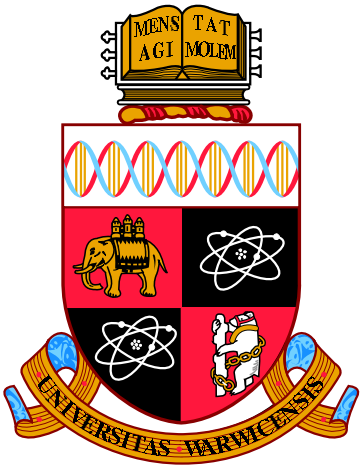


Search for a wrong-flavour contribution to $B_s^0 \to D_s\pi$ decays and study of CP violation in $B_s^0 \to D_sK$ decays

by

Matthew P. Williams

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

January 2014

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

Firstly I would like to thank my supervisors, Professor Tim Gershon, Dr. Sajan Easo and Dr. Michal Kreps. Tim for guiding me through my whole Ph.D. as well as being a tutor and mentor for the four years of my undergraduate studies; Sajan for supervising me during my time at CERN, answering daily questions and helping me with all the work I performed while there; Michal for being my direct contact at Warwick for the last year and a half of my Ph.D. and helping my in my day-to-day work and to all of them for the guidance and many readings they have given of this document over the last few months.

I would like to thank all those who shared my time both at Warwick and at CERN. First Mark Whitehead and Tom Latham for always being willing to answer the many questions I would have and for offering their experience. David Dossett for being a great office-mate through the first year, for making my time at CERN all the more fun and for being my comrade-in-arms through the thesis writing process. I'd also like to thank all the EPP group at Warwick, particularly Matt, Tomas, Nicky, Callum, Dan, Rafael and Charlotte for the fun times, lunches and tea breaks. All those at Warwick without whom my time would have been so much more dull are too many to name but special thanks must go to Steve, Pete and Simon for all the good times we had, to those in the Astronomy group for giving me a second home and to everyone else who helped my in ways large and small.

The whole LHCb collaboration deserves my thanks for without them I would have no thesis. Vava, Eduardo, Moritz and the rest of the B2OC working group for the assistance they have provided my in so many ways and to the Flavour Tagging working group for their help and guidance.

I must thank my family for being there for me my whole life. My parents Denise and Paul for their love, support and for making me the person I am today. My brothers David and Jack and my sister Alice for being there with me from childhood to where we are today.

Finally, all the thanks in the world go to my loving girlfriend Rachel who was there with me through the good and bad, always keeping me on track and always making me be a better person. Without her I would not have got this far.

Declarations

The work presented in this thesis is all of my own work, unless it is specifically referenced to the contrary. The analysis shown in Part II was done as part of a collaboration where my main contribution was to the flavour tagging, though I took part in the whole analysis. This thesis has not been submitted, in any form, to this or any other university for another qualication.

Matt Williams

Abstract

This thesis presents a world-first measurement of the *CP* parameters of $B_s^0 \to D_s\pi$ with the aim of testing the flavour-specific nature of the decay. The measurement is made using $\int \mathcal{L} = 1.0$ fb^{-1} of pp collisions recorded at a centre-of-mass energy of $\sqrt{s} = 7$ TeV collected during 2011.

The measured values of the *CP* parameters are

$$\begin{aligned}
\overline{S} &= \quad 0.197 \pm 0.150 \pm 0.025, \\
\overline{D} &= -0.888 \pm 0.098 \pm 0.541, \\
\Delta S &= \quad 0.066 \pm 0.083 \pm 0.004, \\
\Delta D &= -0.062 \pm 0.050 \pm 0.169,
\end{aligned}$$

where the first uncertainties are statistical and the second are systematic. These results are consistent with the Standard Model prediction.

A measurement of the *CP* parameters of $B_s^0 \to D_sK$ is also presented, producing

$$\begin{aligned}
C &= \quad 1.01 \pm 0.50 \pm 0.23, \\
S_f &= -1.25 \pm 0.56 \pm 0.24, \\
S_{\bar{f}} &= \quad 0.08 \pm 0.68 \pm 0.28, \\
D_f &= -1.33 \pm 0.60 \pm 0.26, \\
D_{\bar{f}} &= -0.81 \pm 0.56 \pm 0.26,
\end{aligned}$$

where the first uncertainties are statistical and the second are systematic.

Part I

Introduction

1

Preface

The LHCb detector at the LHC is dedicated to two primary aims. The study of *CP* violation in the quark sector and searches for new physics, beyond the Standard Model. The two aims are not mutually exclusive as precision measurements of *CP* parameters could shed light on new physics processes. To date the Standard Model of particle physics has withstood repeated attempts to measure inconsistencies, from the earliest days of experimental particle physics up to the latest results from the experiments at the LHC. Most recently, the precise measurement of $B_s^0 \to \mu^+\mu^-$, which so far fantastically agrees with Standard Model predictions despite hopes that it would yield evidence of physics beyond the Standard Model.

This thesis presents an analysis testing one of the assumptions underlying many other precision analyses at LHCb and elsewhere – that there are no wrong-flavour contributions to the decay of $B_s^0 \to D_s\pi$. By searching for wrong-flavour contributions to the $B_s^0 \to D_s\pi$ decay I hope to start work on constraining our level of certainty of this assumption.

