15	-726.71	-3033	10528.26

Table A.2: Vectors from each beam splitter to their respective mirror.

RICH Side	Camera	x	y	z
A-Side	00	-0.0657	1.49	0.1871
	01	-0.1388	0.91	0.3950
	02	-0.0804	0.96	0.2495
	03	-0.0675	0.96	0.2536
	04	-0.0596	0.86	-0.3184
	05	-0.0604	1.17	-0.3230
	06	-0.0641	0.84	-0.3425
	07	-0.0662	1.17	-0.3537
C-Side	08	0.0662	1.16	-0.3536
	09	0.0641	0.84	-0.3424
	10	0.0604	1.17	-0.3229
	11	0.0596	0.87	-0.3183
	12	0.0675	0.96	0.2537
	13	0.0804	0.96	0.2495
	14	0.1388	0.91	0.3950
	15	0.0892	0.80	0.2540

Table A.3: Central coordinate position of each CMOS camera.

RICH Side	Camera	x	y	z
A-Side	00	723.5762	3013.0	10537.1719
	01	798.0663	3013.0	10325.1631
	02	937.8592	3013.0	10527.0586
	03	1315.0054	3013.0	10637.9072
	04	1633.6711	3013.0	10919.8916
	05	1588.3218	3013.0	10677.5625
	06	2063.1873	3013.0	10969.8496
	07	2008.3594	3013.0	10676.8711
C-Side	08	-2008.3593	3013.0	10676.8710
	09	-2063.1873	3013.0	10969.8496
	10	-1588.3219	3013.0	10677.5625
	11	-1633.6710	3013.0	10919.8920
	12	-1315.0054	3013.0	10637.9072
	13	-937.8592	3013.0	10527.0586
	14	-798.0663	3013.0	10325.1631
	15	-723.5762	3013.0	10537.1719

Table A.4: Vectors from each CMOS camera centre to their respective mirror.

RICH Side	Camera	x	y	z
A-Side	00	-0.1022	-0.9513	0.2909
	01	-0.1022	-0.9513	0.2909
	02	-0.0946	-0.9513	0.2935
	03	-0.0793	-0.9513	0.2980
	04	-0.0567	-0.9513	-0.3031
	05	-0.0567	-0.9513	-0.3031
	06	-0.0567	-0.9513	-0.3031
	07	-0.0567	-0.9513	-0.3031
C-Side	08	0.0567	-0.9513	-0.3031
	09	0.0567	-0.9513	-0.3031
	10	0.0567	-0.9513	-0.3031
	11	0.0567	-0.9513	-0.3031
	12	0.0793	-0.9513	0.2980
	13	0.0946	-0.9513	0.2935
	14	0.1022	-0.9513	0.2909
	15	0.1022	-0.9513	0.2909

Table A.5: Rotations of coordinate system to define camera planes required for the ray tracing code.

RICH Side	Camera	Rotation x plane (rad)	Rotation y plane (rad)
A-Side	00	0.3	0.338
	01	0.3	0.338
	02	0.3	0.312
	03	0.3	0.260
	04	-0.3	-0.185
	05	-0.3	-0.185
	06	-0.3	-0.185
	07	-0.3	-0.185
C-Side	08	-0.3	0.185
	09	-0.3	0.185
	10	-0.3	0.185
	11	-0.3	0.185
	12	0.3	-0.260
	13	0.3	-0.312
	14	0.3	-0.338
	15	0.3	-0.338

Appendix B

Cross Contamination

Details of decay channel branching fractions that could potentially contribute to both the $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ lifetime measurements, through either misreconstruction or misidentification are given in Table B.1. The PID efficiencies and misidentification rates used for the misidentified backgrounds, and determined from data are provided in Tables B.2. The relative yields of each $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ background with respect to the specific signal peak, are given in Table B.3.

Decay mode	Sample size	$\mathcal{B}(\times 10^{-6})$	in Kπ	in KK
$B^0_d \to K^+\pi^-$	2114887	$19.55^{+0.54}_{-0.53}$ [36]		
$B^0_d \to \pi^+\pi^-$	2015992	5.11 ± 0.22 [36]		
$B^0_d \to K^+K^-$	527498	0.11 ± 0.78 [96]		
$B^0_s \to K^+K^-$	2024490	25.4 ± 3.7 [36]		
$B^0_s \to \pi^+\pi^-$	721496	0.73 ± 0.14 [36]		
$B^0_s \to \pi^+K^-$	1057992	5.38 ± 0.76 [96]		
$\Lambda_b \to p\pi^-$	523495	4.0 ± 1.3 [22]		
$\Lambda_b \to pK^-$	533998	6.2 ± 1.9 [22]		
$B^0_d \to \pi^+\pi^-\pi^0$	1012496	$50^{(*)a}$	3.6%	1.6%
$B^0_d \to K^+\pi^-\pi^0$	1036995	37.8 ± 3.2 [36]	92.1%	28.7%
$B^0_d \to K^+K^-\pi^0$	115497	$10^{(*)b}$	$< 0.1\%$	24.4%
$B^0_d \to K^0_S\pi^+\pi^-$	2038990	24.8 ± 1.0 [36]	$< 0.1\%$	$< 0.1\%$
$B^0_d \to K^0_S K^+K^-$	2038995	12.4 ± 1.2 [36]	$< 0.1\%$	0.1%
$B^0_d \to K^0_S K^+\pi^-$	2038491	3.2 ± 0.6 [36]	$< 0.1\%$	$< 0.1\%$
$B^+ \to \pi^+\pi^-\pi^+$	2220989	15.2 ± 1.4 [36]	0.2%	$< 0.1\%$
$B^+ \to \pi^+\pi^-K^+$	1690094	16.3 ± 2.0 [36]	2.2%	1.3%
$B^+ \to \pi^+K^-K^+$	1037995	5.0 ± 0.7 [36]	$< 0.1\%$	14.9%
$B^+ \to K^+K^-K^+$	1660492	32.5 ± 1.5 [36]	$< 0.1\%$	0.2%
$B^+ \to pp\pi^+$	1039495	$1.60^{+0.18}_{-0.17}$ [36]	$< 0.1\%$	$< 0.1\%$
$B^+ \to ppK^+$	1532488	5.48 ± 0.34 [36]	$< 0.1\%$	$< 0.1\%$
$B^0_s \to K^+\pi^-\pi^0$	1042000	$5^{(*)}$	1.8%	0.6%
$B^0_s \to K^+K^-\pi^0$	112997	$20^{(*)}$	$< 0.1\%$	27.8%
$B^0_s \to K^0_S K^+K^-$	2052990	2.1 ± 0.9 [114]	$< 0.1\%$	$< 0.1\%$
$B^0_s \to K^0_S K^+\pi^-$	2038488	48.5 ± 6.4 [114]	$< 0.1\%$	$< 0.1\%$
$B^0_s \to K^0_S\pi^+\pi^-$	2089490	6.0 ± 1.8 [114]	$< 0.1\%$	$< 0.1\%$
$B^0_s \to K^0_S p\bar{p}$	204000	$1^{(*)}$	$< 0.1\%$	$< 0.1\%$
$B^0_d \to K^{*0}\gamma$	2678188	43.3 ± 1.5 [22]	$< 0.1\%$	0.1%
$B^0_d \to \phi\gamma$	2629292	39 ± 5 [115]	$< 0.1\%$	$< 0.1\%$

Table B.1: List of decay channels that can contribute to the KK and Kπ as both specific, misidentified and partially reconstructed background. The last two columns give the fractional contribution of that decay to the partially reconstructed background for each signal channel. Footnotes: a) Listed as $< 7.2 \times 10^{-4}$ @ 90% C.L. in [22]. b) Listed as $< 19 \times 10^{-6}$ @ 90% C.L in [22].

Table B.2: PID efficiencies and misidentification rates as determined from the PID calibration software.