The thesis is composed of three main parts. The first provides a background to the analysis work carried out by explaining in Chapter 2 the theoretical underpinning of the models used when analysing LHCb data. This is followed in Chapter 3 by a detailed description of the LHCb detector and the hardware and software analysis that is performed on the data collected.

Since the analysis of $B_s^0 \to D_s\pi$ was based on a previous study of $B_s^0 \to D_sK$, Part II describes the analysis of that decay. The analysis was written up in full [1] and was submitted as a conference paper [2]. Chapter 4 introduces the analysis in the context of the $B_s^0 \to D_s\pi$ analysis. Chapters 5 to 8 then describe the analysis work itself including the selection of events, the fit to the B_s^0 candidate mass distribution

and to the decay-time distribution.

Part III contains a description of the measurement of the *CP* parameters of $B_s^0 \to D_s\pi$ decays using LHCb data in order to measure the wrong-flavour contribution to the decay. Chapter 10 describes the process of selecting the $B_s^0 \to D_s\pi$ events from the LHCb data. Chapter 6 describes the algorithms and techniques used to tag the initial-state flavour of the B_s^0 mesons which is necessary for this type of time-dependent analysis. This is followed in Chapter 11 by a description of the method by which the signal and background yields were extracted. The fit to the decay-time distribution of $B_s^0 \to D_s\pi$ candidates is presented in Chapter 12 with a full study of the possible sources of systematic uncertainty given in Chapter 13.

Chapter 14 gives a summary of the results of the $B_s^0 \to D_s\pi$ analysis, comparing them to the results of the $B_s^0 \to D_sK$ analysis and drawing conclusions.

Information on the combination of uncertainties given large correlations is given in Appendix A.

2

Theory

2.1 The Standard Model

The Standard Model is a theory which describes our understanding of the fundamental particles and their interactions. It has proven to be an excellent model to describe and predict phenomena involving three of the four fundamental forces: electromagnetism and the strong and weak nuclear forces. The force of gravity is not attempted to be explained by the Standard Model and is ignored by it entirely.

The Standard Model describes two classes of fundamental particles: *fermions* and *bosons*. Conventionally the fermions are grouped into three generations, ordered largely by their masses, with four particles in each generation. Each generation can be seen as a copy of the particles in the previous generation but with a larger mass but the same electric charge. Furthermore, they can be categorised into two classes, the *leptons* and the *quarks*. The key properties of the twelve fermions are given in Table 2.1. Further to the twelve fermions listed, each also has a corresponding anti-particle with exactly the same mass but with inverted quantum numbers.

In addition to the fermions, the Standard Model also describes a set of *gauge bosons* which act as the force mediators in interaction processes. They are summarised in Table 2.2.

All gauge bosons and fermions have been observed in experiments with the last quark to be observed being the top quark which was measured by the CDF and D0 experiments at on the TEVATRON at Fermilab in 1995 [4]. The last of the leptons, the tau neutrino, was observed by the DONUT experiment (again at Fermilab) in 2000 [5].

There are a number of theories which make up the Standard Model, each

Generation	Fermion name	Mass (MeV/c^2)	Electric charge
I	Electron	0.510998910(13)	-1
	Electron neutrino	Undefined	0
	Up quark	$2.5^{+0.6}_{-0.8}$	$2/3$
	Down quark	$5.0^{+0.7}_{-0.9}$	$-1/3$
II	Muon	105.658367(4)	-1
	Muon neutrino	Undefined	0
	Charm quark	1290^{+50}_{-110}	$2/3$
	Strange quark	100^{+30}_{-20}	$-1/3$
III	Tau	1776^{+600}_{-900}	-1
	Tau neutrino	Undefined	0
	Top quark	172900^{+600}_{-900}	$2/3$
	Bottom quark	4190^{+180}_{-60}	$-1/3$

Table 2.1: Summary of the fermions, organised by generation. Masses are taken from Ref [3].

Boson name	Associated force	Mass (GeV/c^2)
Photon	Electromagnetic	0
Gluon	Strong nuclear	0
$W^{\pm}$	Weak nuclear	80.399 ± 0.023
Z^0	Weak nuclear	91.1876 ± 0.0021

Table 2.2: Summary of the Standard Model gauge bosons. Masses are taken from Ref [3].

describing a different type of process. Charged particle interactions are described by Quantum Electrodynamics (QED) for which the Nobel Prize was awarded in 1965. Later work was able to combine the effects of QED with a theory of the weak interaction into a unified theory known as Electroweak theory (EW). This theory was able to describe nuclear beta decay and other processes involving quarks. As quark theory evolved, it became clear that quarks had an extra quantum number, named *colour*, which affected how quarks interacted with each other via the strong force. Quantum Chromodynamics (QCD) was developed to describe these strong interactions.

The strong force is mediated by the gluon and quarks are the only fermions to feel the effects of it. It can only affect particles with colour charge which comes in three types: *red*, *green* and *blue* (along with three anti-colours *anti-red*, *anti-green* and *anti-blue*). A feature of the strong force which differentiates it from the others is a concept known as *confinement*. This arises from the fact that the strength of the strong force increases as the distance between two coloured particles increases meaning that only colourless composite particles can exist freely. These colourless particles are called *hadrons* and fall into two classes. Those containing a quark and and an anti-quark are called *mesons* and those containing three quarks (or three anti-quarks) are called *baryons*.

These six quarks can be combined in a number of different ways to form hadrons and there is a host of observed particles formed this way. For example the particles which make up all matter are protons and neutrons which are baryons formed of, respectively, *uud* and *udd* quarks.

Heavy mesons follow a relatively consistent naming scheme. In the case where the lighter quark is one of an up or down quark, mesons are generally named after the heaviest quark inside them so mesons containing a strange, charm or bottom quark are named K, D and B respectively. Where the lighter quark is something heaver than a u or a d, the meson's name is augmented so that a meson formed of a bottom and a strange quark would be named a strange B meson (B_s).

2.2 *CP* and its violation

2.2.1 *C* and *P* symmetries and their violation

CP is the combination of the *charge conjugation* and *parity* transformations. The charge conjugation transformation inverts the sign of the electric charge and all the internal quantum numbers of a particle — turning a particle into an antiparticle (or vice versa) and the parity transformation inverts all spatial directions — effectively

looking at the particle in a mirror. If all processes were invariant under these transformations individually, then there would be no difference in behaviour between particles and antiparticles. To say a symmetry is violated is to say that physical processes are not invariant under their associated transformations. Invariance under charge conjugation is termed C *symmetry* and invariance under the parity transformation is termed P *symmetry*.

For a while, experiments seemed to confirm that nature obeyed P symmetry but in 1956, after a careful review of the experiments to date by T. D. Lee and C. N. Yang, it was found that none of those experiments had proven its invariance in weak interactions [6]. Later that year experimental evidence for parity violation was found by measuring the direction in which an electron is emitted from decaying Cobalt-60 atoms [7]. If parity were conserved, it would be expected that when a Cobalt atom undergoes radioactive decay, the emitted electron would have no preference for travelling parallel or anti-parallel to the spin vector of the atom. The experiment proceeded by placing a large number of Cobalt-60 atoms in a strong magnetic field and lowering their temperature in order to align their spin vectors and then measuring their decay. Instead of seeing equal numbers of electrons being emitted in each direction with respect to the magnetic field, almost all the electrons were detected travelling in one direction. Since there wasn't just a small imbalance but rather a great preference for emission in one direction the violation is said to be *maximal*.

C symmetry is conserved in the strong and electromagnetic interactions but is not in the weak interaction [8]. This can be shown by looking at the effect that charge conjugation has on neutrinos. Consider a left-handed neutrino (a neutrino with its spin being anti-parallel to its momentum) which is able to interact with the weak force. Under charge conjugation this neutrino becomes a left-handed antineutrino which is unable to feel the weak interaction and in fact has never been observed in nature.

Even though the parity symmetry was found to be broken, the overall symmetry of the system could be kept if another symmetry were found which, when combined with parity, resulted in a non-violated symmetry. In 1957 Landau suggested [9] that the combination of charge conjugation and parity (CP) was in fact the conservation law for all (including weak) processes. In the case of the left-handed neutrino, charge conjugation changes it into a left-handed antineutrino. When the parity transformation is also applied the particle becomes a right-handed antineutrino. This is consistent with observations of neutrinos in nature which interact via the weak interaction where only left-handed neutrinos and right-handed antineutri-

nos are seen.

Indirect *CP* violation

However, it was shown in 1964 by Cronin and Fitch [10] that not only does the weak interaction violate the two symmetries individually but also their combination, *CP symmetry*. The symmetry was not found to be violated maximally (as was the case for parity) but instead only very slightly.

The method used by Cronin and Fitch to measure this violation involved the ability of neutral K mesons (the $K^0(d\bar{s})$ and $\overline{K}^0(\bar{d}s)$) to change into their antiparticles. This is fairly unique to this kind of particle since it relies on the fact that there is no conserved quantum number that distinguishes between the K^0 and the $\overline{K}^0$. Because of this, the propagating state is actually a combination of these two states, this is known as the K^0–$\overline{K}^0$ *mixing phenomenon*. The neutral B mesons also exhibit this behaviour and is the basis for some of the measurements to be performed at LHCb.

There are two observed kaon states known as the K^0_{L} and K^0_{S} (with relatively long ($\tau = 5 \times 10^{-6}\,\mathrm{s}$) and short ($\tau = 1 \times 10^{-10}\,\mathrm{s}$) lifetimes respectively, hence the index) which predominantly decay to three or two pion systems. The K^0_{S} is dominated by $K^0_{\mathrm{S}} \to \pi^+\pi^-, \pi^0\pi^0$ (both of which resultant systems have a *CP* eigenvalue of $+1$) and the K^0_{L} is dominated by $K^0_{\mathrm{L}} \to \pi^+\pi^-\pi^0, \pi^0\pi^0\pi^0$ (both with $CP = -1$).