Efficiencies	
$\epsilon_{\mathrm{K}\pi}^{\mathrm{B}_d^0\to\mathrm{K}^+\pi^-}$	$42.28 \pm 0.004\ (\mathrm{stat}) \pm 0.98\ (\mathrm{syst})\ \%$
$\epsilon_{\mathrm{KK}}^{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-}$	$49.48 \pm 0.001\ (\mathrm{stat}) \pm 0.80\ (\mathrm{syst})\ \%$
$\epsilon_{pK}^{\Lambda_b\to p\mathrm{K}^-}$	$43.45 \pm 0.03\ (\mathrm{stat}) \pm 11.3\ (\mathrm{syst})\ \%$
Mis-ID rates	
$\omega_{\mathrm{K}\pi}^{\mathrm{B}_s^0\to\mathrm{K}^+\mathrm{K}^-}$	$0.73 \pm 0.002\ \ (\mathrm{stat}) \pm 0.07\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{K}\pi}^{\mathrm{B}_d^0\to\pi^+\pi^-}$	$1.11 \pm 0.002\ \ (\mathrm{stat}) \pm 0.08\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{K}\pi}^{\Lambda_b\to p\pi^-}$	$5.13 \pm 0.001\ \ (\mathrm{stat}) \pm 1.23\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{K}\pi}^{\Lambda_b\to p\mathrm{K}^-}$	$0.94 \pm 0.007\ \ (\mathrm{stat}) \pm 0.61\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{KK}}^{\mathrm{B}_d^0\to\mathrm{K}^+\pi^-}$	$1.31 \pm 0.001\ \ (\mathrm{stat}) \pm 0.07\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{KK}}^{\mathrm{B}_s^0\to\pi^+\mathrm{K}^-}$	$1.31 \pm 0.001\ \ (\mathrm{stat}) \pm 0.07\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{KK}}^{\Lambda_b\to p\mathrm{K}^-}$	$5.94 \pm 0.001\ \ (\mathrm{stat}) \pm 1.56\ (\mathrm{syst})\ \%$
$\omega_{\mathrm{KK}}^{\Lambda_b\to p\pi^-}$	$0.16 \pm 0.0004\ (\mathrm{stat}) \pm 0.09\ (\mathrm{syst})\ \%$

Table B.3: Yields of the main $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ backgrounds expressed relative to the main signal peak. The values in the second column are calculated with Equations 6.8 and 6.9. The values in the third column are determined from fits to yields in the KK, $\mathrm{K}\pi$ and pK spectra using Equation 6.10.

Relative yields in the $\mathrm{K}\pi$ fit compared to $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$

	from $\mathcal{B}$	from fits
$\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$	$(7.15 \pm 0.92) \times 10^{-2}$	$(7.24 \pm 0.35) \times 10^{-2}$
$\Lambda_b \to p\pi^-$	$(1.22 \pm 0.57) \times 10^{-2}$	-
$\mathrm{B}^0_d \to \pi^+\pi^-$	$(6.84 \pm 0.61) \times 10^{-3}$	-
$\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$	$(5.80 \pm 0.80) \times 10^{-3}$	$(6.10 \pm 0.60) \times 10^{-3}$
$\Lambda_b \to p\mathrm{K}^-$	$(3.5 \pm 2.6) \times 10^{-3}$	-
$\mathrm{B}^0_s \to \pi^+\pi^-$	$(2.54 \pm 0.55) \times 10^{-4}$	-
$\mathrm{B}^0_d \to \mathrm{K}^+\mathrm{K}^-$	$(1.06 \pm 0.45) \times 10^{-4}$	-

Relative yields in the KK fit compared to $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$

	from $\mathcal{B}$	from fits
$\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$	$(7.88 \pm 0.87) \times 10^{-2}$	$(7.50 \pm 0.45) \times 10^{-2}$
$\Lambda_b \to p\mathrm{K}^-$	$(5.6 \pm 2.6) \times 10^{-2}$	$(2.52 \pm 0.93) \times 10^{-2}$
$\mathrm{B}^0_d \to \mathrm{K}^+\mathrm{K}^-$	$(1.82 \pm 0.78) \times 10^{-2}$	-
$\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$	$(5.63 \pm 0.81) \times 10^{-3}$	$(5.43 \pm 0.42) \times 10^{-3}$
$\Lambda_b \to p\pi^-$	$(9.6 \pm 6.7) \times 10^{-4}$	-
$\mathrm{B}^0_d \to \pi^+\pi^-$	$(2.69 \pm 0.39) \times 10^{-4}$	-
$\mathrm{B}^0_s \to \pi^+\pi^-$	$(9.7 \pm 2.3) \times 10^{-6}$	-

Appendix C

Mass Fitter Parameters

Parameter details from the mass fits to 2011 data, using the 2011 lifetime fitter. The $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ parameters are given in Table C.1, with the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ parameters given in Table C.2. Parameter details from the $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ channel using 2011 data extracted using the $B_d^0 \to p\bar{p}$ mass fitter are given in Table C.3.

Parameter	Value		Origin	Description
$f_{\mathrm{B}_s^0}$	0.4654 ± 0.0044		Data	B_s^0 signal fraction
$f_{\mathrm{B}_d^0}$	0.0084 ± 0.0001		Data	$\mathrm{B}_d^0 \to \mathrm{K}^+\mathrm{K}^-$ signal fraction
$f_{\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-}$	0.0349 ± 0.0003		Data	$\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ Mis-ID fraction
$f_{\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-}$	0.00251 ± 0.00002		Data	$\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ Mis-ID fraction
$f_{\Lambda_b \to p\mathrm{K}^-}$	0.0117 ± 0.0001		Data	$\Lambda_b \to p\mathrm{K}^-$ Mis-ID fraction
f_{part}	0.0912 ± 0.0101		Data	Partially reconstructed fraction
$\mu_{\mathrm{B}_s^0}$	5372.86 ± 0.26	MeV	Data	B_s^0 mass
σ_{B}	20.97 ± 0.25	MeV	Data	Signal mass resolution
$\alpha_{Low}^{\mathrm{B}_s^0}$	1.387 ± 0.089		MC11a	Low CB part boundary (B_s^0)
$\alpha_{High}^{\mathrm{B}_s^0}$	1.885 ± 0.070		MC11a	High CB part boundary (B_s^0)
$n_{Low}^{\mathrm{B}_s^0}$	1.647 ± 0.059		MC11a	Low CB part tail (B_s^0)
$n_{High}^{\mathrm{B}_s^0}$	6.772 ± 1.272		MC11a	High CB part tail (B_s^0)
$fraction_{Low}^{\mathrm{B}_s^0}$	0.2958 ± 0.0421		MC11a	MC11a fraction of low CB (B_s^0)
μ_{KK}^{EMG}	5208.60 ± 11.01	MeV	Data	EMG peak
$\sigma_{\mathrm{KK}}^{EMG}$	60 ± 22	MeV	MC11a	EMG smearing
$\lambda_{\mathrm{KK}}^{EMG}$	0.0025 ± 0.0014	MeV^{-1}	Data	EMG tail
$\nabla_{comb}^{\mathrm{KK}}$	$(-3.605 \pm 1.632) \cdot 10^{-7}$	MeV^{-1}	Data	Gradient of the combinatorial

Table C.1: Description and fitted values of the parameters of the mass fit to the KK spectrum. The "Origin" column states if the value is determined from MC, and therefore fixed in the fit, or from the fit to the data.