If it is assumed that *CP* is an invariant operation then it can be shown that there are two *CP* eigenstates of the $K^0/\overline{K}^0$, one with $CP = 1$ (labelled K^0_1) and one with $CP = -1$ (labelled K^0_2). Given the matching *CP* eigenvalues, this would immediately suggest that $K^0_1 = K^0_{\mathrm{S}}$ and $K^0_2 = K^0_{\mathrm{L}}$.

However, Cronin and Fitch measured decays of the K^0_{L} into a $\pi\pi$ system ($\sim 0.1\%$ of the time), a change of *CP* from -1 to $+1$, demonstrating that the observed neutral kaon states must be combinations of the K^0_1 and K^0_2 as given by,

$$\begin{aligned} K^0_{\mathrm{L}} &= \left(1 + |\varepsilon|^2\right)^{-1/2} \left[\varepsilon K^0_1 + K^0_2\right], \\ K^0_{\mathrm{S}} &= \left(1 + |\varepsilon|^2\right)^{-1/2} \left[K^0_1 - \varepsilon K^0_2\right], \end{aligned} \tag{2.1}$$

where ε is the mixing parameter ($\sim 10^{-3}$) and the term in front is for normalisation. They supposed that it is the K^0_1 component which is decaying in the case they observed. It is the existence of the *CP* forbidden K^0_1 component in the K^0_{L} wave function which provides evidence for weak processes violating *CP* symmetry indirectly and hence is called *indirect CP violation*.

Direct *CP* violation

For a long time, evidence for indirect *CP* violation was all that had been seen but in the 1990's, experiments from Fermilab and the NA48 collaboration at CERN demonstrated *CP* violation in processes which did not include particle-antiparticle oscillations but rather an asymmetry which occurs in the decay of the kaons themselves [11]. That is, rather than it being the *CP*-forbidden component decaying via a *CP*-allowed process, it is instead the *CP*-permitted component decaying via a *CP*-violating process.

At this point, there was no experimental evidence to suggest that *CP* violation was present anywhere except the kaon sector. "*B factory*" experiments such as BABAR and BELLE were designed to look for *CP* violation in B mesons. First, in 2001, BABAR and BELLE measured indirect *CP* violation in B meson systems [12, 13], confirming that *CP* violation was not confined to neutral kaons, then in 2004 evidence for direct *CP* violation was seen in the decay of $B^0 \to K^+\pi^-$ [14, 15].

The 2004 BABAR result (and similarly the 1999 NA48 kaon experiment) worked not by inferring the violation of *CP* symmetry through the requirement of a kaon-mixing system but rather by directly measuring the difference in yields of different decays and so is referred to as *direct CP violation*. The BABAR experiment worked by creating $B^0\bar{B}^0$ pairs and measuring the number of decays of $B^0 \to K^+\pi^-$ (called $n_{K^+\pi^-}$) and $\bar{B}^0 \to K^-\pi^+$ (called $n_{K^-\pi^+}$). If *CP* invariance was assumed, the two yields should be identical (since particles and their antiparticles should behave identically) and so any difference indicates *CP* violation. The *CP* asymmetry is quantified as

$$A_{K\pi} = \frac{n_{K^-\pi^+} - n_{K^+\pi^-}}{n_{K^-\pi^+} + n_{K^+\pi^-}}. \tag{2.2}$$

The value of $A_{K\pi}$ measured by the BABAR collaboration was $-0.133 \pm 0.030\,(\text{stat.}) \pm 0.009\,(\text{syst.})$ demonstrating a clear non-zero value (the probability of observing such asymmetry without the effect of direct *CP* violation is 1.3×10^{-5}, or 4.2 standard deviations).

The methods used by LHCb to measure the level of *CP* violation builds upon those used by BABAR, BELLE and CDF [16] and will be described later in section 2.3.3.

CPT symmetry

Much like Landau suggesting that by combining C symmetry with P symmetry, *CP* was the symmetry for weak interactions, it has been shown that including the time reversal transformation (to create *CPT symmetry*) makes all interactions invariant.

CPT requires (among other things) that the *mass* of a particle must be the same as its antiparticle. Measurements of the masses of neutral K mesons have shown that $|m_{K^0} - m_{\bar{K^0}}| < 5.1 \times 10^{-19}\,\mathrm{GeV}/c^2$ at a 90% confidence level [3].

2.2.2 Why is *CP* violation important?

The violation of *CP* symmetry is an important area of research for many reasons. The most prominent of which is the problem of the large imbalance of matter and antimatter in the universe — that is why do we have predominantly baryons rather that antibaryons. This apparent 'production' of matter is know as *baryogenesis*. The problem is that in the early universe it is assumed that matter and antimatter were created in equal amounts by a process like $boson \rightarrow matter + antimatter$ and destroyed in equal amounts as $matter + antimatter \rightarrow boson$ — a perfectly symmetric process. Clearly there is an asymmetry and so there must be something outside this simplistic framework responsible for it. Either there had always been an imbalance of matter and antimatter (which is thought not to be possible) or there is some process by which matter and antimatter can act differently.