Parameter	Value		Origin	Description
$f_{\mathrm{B}^0_d}$	0.4327 ± 0.0028		Data	B^0_d signal fraction
$f_{\mathrm{B}^0_s}$	0.0312 ± 0.0014		Data	B^0_s signal fraction
$f_{\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-}$	0.00264 ± 0.00002		Data	$\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ mis ID fraction
$f_{\mathrm{B}^0_d \to \pi^+\pi^-}$	0.00294 ± 0.00002		Data	$\mathrm{B}^0_d \to \pi^+\pi^-$ mis ID fraction
$f_{\Lambda_b \to p\pi^-}$	0.00528 ± 0.00003		Data	$\Lambda_b \to p\pi^-$ mis ID fraction
f_{part}	0.08506 ± 0.00291		Data	Partially reconstructed fraction
$\mu_{\mathrm{B}^0_d}$	5285.50 ± 0.17	MeV/c^2	Data	B^0_d mass
$\mu_{\mathrm{B}^0_s}$	5369.18 ± 1.25	MeV/c^2	Data	B^0_s mass
$\sigma_{\mathrm{B}^0_d}$	20.85 ± 0.16	MeV/c^2	Data	B^0_d signal mass resolution
$\sigma_{\mathrm{B}^0_s}$	21.21 ± 0.17	MeV/c^2	Data	B^0_s signal mass resolution
$\alpha^{\mathrm{B}^0_d}_{Low}$	1.172 ± 0.091		MC11a	Low CB part boundary (B^0_d)
$\alpha^{\mathrm{B}^0_d}_{High}$	1.791 ± 0.059		MC11a	High CB part boundary (B^0_d)
$n^{\mathrm{B}^0_d}_{Low}$	1.854 ± 0.089		MC11a	Low CB part tail (B^0_d)
$n^{\mathrm{B}^0_d}_{High}$	8.335 ± 1.480		MC11a	High CB part tail (B^0_d)
$fraction^{\mathrm{B}^0_d}_{Low}$	0.2911 ± 0.0412		MC11a	MC11a fraction of low CB (B^0_d)
$\alpha^{\mathrm{B}^0_s}_{Low}$	0.9951 ± 0.1470		MC11a	Low CB part boundary (B^0_s)
$\alpha^{\mathrm{B}^0_s}_{High}$	1.949 ± 0.088		MC11a	High CB part boundary (B^0_s)
$n^{\mathrm{B}^0_s}_{Low}$	2.033 ± 0.137		MC11a	Low CB part tail (B^0_s)
$n^{\mathrm{B}^0_s}_{High}$	5.898 ± 1.166		MC11a	High CB part tail (B^0_s)
$fraction^{\mathrm{B}^0_s}_{Low}$	0.2244 ± 0.0528		MC11a	MC11a fraction of low CB B^0_s)
$\mu^{EMG_1}_{\mathrm{K}\pi}$	5137.3 ± 1.3	MeV/c^2	MC11a	EMG peak
$\sigma^{EMG_1}_{\mathrm{K}\pi}$	19.95 ± 0.94	MeV/c^2	MC11a	EMG smearing
$\lambda^{EMG_1}_{\mathrm{K}\pi}$	$(4.16 \pm 0.55) \cdot 10^{-3}$	$(\mathrm{MeV}/c^2)^{-1}$	MC11a	EMG tail
$\mu^{EMG_2}_{\mathrm{K}\pi}$	5166.6 ± 7.7	MeV/c^2	MC11a	EMG peak
$\sigma^{EMG_2}_{\mathrm{K}\pi}$	43.6 ± 4.3	MeV/c^2	MC11a	EMG smearing
$\lambda^{EMG_2}_{\mathrm{K}\pi}$	$(7.3 \pm 7.4) \cdot 10^{-5}$	$(\mathrm{MeV}/c^2)^{-1}$	MC11a	EMG tail
$R^{EMG_{12}}_{\mathrm{K}\pi}$	11.7		MC11a	Ratio of the two EMGs
$\nabla^{\mathrm{K}\pi}_{comb}$	$(-1.099 \pm 0.052) \cdot 10^{-6}$	$(\mathrm{MeV}/c^2)^{-1}$	Data	Gradient of the combinatorial

Table C.2: Description and fitted values of the parameters of the mass fit to the $\mathrm{K}\pi$ spectrum. The "Origin" column states if the value is determined from MC, and therefore fixed in the fit, or from the fit to the data.

Parameter	Value		Origin	Description
$f_{\mathrm{B}^0_d}$	0.4304 ± 0.0029		Data	B^0_d signal fraction
$f_{\mathrm{B}^0_s}$	0.03053 ± 0.00141		Data	B^0_s signal fraction
f^{KK}_{misID}	$(2.5839 \pm 0.0171) \cdot 10^{-3}$		Data/MC	$\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ misID fraction
$f^{\pi\pi}_{misID}$	$(2.9545 \pm 0.0196) \cdot 10^{-3}$		Data/MC	$\mathrm{B}^0_d \to \pi^+\pi^-$ misID fraction
$f^{p\pi}_{misID}$	$(5.952 \pm 0.040) \cdot 10^{-3}$		Data/MC	$\Lambda_b \to p\pi^-$ misID fraction
f_{part}	0.0847 ± 0.0030		Data	Partially reconstructed fraction
$\mu_{\mathrm{B}^0_d}$	5285.600 ± 0.047	MeV/c^2	Data	B^0_d signal peak mean
$\mu_{\mathrm{B}^0_s}$	5369.19 ± 1.30	MeV/c^2	Data	B^0_s signal peak mean
$\sigma_{\mathrm{B}^0_d}$	20.955 ± 0.168	MeV/c^2	Data	B^0_d signal width
$\sigma_{\mathrm{B}^0_s}$	21.332 ± 0.171	MeV/c^2	MC11a	B^0_s signal width (from B^0_d width)
$\alpha^{\mathrm{B}^0_d}_{Low}$	1.161 ± 0.091		MC11a	Low CB part boundary (B^0_d)
$\alpha^{\mathrm{B}^0_d}_{High}$	1.806 ± 0.059		MC11a	High CB part boundary (B^0_d)
$n^{\mathrm{B}^0_d}_{Low}$	1.850 ± 0.070		MC11a	Low CB part tail (B^0_d)
$n^{\mathrm{B}^0_d}_{High}$	8.080 ± 1.400		MC11a	High CB part tail (B^0_d)
$frac^{\mathrm{B}^0_d}_{Low}$	0.2844 ± 0.0401		MC11a	MC fraction of low CB (B^0_d)
$\alpha^{\mathrm{B}^0_s}_{Low}$	0.9989 ± 0.1460		MC11a	Low CB part boundary (B^0_s)
$\alpha^{\mathrm{B}^0_s}_{High}$	1.960 ± 0.0090		MC11a	High CB part boundary (B^0_s)
$n^{\mathrm{B}^0_s}_{Low}$	2.036 ± 0.135		MC11a	Low CB part tail (B^0_s)
$n^{\mathrm{B}^0_s}_{High}$	5.777 ± 1.154		MC11a	High CB part tail (B^0_s)
$frac^{\mathrm{B}^0_s}_{Low}$	0.2265 ± 0.0531		MC11a	MC fraction of low CB (B^0_s)
$\mu^{EMG_1}_{\mathrm{K}\pi}$	5137.3 ± 1.3	MeV/c^2	MC11a	EMG peak
$\sigma^{EMG_1}_{\mathrm{K}\pi}$	19.95 ± 0.95	MeV/c^2	MC11a	EMG smearing
$\lambda^{EMG_1}_{\mathrm{K}\pi}$	$(4.16 \pm 0.55) \cdot 10^{-3}$	$(\mathrm{MeV}/c^2)^{-1}$	MC11a	EMG tail
$\mu^{EMG_2}_{\mathrm{K}\pi}$	5166.7 ± 7.7	MeV/c^2	MC11a	EMG peak
$\sigma^{EMG_2}_{\mathrm{K}\pi}$	43.6 ± 4.3	MeV/c^2	MC11a	EMG smearing
$\lambda^{EMG_2}_{\mathrm{K}\pi}$	$(7.3 \pm 7.4) \cdot 10^{-5}$	$(\mathrm{MeV}/c^2)^{-1}$	MC11a	EMG tail
$R^{EMG_{12}}_{\mathrm{K}\pi}$	11.7		MC11a	Ratio of the two EMGs
$\nabla^{\mathrm{K}\pi}_{comb}$	$(-1.089 \pm 0.052) \cdot 10^{-6}$	$(\mathrm{MeV}/c^2)^{-1}$	Data	Gradient of the combinatorial

Table C.3: Description and fitted values of the parameters of the mass fit to the $\mathrm{K}\pi$ spectrum for the $B^0_{d/s} \to p\bar{p}$ control channel. The "Origin" column states if the value is determined from MC, and therefore fixed in the fit, or from the fit to the data.

Appendix D

Verification Fit Parameters

Pull distributions for the mass fit variables from $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC using the 2010 lifetime fitter, Figure D.1. Pull distributions for the mass fit variables from $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ toy MC using the 2011 lifetime fitter, Figure D.2. Pull distributions for the mass fit variables from $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC using the 2011 lifetime fitter, Figure D.3.