In 1967, Andrei Sakharov came up with three necessary conditions for the universe to have been able to create matter and antimatter at different rates [17]. One of these so-called *Sakharov Conditions* is that charge conjugation symmetry and *CP* symmetry must be violated. *CP* violation in this context allows matter and antimatter to act slightly (or completely) differently, shifting the balance of matter and allowing a significant residual amount of matter to remain after maximum annihilation. Within this framework, the amount of residual matter remaining should be equal to the amount of matter in our universe today.

While *CP* violation has been observed experimentally in the weak interaction, the *CP* violation is not large enough by a long way to account for the large matter asymmetry using models based on the number of photons in the universe [18]. The level of *CP* violation required in the early universe to observe the asymmetry in the universe today is several orders of magnitude smaller than that in the kaon system. However, since *CP* violation is dependent on there being a difference in mass between the quarks, the early universe's high energy density would make this effect appear negligible. Theories outside the current Standard Model such as the *Minimal Supersymmetric Standard Model* [19] provide a promising framework for weak baryogenesis.

Another method to explain baryogenesis, while remaining largely within the Standard Model, is through the *CP* violation in neutrino mixing [20]. The mixing itself creates an asymmetry between different species of leptons (*leptogenesis*) which

in turn is responsible for baryogenesis.

2.3 Unitary triangle

2.3.1 The CKM matrix

In 1963 Nicola Cabibbo introduced [21] the Cabibbo angle, θ_C, to parametrise the probability of a down or a strange quark decaying into a up quark via the weak interaction. What is really represented was the extent to which down and strange mass eigenstates (d and s) were present in the weak interaction eigenstate (d') — that is the level of mixing of the mass eigenstates. When the charm quark was discovered it turned out that the mixing of down and strange mass states in a charm weak eigenstate could also be parametrised by θ_C. This can be represented in matrix notation as

$$\begin{bmatrix} |d'\rangle \\ |s'\rangle \end{bmatrix} = \begin{bmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{bmatrix} \begin{bmatrix} |d\rangle \\ |s\rangle \end{bmatrix}, \tag{2.3}$$

or, more suggestively as

$$\begin{bmatrix} |d'\rangle \\ |s'\rangle \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{bmatrix} \begin{bmatrix} |d\rangle \\ |s\rangle \end{bmatrix}. \tag{2.4}$$

In order to incorporate a proposed third generation of quarks, Kobayashi and Maskawa expanded the matrix to be 3×3 with quark mixing as

$$\begin{bmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{bmatrix}. \tag{2.5}$$

The elements of this matrix (known as the *CKM matrix*) represent the probability of a transition of a quark of one type i to another type j with Prob $= |V_{ij}|^2$. The elements are not entirely real but rather have a complex phase which represents the level of *CP* violation present. The currently published magnitudes of the elements are given by [3, p. 145–152],

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.999133^{+0.000044}_{-0.000043} \end{bmatrix}. \tag{2.6}$$

It can be seen that the matrix is almost, but not quite, diagonal telling us that quarks are most likely to decay within their generation. However, it *is* unitary such that

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.7}$$

This relation is simply saying that the sum of three complex number must add up to zero and so, if drawn on the complex plane as vectors, would form a triangle. The internal angles of the triangle would parametrise the mixing and the area would be a measure of *CP* violation. This triangle is called a unitarity triangle (as it represents the unitarity of the CKM matrix).

2.3.2 Unitarity triangles

In addition to the relation in Equation 2.7 there are other 5 other relations leading to 5 other triangles, one labelled for each of the quark flavours — three from the columns of the CKM matrix (d, s and b) and three from the rows (u, c and t). These 6 triangles are not completely separate though — each triangle has each of its sides present in one of the other 5 triangles. This means that instead of there being $6\,(\text{triangles}) \times 3\,(\text{angles per triangle}) = 18$ unique angles there are only 9.

The three angles that have historically received the most attention are α (also known as ϕ_2), $\beta(\phi_1)$ and γ (ϕ_3) and are defined as

$$\begin{aligned} \alpha &= \arg\left(-V_{td}V_{tb}^*/V_{ud}V_{ub}^*\right), \\ \beta &= \arg\left(-V_{cd}V_{cb}^*/V_{td}V_{tb}^*\right), \\ \gamma &= \arg\left(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*\right). \end{aligned} \tag{2.8}$$

These are the angles which compose the standard unitarity triangle (Figure 2.1).

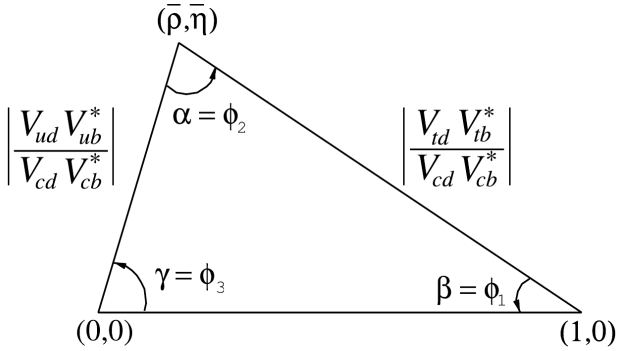


Figure 2.1: The unitarity triangle, created from Equation 2.7. Reproduced from Ref [3].