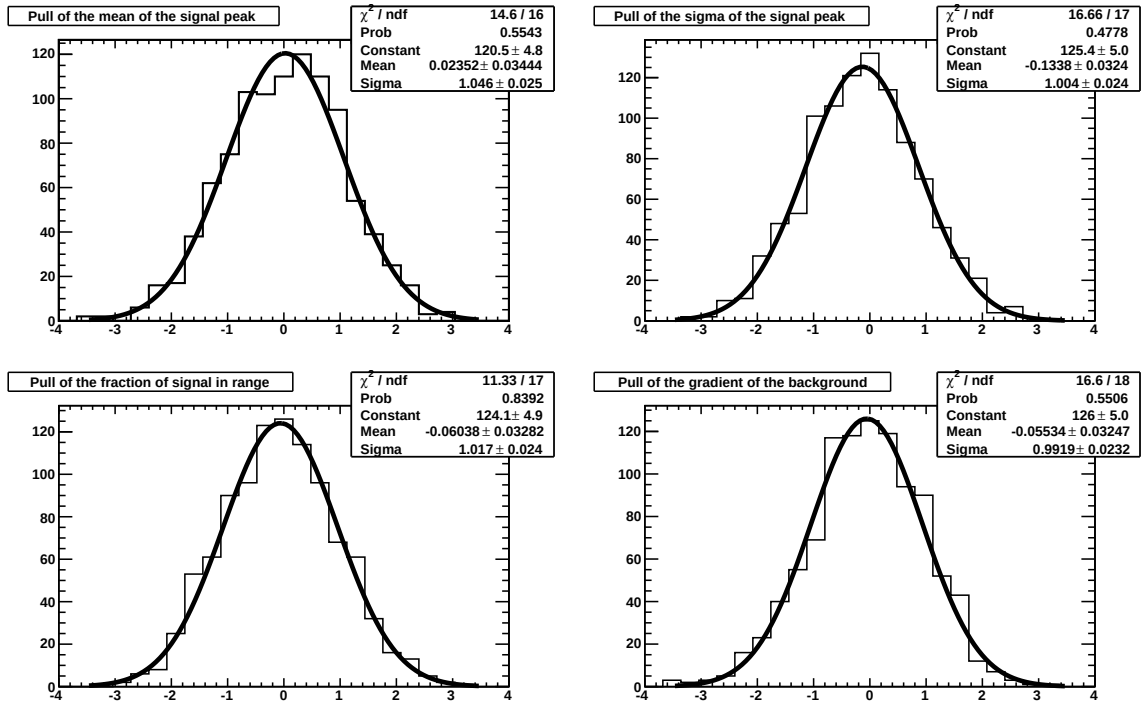


Figure D.1: Verification mass fit results from 1000 $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ toy MC datasets for the 2010 analysis.

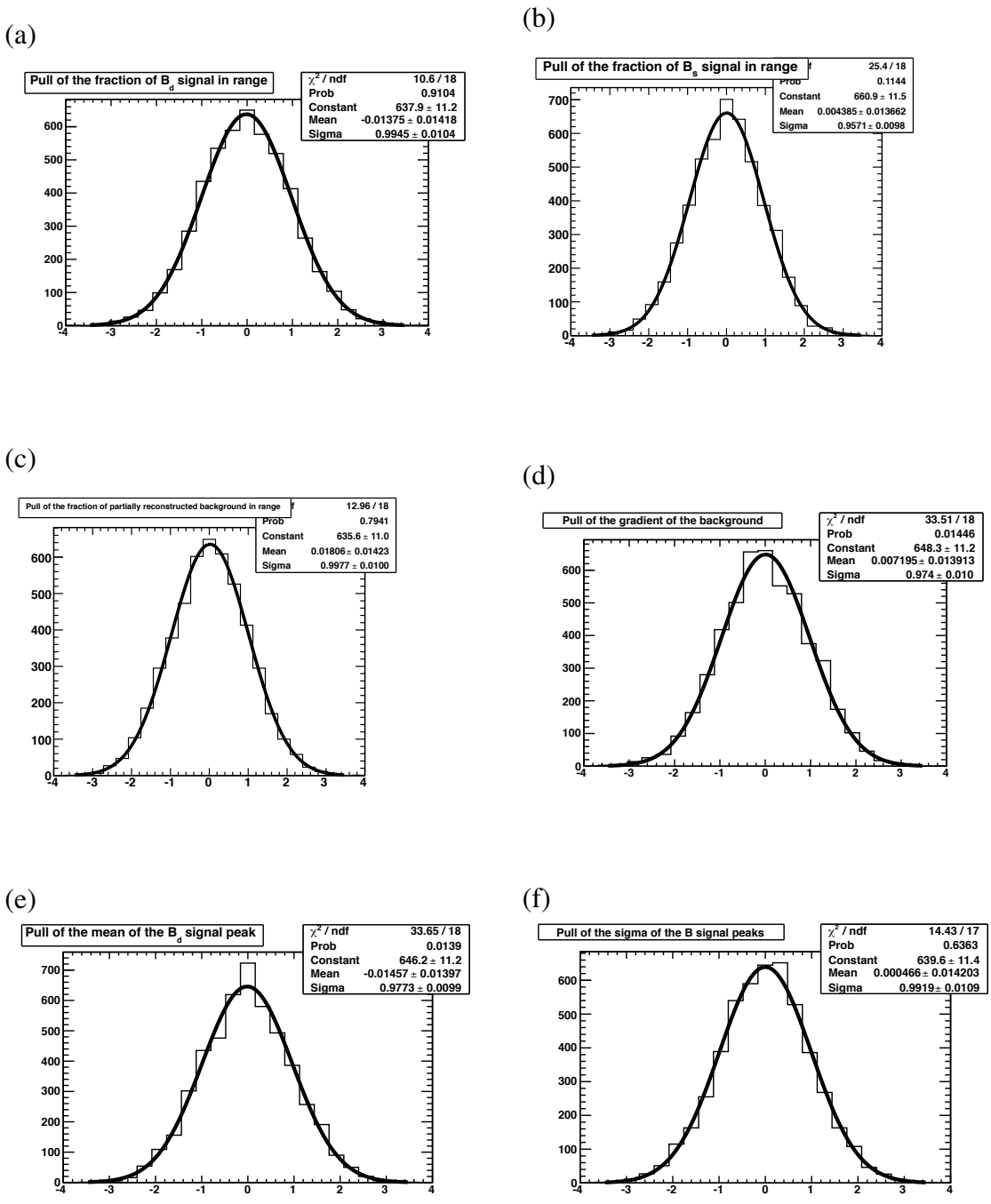


Figure D.2: Pull distributions of the fitted free parameters from 5000 $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ toy MC datasets for the 2011 analysis. $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ signal fraction (a), $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ signal fraction (b), partially reconstructed background signal fraction (c), gradient of the combinatorial background (d), $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ signal mean (e), $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ signal σ (f).

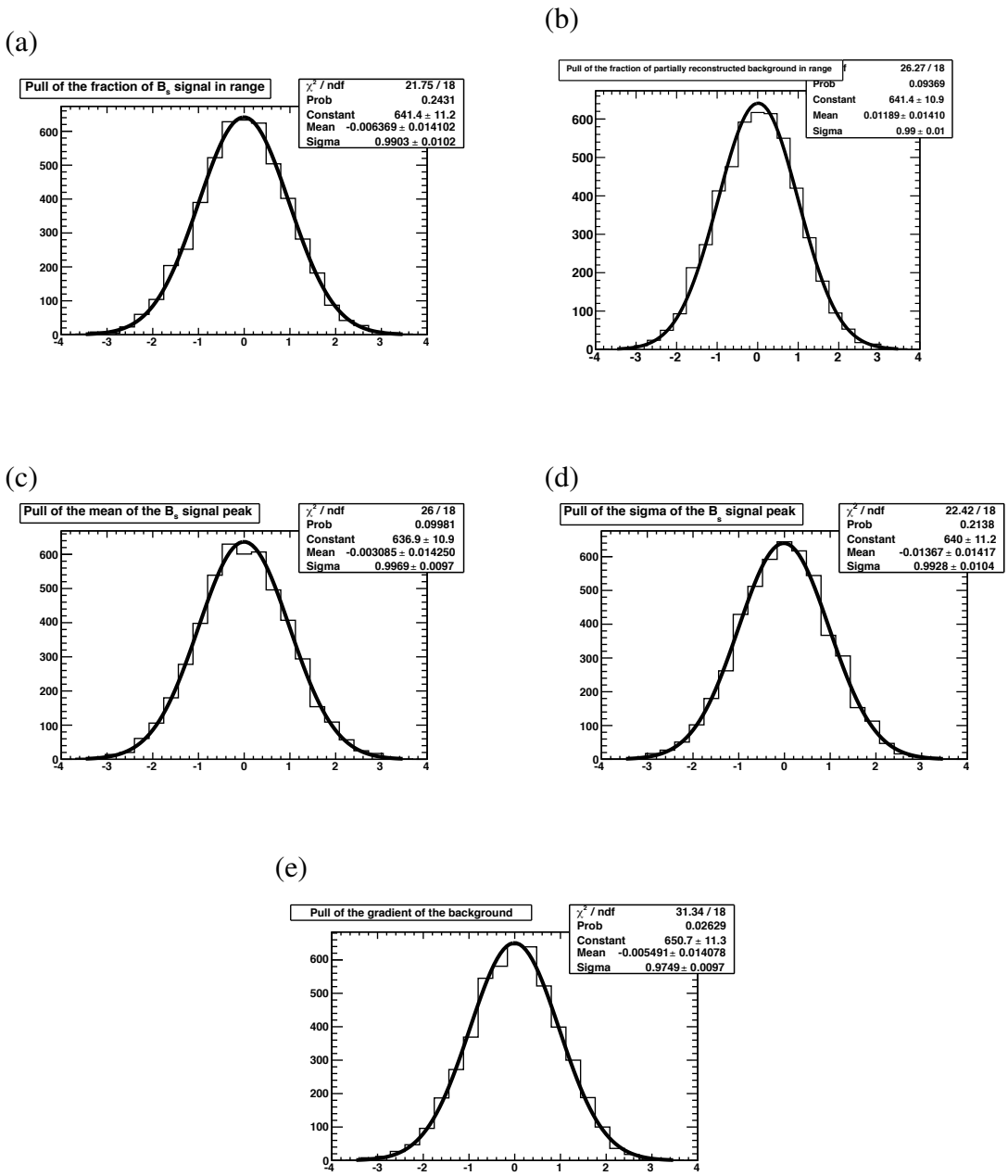


Figure D.3: Pull distributions of the fitted free parameters from 5000 $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ toy MC datasets for the 2011 analysis. $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ signal fraction (a), partially reconstructed signal fraction (b), $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ mass mean (c), $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ mass sigma (d), combinatorial background gradient (e).