The angles and lengths of the sides can all be measured directly or indirectly by various processes. For example angle β is measured through the interference between decays with and without mixing.

One of the key goals of LHCb is to improve the measurement of the angle γ. In 2009 the collaboration published a document [16] reviewing the measurements at other experiments such as BABAR, BELLE and CDF as well as detailing the proposed methods of measuring the angle γ at LHCb.

Published values of γ to date by groups such as CKMfitter [22] by averaging results from many different experiments have yielded values of $(66 \pm 12)^\circ$. Example plots showing some of the constraints placed on the unitarity triangle are shown in Figure 2.2.

2.3.3 Methods to measure angle γ

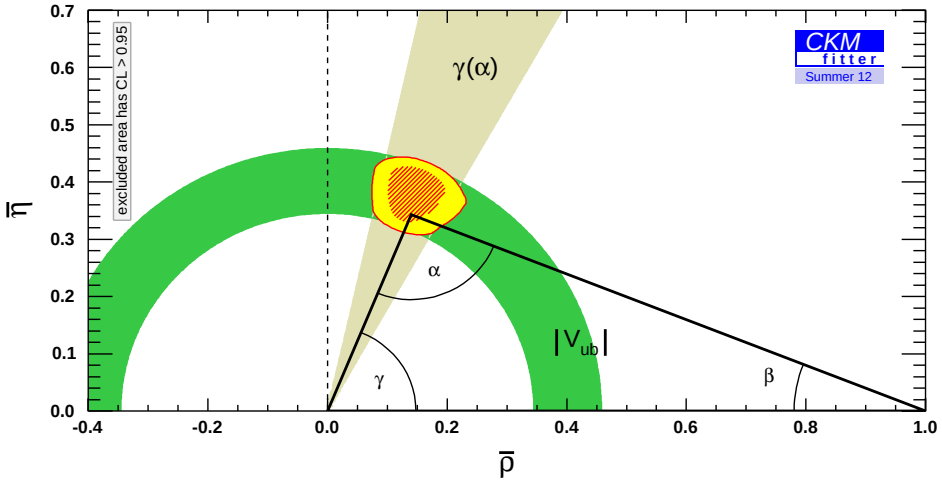
There have been many methods used to measure γ (and other angles) to date but they all fall in to one of two categories: *loop-level* and *tree-level*. With respect to measuring γ from tree-level processes there are two main techniques used: time-dependent measurements of *CP*-violation in $B^0 \to D^{(*)}\pi$ and $B_s^0 \to D_s K$ decays and measurement of direct *CP*-violation in families of decays such as $B \to DK^{(*)}$.

Time-dependent

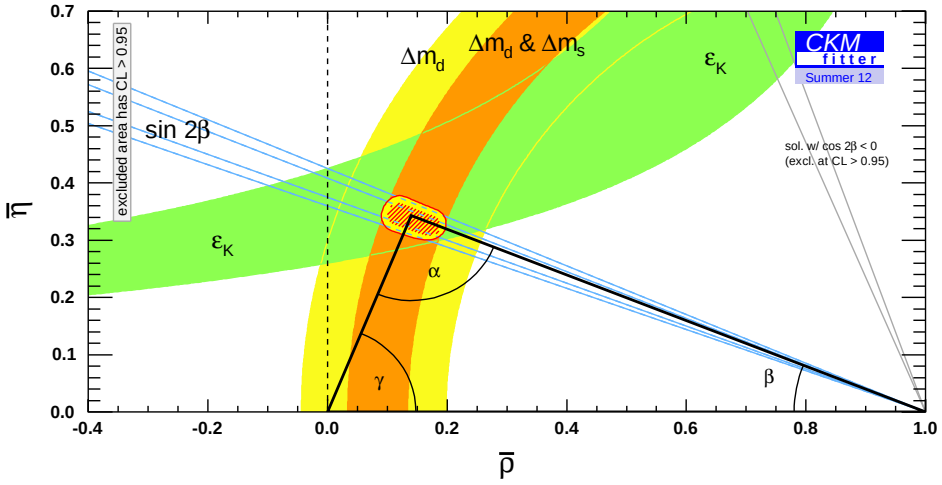
In the decays show in Figure 2.3, the B mesons start with opposite flavour states (B_s^0 and $\overline{B}_s^0$) but have identical final states ($K^- D_s^+$). Measurements of their time-dependent *CP* asymmetries allow $\gamma - 2\beta_s$ to be measured (where β_s is the B_s^0 mixing phase). The sensitivity arises from the interference between the immediate decay of the B meson initial state to $D_s^+ K^-$ and the decay to that same final state after B meson mixing. The value of γ can be extracted from $\gamma - 2\beta_s$ since β_s will be well known due to other studies of $B_s \to J/\psi\phi$ [23].

Time-integrated interference measurement

The time integrated measurements work by measuring many decays of a particle into two different intermediate states which only differ by the type of quark transition ($b \to c$ and $b \to u$) and then measuring the interference of those particles decaying to identical final states. An example of this is the $B^- \to D^0/\overline{D}^0 K^-$ decays shown in Figure 2.4 where the $D^0/\overline{D}^0$ would decay to a final state accessible to both. The decay amplitude of diagram (a) depends on the CKM matrix element V_{cb} while the amplitude of diagram (b) depends on the V_{ub} element and it is therefore suppressed



(a) Constraints on γ from tree-level processes.



(b) Constraints from loop-level processes.

Figure 2.2: Constraints on the unitarity triangle as published by the CKMfitter collaboration [22] from Summer 2012.

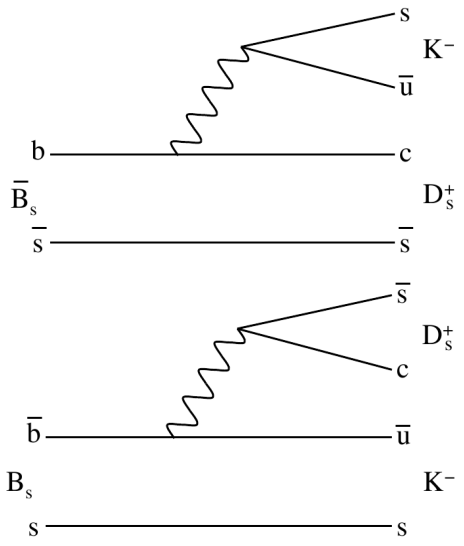


Figure 2.3: Diagrams for (a) $\bar{B}_s \to K^- D_s^+$ and (b) $B_s \to K^- D_s^+$ (reproduced from Ref [16]).

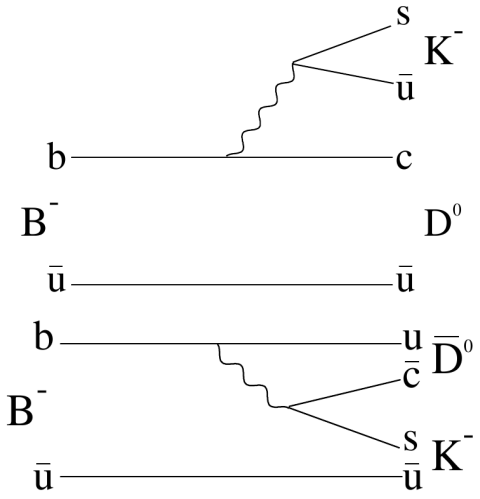


Figure 2.4: Diagrams for (a) $B^- \to D^0 K^-$ and (b) $B^- \to \overline{D}^0 K^-$ (reproduced from Ref [16]).

(see Equation 2.6). The weak phase difference between V_{cb} and V_{ub} is $-\gamma$ and so interference between identical final states of the $D^0/\overline{D}^0$ decay gives sensitivity to γ.

There are other factors that feed into the measurement such as the ratio of the absolute value of the suppressed decay amplitude to the favoured decay, r_B, the strong phase difference, δ_B as well as parameters of the specific decay the D meson undergoes. These must all be measured by other methods or from previous measurements.

The largest correction to the measurement is due to $D^0\overline{D}^0$ mixing but this is observed to have a small effect on γ [24].

2.4 Time evolution of a B_s^0 meson

In the case of $B_s^0 \to D_s\pi$ decays, the observables are the four decay rates of the two initial states, B_s^0 and $\overline{B}_s^0$ decaying to the two final states, $D_s^-\pi^+$ and $D_s^+\pi^-$. which are dependent on the level of *CP* violation in the system. The discussion that follows is valid for any neutral B meson system with the mesons denoted as B and $\overline{B}$ and final states f and $\overline{f}$. For the purposes of the analysis detailed in this thesis, the mesons are B_s^0 and $\overline{B}_s^0$ and the final states are $D_s^-\pi^+$ and $D_s^+\pi^-$.

2.4.1 *CP* violation

As with the kaon system, the B and $\overline{B}$ states also mix. Similarly to Equation 2.1, we can define two *CP* eigenstates in terms of the two observed flavour eigenstates as

$$\begin{aligned} |B_L\rangle &= p|B\rangle + q|\overline{B}\rangle, \\ |B_H\rangle &= p|B\rangle - q|\overline{B}\rangle, \end{aligned} \tag{2.9}$$

with the constraint that $|p|^2 + |q|^2 = 1$. At this point it it also useful to define a shorthand for the various decay amplitudes

$$\begin{aligned} A_f = A(B \to f) &= \langle f|\mathcal{H}|B\rangle, \\ \overline{A}_f = A(\overline{B} \to f) &= \langle f|\mathcal{H}|\overline{B}\rangle, \end{aligned} \tag{2.10}$$

and based on that it is possible to define a key *CP* quantity,

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{2.11}$$

The quantity λ_f contains the important feature of the interference of the two meson states to final state f: the relative phase between q and p (based on the

mixing) and the relative amplitudes of the two decays. If $\lambda_f = 0$ then the decay is flavour-specific.