Appendix E

Consistency across Magnet and Trigger Configurations

Due to the high precision of the 2011 lifetime measurements, studies, in addition to systematics, are required to ensure consistency between experimental configurations. These studies were conducted to ensure that changes in magnet polarity and trigger configuration during data taking did not have an impact on the fitted lifetime.

To ascertain if there was an effect that could originate from magnet polarity changes during data taking, the full datasets for the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ were split into data subsets collected with each magnet polarity. A fit to each individual data subset was performed, with a weighted average lifetime then determined for each magnet polarity. These fits, along with the calculated weighted average lifetime and default lifetime fit on the full dataset, are shown for the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ decays in Figure E.1.

The fitted lifetimes for each category, for each decay, are all within 2 sigma of the weighted average lifetime for each polarity and fit using the full dataset. Thus we state that this has no effect on the measured lifetime.

As discussed in Sections 2.10 and 5.3.1, the trigger can be split into a lower level hardware and higher level software trigger. To check for consistencies across different trigger configurations, studies are performed using the fitted lifetime subject to different trigger categories. These categories are split over the L0 and HLT. The studies were not able to be conducted for the 2010 lifetime analysis, due to the size of the dataset.

The L0 trigger employed in the 2011 analysis is composed of a TIS event *or* TOS event (defined in Section 5.3.1), with TIS events not subject to the same biases that can effect TIS events. The data is split into three categories: The events must be selected by the L0_TIS only, the L0_TOS only and the default trigger configuration, L0_TIS *or* L0_TOS. The resultant fits of the three lifetimes, $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$, for each

category are shown in Figure E.2. As the fitted lifetimes for each category, for each decay, are consistent within the uncertainties, it can be determined that the L0 triggering does not introduce any bias on the fitted lifetime.

To study the effect of HLT trigger variations, it is necessary to split the total dataset up into 7 number of datasets that relate to a unique trigger configuration. This is done as trigger configurations change over time, due to reoptimisations based on data taking conditions, so inconsistencies between these can be determined.

The resultant fits of the three lifetimes, for each trigger configuration, are shown in Figure E.3. As the fitted lifetimes for each category, for each decay, are consistent within the uncertainties, we can state that changes in the HLT trigger during data taking has negligible impact.

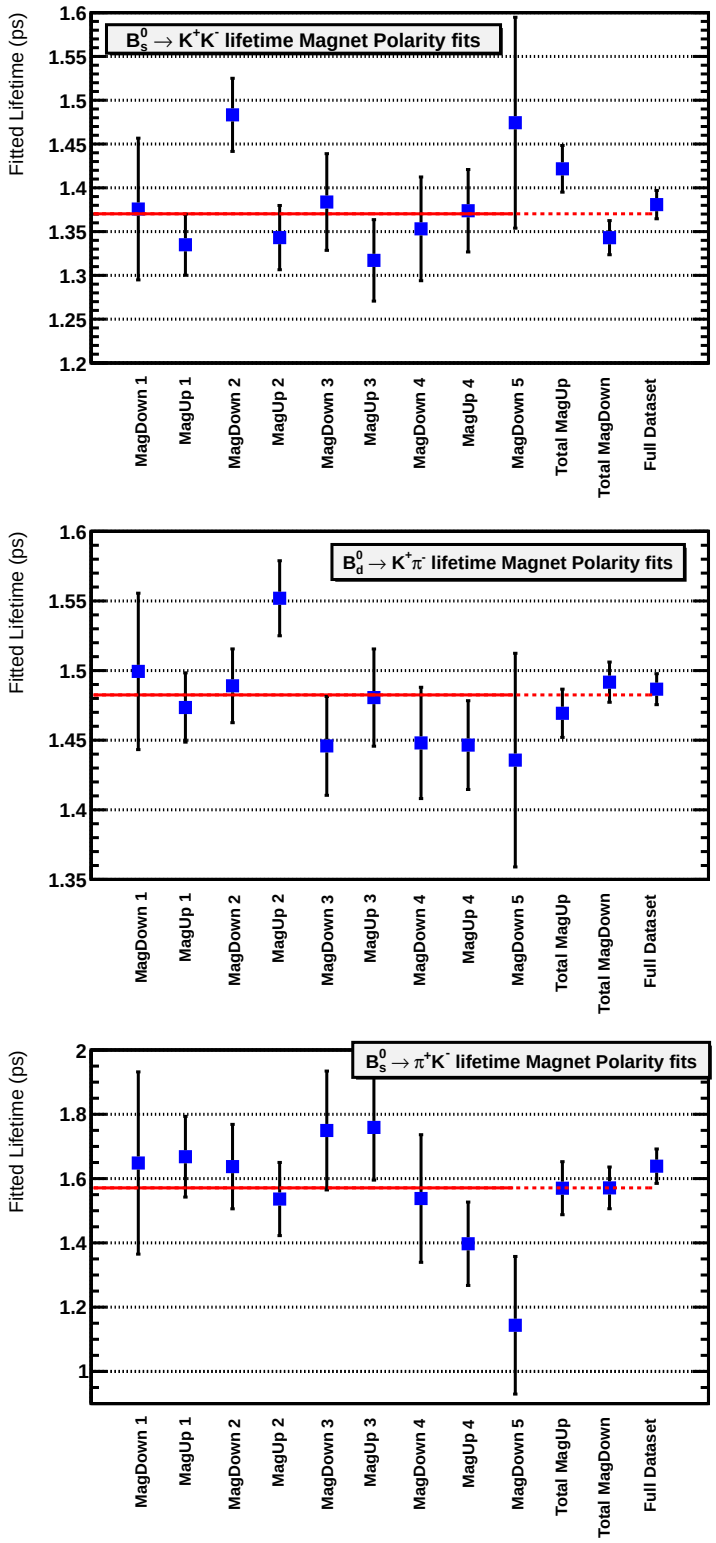


Figure E.1: The (blind) fitted lifetime of $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ (top), $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ (middle) and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ (bottom) as a function data subsets split by magnet polarity (details provided in [89]). The red line indicates a fit of a constant value to the lifetime distribution from the magnet polarity separated datasets, with the dotted line being an extrapolation to the lifetime for each polarity and the full dataset.

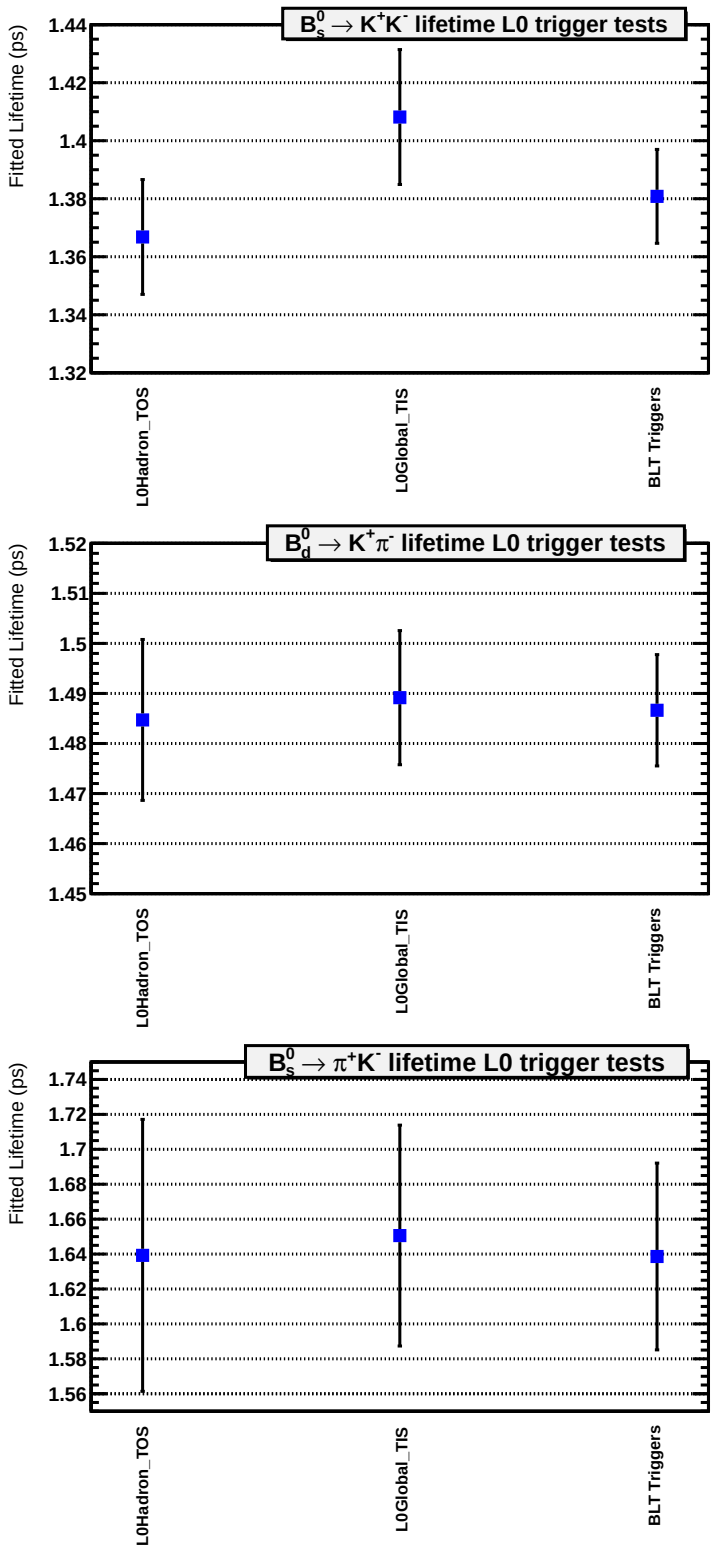


Figure E.2: The (blind) fitted lifetime of $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ (top), $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ (middle) and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ (bottom) as a function of the L0 trigger criteria applied to the $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ data. The two first categories are exclusive trigger criteria and the one labelled 'BLT' (B lifetime) is the logical OR of the other two trigger lines and the default trigger selection.

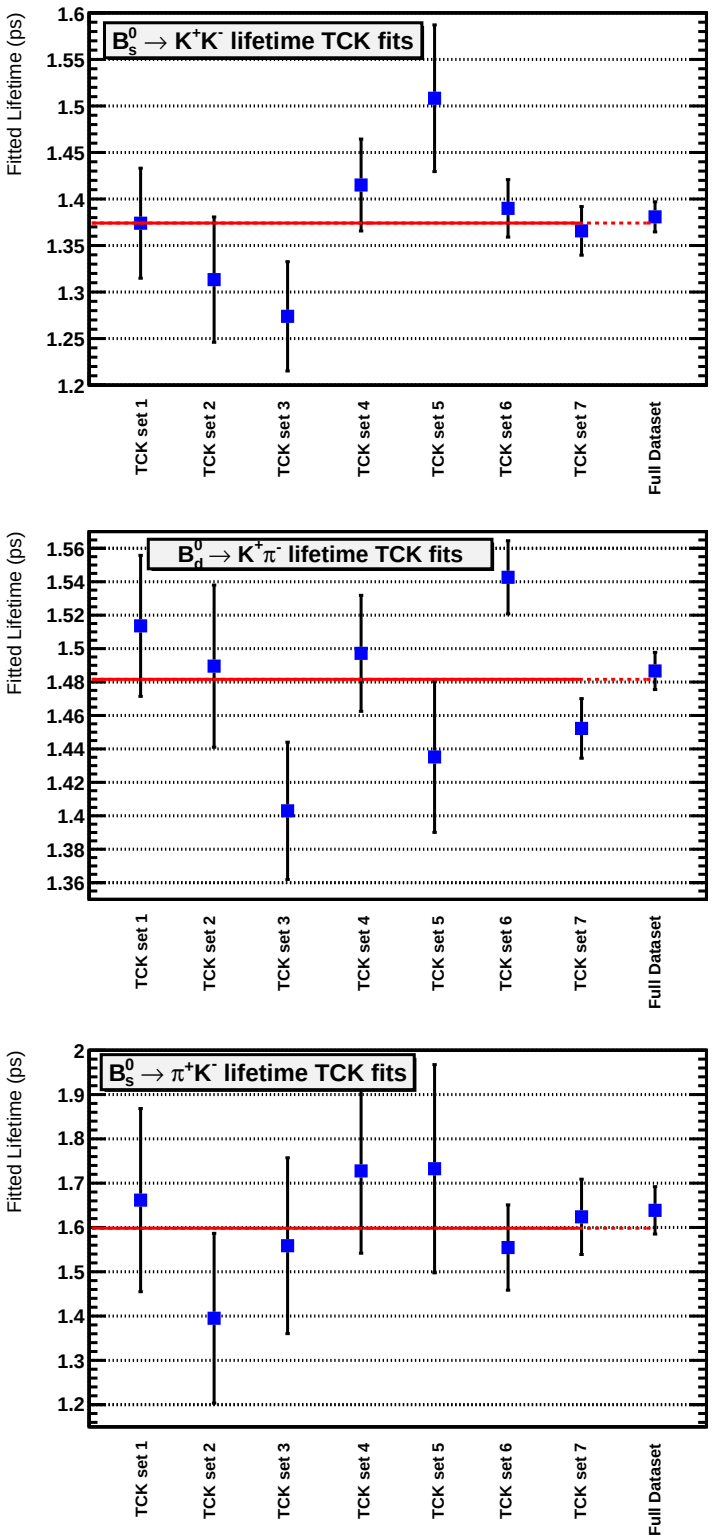


Figure E.3: The (blind) fitted lifetime of $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$ (top), $\mathrm{B}_d^0 \to \mathrm{K}^+\pi^-$ (middle) and $\mathrm{B}_s^0 \to \pi^+\mathrm{K}^-$ (bottom) for the seven categories of trigger configurations. The red line indicates a fit of a constant value to the lifetimes from the different trigger configurations, with the dotted line being an extrapolation to the lifetime fitted from the full dataset.

Bibliography

[1] J. Dalton, "A New System of Chemical Philosophy", S. Russell, 1808, {http://books.google.com/books?id=Wp7QAAAAMAAJ}.

[2] E. Rutherford, "The Scattering of α and β rays by Matter and the Structure of the Atom", *Philos. Mag.* **6** (1911) 21.

[3] J. J. Thomson, "The Positive Electron", *Philos. Mag.* **39** (1897) 293.

[4] P. A. M. Dirac, "The Quantum Theory of the Electron", *Proceedings of the Royal Society of London: Series A* **117(778)** (1928) 610–624.

[5] C. D. Anderson, "Cathode Rays", *Phys. Rev.* **43** (1933) 491–494.

[6] E598 Collaboration, J. Aubert *et al*, "Experimental Observation of a Heavy Particle J", *Phys.Rev.Lett.* **33** (1974) 1404–1406.

[7] SLAC-SP-017 Collaboration, J. Augustin *et al*, "Discovery of a Narrow Resonance in e+ e- Annihilation", *Phys.Rev.Lett.* **33** (1974) 1406–1408.

[8] S. W. Herb *et al*, "Observation of a dimuon resonance at 9.5 GeV in 400 GeV proton-nucleus collisions", *Phys. Rev. Lett.* **39** (Aug, 1977) 252–255, http://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[9] CDF Collaboration, F. Abe *et al*, "Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab", *Phys. Rev. Lett.* **74** (Apr, 1995) 2626–2631, http://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[10] D0 Collaboration, S. Abachi *et al*, "Search for high mass top quark production in $p\overline{p}$ collisions at $\sqrt{s} = 1.8\,\mathrm{TeV}$", *Phys. Rev. Lett.* **74** (Mar, 1995) 2422–2426, http://link.aps.org/doi/10.1103/PhysRevLett.74.2422.

[11] DONUT Collaboration, "Observation of tau neutrino interactions", *Physics Letters B* **504** (April, 2001) 218–224, arXiv:hep-ex/0012035.

[12] ATLAS Collaboration, G. Aad *et al*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[13] CMS Collaboration, S. Chatrchyan *et al*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[14] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132–133.

[15] I. I. Bigi and A. I. Sanda, "CP Violation", Cambridge University Press, 2000.

[16] A. D. Sakharov, "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe", *J. Exp. Theor. Phys. Lett.* **5** (1967) 24–27.

[17] Burles, Scott and Nollett, Kenneth M. and Turner, Michael S, "Big bang nucleosynthesis predictions for precision cosmology", *Astrophys.J.* **552** (2001) L1–L6, arXiv:astro-ph/0010171 [astro-ph].

[18] Riotto, Antonio and Trodden, Mark, "Recent progress in baryogenesis", *Ann.Rev.Nucl.Part.Sci.* **49** (1999) 35–75, arXiv:hep-ph/9901362 [hep-ph].

[19] http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, April, 2012.

[20] Pauli, W, "The Connection Between Spin and Statistics", *Phys.Rev.* **58** (1940) 716–722.

[21] http://en.wikipedia.org/wiki/File:Bosons-Hadrons-Fermions-RGB-png3.png, February, 2013.

[22] Particle Data Group, J. Beringer *et al*, "Review of Particle Physics (RPP)", *Phys.Rev.* **D86** (2012) 010001.

[23] Schwinger, Julian S, "A Theory of the Fundamental Interactions", *Annals Phys.* **2** (1957) 407–434.

[24] Kobayashi, Makoto and Maskawa, Toshihide, "CP Violation in the Renormalizable Theory of Weak Interaction", *Prog.Theor.Phys.* **49** (1973) 652–657.

[25] Maki, Ziro and Nakagawa, Masami and Sakata, Shoichi, "Remarks on the unified model of elementary particles", *Prog.Theor.Phys.* **28** (1962) 870–880.

[26] Cabibbo, Nicola, "Unitary Symmetry and Leptonic Decays", *Phys.Rev.Lett.* **10** (1963) 531–533.

[27] Glashow, S.L. and Iliopoulos, J. and Maiani, L, "Weak Interactions with Lepton-Hadron Symmetry", *Phys.Rev.* **D2** (1970) 1285–1292.

[28] Chau, Ling-Lie and Keung, Wai-Yee, "Comments on the Parametrization of the Kobayashi-Maskawa Matrix", *Phys.Rev.Lett.* **53** (1984) 1802.

[29] `http://ckmfitter.in2p3.fr/www/results/plots_fpcp13/ckm_res_fpcp13.html`, July, 2013.

[30] LHCb Collaboration, B. Adeva *et al*, "Roadmap for selected key measurements of LHCb", `arXiv:0912.4179 [hep-ex]`.

[31] Lenz, Alexander and Nierste, Ulrich, "Theoretical update of $B_s - \bar{B}_s$ mixing", *JHEP* **0706** (2007) 072, `arXiv:hep-ph/0612167 [hep-ph]`.

[32] Donega, Mauro, "Measurement of the $B_s \to K^+K^-$ lifetime and extraction of the $\Delta\Gamma_{CP}/\Gamma_{CP}$ at CDF Run II and development of the ATLAS-SCT endcap modules", PhD thesis, Geneva U, 2006, CERN-THESIS-2006-096.

[33] K. Anikeev *et al*, "B physics at the Tevatron: Run II and beyond", `arXiv:hep-ph/0201071 [hep-ph]`.

[34] K. Hartkorn and H. G. Moser, "A new method of measuring $\Delta\Gamma_s/\Gamma_s$ in the B_s^0 $\bar{\mathrm{B}}_s^0$ system", *Eur. Phys. J. C* **8** no. 381, (1999) .

[35] Fleischer, Robert and Knegjens, Robert, "In Pursuit of New Physics with $\mathrm{B}_s^0 \to \mathrm{K}^+\mathrm{K}^-$", *Eur.Phys.J.* **C71** (2011) 1532, `arXiv:1011.1096 [hep-ph]`.

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

[37] Evans, Lyndon and Bryant, Philip, "LHC Machine", *JINST* **3** (2008) S08001.

[38] "CERN: The European Organization for Nuclear Research", `http://www.cern.ch`, March, 2013.

[39] "The LEP Collider", `http://sl-div.web.cern.ch/sl-div/history/lep_doc.html`, November, 1990.

[40] http://keyhole.web.cern.ch/keyhole/reception/CERN.html, March, 2013.

[41] "CERN FAQ : LHC the guide", http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf, February, 2009.

[42] ATLAS Collaboration, G. Aad *et al*, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003.

[43] CMS Collaboration, S. Chatrchyan *et al*, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004.

[44] LHCb Collaboration, J. Alves, A. Augusto *et al*, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005.

[45] ALICE Collaboration, K. Aamodt *et al*, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002.

[46] TOTEM Collaboration, G. Anelli *et al*, "The TOTEM experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08007.

[47] LHCf Collaboration, O. Adriani *et al*, "The LHCf detector at the CERN Large Hadron Collider", *JINST* **3** (2008) S08006.

[48] LHCb Collaboration, R. Aaij *et al*, "Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region", *Phys.Lett.* **B694** (2010) 209–216, arXiv:1009.2731 [hep-ex].

[49] LHCb Collaboration, R. Aaij *et al*, "Absolute luminosity measurements with the LHCb detector at the LHC", *JINST* **7** (2012) P01010, arXiv:1110.2866 [hep-ex].

[50] LHCb Collaboration, S. Amato *et al*, "LHCb technical proposal", CERN-LHCC-98-04, CERN-LHCC-P-4.

[51] https://lbtwiki.cern.ch/bin/view/VELO/VELOConferencePlots#EventDisplay, June, 2011.

[52] https://lbtwiki.cern.ch/pub/VELO/VELOConferencePlots/randphisensors.pdf, August, 2013.

[53] P. Rodríguez Pérez, "The LHCb Vertex Locator performance and Vertex Locator upgrade", *Journal of Instrumentation* **7** (Dec, 2012) C2008, arXiv:1209.4845 [physics.ins-det].

[54] https://lbtwiki.cern.ch/pub/VELO/VELOConferencePlots/IPX-Resolution-Vs-PT-Compare2011DataToSimulation.pdf, May, 2013.

[55] "ST Material for Publications", http://lhcb.physik.uzh.ch/ST/public/material/index.php, August, 2011.

[56] J. D. Jackson, "Classical Electrodynamics", John Wiley and Sons, 1975.

[57] M. Adinolfi *et al*, "Performance of the LHCb RICH detector at the LHC", *European Physical Journal C* **73** (May, 2013) 24–31, arXiv:1211.6759 [physics.ins-det].

[58] J. Badier *et al*, "Shashlik calorimeter Beam-test results", *Nuclear Instruments and Methods in Physics Research A* **348** (Aug, 1994) 74–86.

[59] LHCb Collaboration, "LHCb calorimeters: Technical design report", CERN-LHCC-2000-036.

[60] http://www2.dupont.com/Tyvek/en_US/index.html, January, 2012.

[61] Patel, M. and Losasso, M. and Gys, T, "Magnetic shielding studies of the RICH photon detectors", CERN-LHCB-2005-055.

[62] LHCb Collaboration, "LHCb muon system technical design report", CERN-LHCC-2001-010.

[63] A. A. Alves, Jr *et al*, "Performance of the LHCb muon system", *Journal of Instrumentation* **8** (Feb, 2013) 2022P, arXiv:1211.1346 [physics.ins-det].

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

[65] http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/, December, 2013.

[66] http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/, December, 2013.

[67] LHCb Collaboration, "LHCb: RICH technical design report", CERN-LHCC-2000-037.

[68] Aglieri Rinella, Gianluca, "Development of the photon detection, acquisition and optical systems of modern ring imaging Cherenkov detectors", PhD thesis, University of Palermo, 2005, CERN-THESIS-2005-060.

[69] A. Papanestis, "Laser Alignment Monitoring System (LAMS) for the RICH", 2nd LHC detector alignment workshop, June, 2007.

[70] Macgregor, Andrew Alan, "The laser mirror alignment system for the LHCb RICH detectors", PhD thesis, University of Glasgow, 2006.

[71] Blanks, C. and Soomro, F. and Nardulli, J, "The Laser Alignment and Monitoring System for the LHCb RICH detectors", LHCb-PUB-2011-014, CERN-LHCb-PUB-2011-014.

[72] "Sellmeier Equation", http://en.wikipedia.org/wiki/Sellmeier_equation.

[73] "RefractiveindexInfo", http://refractiveindex.info/.

[74] Dr Christoph Frei, "Private correspondence", 2013.

[75] The LHCb Collaboration, "Measurement of the effective $B_s^0 \to K^+K^-$ lifetime", LHCb-CONF-2011-018.

[76] LHCb Collaboration, R. Aaij *et al*, "Measurement of the effective $B_s^0 \to K^+K^-$ lifetime", *Phys.Lett.* **B707** (2012) 349–356, arXiv:1111.0521 [hep-ex].

[77] CDF collaboration, "Measurement of branching fractions and direct CP asymmetries of $\mathrm{B} \to \mathrm{h}^+\mathrm{h}^-$ decays in 1 fb^{-1}", 2006 CDF public note 8579v1.

[78] "MINUIT : Function Minimization and Error Analysis", http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html.

[79] M. Pivk and F. R. Le Diberder, "sPlots: A statistical tool to unfold data distributions", *Nuclear Instruments and Methods in Physics Research A* **555** (Dec, 2005) 356–369, arXiv:physics/0402083.

[80] Y. Xie, "sFit: a method for background subtraction in maximum likelihood fit", *ArXiv e-prints* (May, 2009) , arXiv:0905.0724 [physics.data-an].

[81] R. Bailey *et al*, "Measurement of the lifetime of charged and neutral D mesons with high resolution silicon strip detectors", *Z.Phys.* **C28** (1985) 357–363.

[82] Rademacker, Jonas, "Reduction of statistical power per event due to upper lifetime cuts in lifetime measurements", *Nucl.Instrum.Meth.* **A570** (2007) 525–528, arXiv:hep-ex/0502042 [hep-ex].

[83] CDF Collaboration, T. Aaltonen *et al*, "Measurement of the B^- lifetime using a simulation free approach for trigger bias correction", *Phys.Rev.* **D83** (2011) 032008, arXiv:1004.4855 [hep-ex].

[84] Gligorov, Vladimir Vava, "Measurement of the CKM angle gamma and B meson lifetimes at the LHCb detector", PhD thesis, Oxford University, 2008.

[85] Gersabeck, M, "Alignment of the LHCb Vertex Locator and Lifetime Measurements of Two-Body Hadronic Final States", PhD thesis, University of Glasgow, 2009.

[86] Alexander, Michael Thomas, "Constraints on Mixing and CP-Violation in the Neutral Charmed Meson System at LHCb", PhD thesis, University of Glasgow, 2012.

[87] LHCb Collaboration, R. Aaij *et al*, "Measurement of J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV", *Eur.Phys.J.* **C71** (2011) 1645, arXiv:1103.0423 [hep-ex].

[88] Cranmer, Kyle S, "Kernel estimation in high-energy physics", *Comput.Phys.Commun.* **136** (2001) 198–207, arXiv:hep-ex/0011057 [hep-ex].

[89] R. Aaij *et al*, "Measurement of the $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$, $\mathrm{B}^0_d \to \mathrm{K}^+\pi^-$ and $\mathrm{B}^0_s \to \pi^+\mathrm{K}^-$ Lifetimes", LHCb-ANA-2012-057.

[90] "The MINT Package", Source code available from http://www.phy.bris.ac.uk/people/rademacker_j/index.html.

[91] James, F. and Roos, M, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput.Phys.Commun.* **10** (1975) 343–367.

[92] H. Albrecht et al, "Search for rare B decays", *Physics Letters B* **353** no. 4, (1995) 554–562, {http://www.sciencedirect.com/science/article/pii/037026939500626V}.

[93] CLEO Collaboration, R. Godang *et al*, "Observation of exclusive two-body B decays to kaons and pions", *Phys. Rev. Lett.* **80** (Apr, 1998) 3456–3460, http://link.aps.org/doi/10.1103/PhysRevLett.80.3456.

[94] CDF Collaboration, A. Abulencia *et al*, "Observation of B^0 $(s) \to K^+K^-$ and Measurements of Branching Fractions of Charmless Two-body Decays of B^0 and B_s^0 Mesons in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1.96 TeV", *Phys.Rev.Lett.* **97** (2006) 211802, arXiv:hep-ex/0607021 [hep-ex].

[95] Belle Collaboration, Louvot, Remi, "B0(S) Decays at Belle", *PoS* **EPS-HEP2009** (2009) 170, arXiv:hep-ex/0909.2160 [hep-ex].

[96] R. Aaij *et al*, "Measurement of B-hadron branching fractions for two-body decays into charmless charged hadrons", *J. High Energy Phys.* **volume 10** (Jun, 2012) 037. 21 p arXiv:1206.2794. LHCB-PAPER-2012-002. CERN-PH-EP-2012-161.

[97] LHCb Collaboration, R. Aaij *et al*, "First evidence of direct CP violation in charmless two-body decays of B_s^0 mesons", *Phys.Rev.Lett.* **108** (2012) 201601, arXiv:1202.6251 [hep-ex].

[98] LHCb Collaboration, "Measurement of time-dependent CP violation in charmless two-body B decays", LHCb-ANA-2012-020.

[99] R. Aaij *et al*, "Measurement of the effective $B_s^0 \to K^+K^-$ lifetime", *Phys. Lett. B* **716** no. arXiv:1207.5993. LHCB-PAPER-2012-013. CERN-PH-EP-2012-172, (Jul, 2012) 393–400. 16 p.

[100] "NeuroBayes-HowTo", http://neurobayes.phi-t.de/.

[101] "TMVAUsersGuide", http://tmva.sourceforge.net/.

[102] Skwarnicki, Tomasz, "A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances", PhD thesis, Insitute of Nuclear Physics, Cracow, 1986.

[103] "Exponentially Modified Gaussian Distribution", http://en.wikipedia.org/wiki/Exponentially_modified_Gaussian_distribution.

[104] "ROOT : A Data Analysis Framework", http://root.cern.ch/drupal/.

[105] LHCb Collaboration, R. Aaij *et al*, "Measurement of mixing and CP violation parameters in two-body charm decays", *JHEP* **1204** (2012) 129, arXiv:1112.4698 [hep-ex].

[106] The LHCb Collaboration, "Search for the Rare Two-Body Charmless Baryonic B-Decays $\mathrm{B}^0_d \to \mathrm{p}\overline{\mathrm{p}}$ and $\mathrm{B}^0_s \to \mathrm{p}\overline{\mathrm{p}}$", LHCb-ANA-2012-064.

[107] R. Aaij *et al*, "First evidence for the two-body charmless baryonic decay $B^0 \to p\bar{p}$", Tech. Rep. arXiv:1308.0961. LHCB-PAPER-2013-038. CERN-PH-EP-2013-138, CERN, Geneva, Aug, 2013.

[108] The LHCb Collaboration, "Observation of the decay $\mathrm{B}^+_c \to \mathrm{B}^0_s\pi^+$", LHCb-PAPER-2013-044.

[109] LHCb Collaboration, "Measurement of the CP Violation Parameter A_Γ in Two-Body Charm Decays", LHCb-ANA-2011-066, LHCb-CONF-2011-046.

[110] LHCb collaboration, R. Aaij *et al*, "First measurement of time-dependent CP violation in $\mathrm{B}^0_s \to \mathrm{K}^+\mathrm{K}^-$ decays", `arXiv:1308.1428 [hep-ex]`.

[111] LHCb collaboration, R. Aaij *et al*, "First observation of CP violation in the decays of B^0_s mesons", *Phys.Rev.Lett.* **110** (2013) 221601, `arXiv:1304.6173 [hep-ex]`.

[112] LHCb Collaboration, R. Aaij *et al*, "Measurement of the $\overline{B}^0_s$ Effective Lifetime in the $J/\psi f_0(980)$ Final State", *Phys. Rev. Lett.* **109** (Oct, 2012) 152002, `http://link.aps.org/doi/10.1103/PhysRevLett.109.152002`.

[113] LHCb collaboration, R. Aaij *et al*, "Measurement of CP violation and the B^0_s meson decay width difference with $B^0_s \to J/\psi K^+K^-$ and $B^0_s \to J/\psi\pi^+\pi^-$ decays", *Phys. Rev. D 87,* **112010** (2013) , `arXiv:1304.2600 [hep-ex]`.

[114] LHCb collaboration, "Branching fraction measurements of $B^0_{d,s}$ decays to $K^0_{\mathrm{S}}h^\pm h'^\mp$ final states, including first observation of $B^0_s \to K^0_{\mathrm{S}}K^\pm\pi^\mp$", LHCb-CONF-2012-023.

[115] LHCb Collaboration, R. Aaij *et al*, "Measurement of the ratio of branching fractions $\mathcal{B}(\mathrm{B}^0_d \to \mathrm{K}^{*0}\gamma)/\mathcal{B}(\mathrm{B}^0_d \to \phi\gamma)$", `arXiv:1202.6267 [hep-ex]`.