2.4.2 Time-dependent meson states

The decay amplitude is a function of decay time, t, with initial conditions defined as $|B(t=0)\rangle = |B\rangle$ and $|\overline{B}(t=0)\rangle = |\overline{B}\rangle$. As the mesons propagate, for $t > 0$ the states are defined as superpositions of $|B\rangle$ and $|\overline{B}\rangle$ and are described by a Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix}|B(t)\rangle\\|\overline{B}(t)\rangle\end{pmatrix} = \Sigma\begin{pmatrix}|B(t)\rangle\\|\overline{B}(t)\rangle\end{pmatrix}, \tag{2.12}$$

where Σ is a 2×2 matrix which can be written as the sum two matrices

$$\Sigma = M - i\frac{\Gamma}{2}, \tag{2.13}$$

where $M = M^\dagger$ is the mass matrix and $\Gamma = \Gamma^\dagger$ is the decay matrix. The mass eigenstates $|B_L\rangle$ and $|B_H\rangle$ are the eigenvectors of Σ.

The time evolution of the mass eigenstates is governed by the eigenvalues $B_H - i\Gamma_H/2$ and $B_L - i\Gamma_L/2$ so that

$$|B_{H,L}(t)\rangle = e^{-(iB_{H,L}-\Gamma_{H,L}/2)t}|B_{H,L}(t=0)\rangle, \tag{2.14}$$

where, $|B_{H,L}(t=0)\rangle$, as with the flavour eigenstates, denotes the mass eigenstate at $t = 0$. By inverting Equation 2.9, the flavour eigenstates are found in terms of the mass eigenstates to be

$$\begin{aligned}|B(t)\rangle &= \frac{1}{2p}\left(|B_H(t)\rangle + |B_L(t)\rangle\right),\\ |\overline{B}(t)\rangle &= \frac{1}{2q}\left(|B_H(t)\rangle - |B_L(t)\rangle\right),\end{aligned} \tag{2.15}$$

inserting the time evolution as defined in Equation 2.14 we obtain

$$\begin{aligned}|B(t)\rangle &= \frac{1}{2p}\left[e^{-iB_Lt-\Gamma_Lt/2}|B_L(t=0)\rangle + e^{-iB_Ht-\Gamma_Ht/2}|B_H(t=0)\rangle\right],\\ |\overline{B}(t)\rangle &= \frac{1}{2q}\left[e^{-iB_Lt-\Gamma_Lt/2}|B_L(t=0)\rangle - e^{-iB_Ht-\Gamma_Ht/2}|B_H(t=0)\rangle\right].\end{aligned} \tag{2.16}$$

Experimentally one does not measure the mass eigenstates but rather the initial and final flavour composition of the meson, so, using once more Equation 2.9

to express mass eigenstates in terms of flavour eigenstates at $t = 0$ we get

$$\begin{aligned} |B(t)\rangle &= g_+(t)|B(t=0)\rangle + \frac{q}{p} g_-(t)|\overline{B}(t=0)\rangle, \\ |\overline{B}(t)\rangle &= \frac{p}{q} g_-(t)|B(t=0)\rangle + g_+(t)|\overline{B}(t=0)\rangle, \end{aligned} \tag{2.17}$$

where

$$\begin{aligned} g_+(t) &= e^{-imt} e^{-\Gamma t/2} \left[\cosh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} - i \sinh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right], \\ g_-(t) &= e^{-imt} e^{-\Gamma t/2} \left[-\sinh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} + i \cosh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right]. \end{aligned} \tag{2.18}$$

In these equations we have used the following definitions of the average and difference for the mass and decay widths:

$$\begin{aligned} m &= \frac{B_H + B_L}{2} = M_{11}, \qquad & \Gamma &= \frac{\Gamma_H + \Gamma_L}{2} = \Gamma_{11}, \\ \Delta m &= B_H - B_L, & \Delta\Gamma &= \Gamma_H - \Gamma_L. \end{aligned} \tag{2.19}$$

It should be noted here that there is a sign convention when defining $\Delta\Gamma$. Under the convention used here, the Standard Model prediction is for $\Delta\Gamma$ to be negative for B_s^0 mesons (and measurements at LHCb confirm this [23]) while some authors use the opposite convention.

From the equations above, the following equations can also be defined:

$$\begin{aligned} |g_\pm(t)|^2 &= \frac{e^{-\Gamma t}}{2} \left[\cosh \frac{\Delta\Gamma t}{2} \pm \cos(\Delta M t) \right], \\ g_+^*(t) g_-(t) &= \frac{e^{-\Gamma t}}{2} \left[-\sinh \frac{\Delta\Gamma t}{2} + i \sin(\Delta M t) \right]. \end{aligned} \tag{2.20}$$

2.4.3 Time-dependent decay rates

The time-dependent decay rate of a sample of N_B mesons which were created as B is then:

$$\Gamma(B(t) \to f) = \frac{1}{N_B} \frac{dN(B(t) \to f)}{dt}, \tag{2.21}$$

where $dN(B(t) \to f)$ is the number of decays into final state f within the infinitesimal time from t to $t + dt$. With an analogous definition possible for $\Gamma(\overline{B}(t) \to f)$

we can restate them as

$$\begin{aligned}\Gamma(B(t)\to f) &= \mathcal{N}_f \left|\langle f|\mathcal{H}|B(t)\rangle\right|^2, \\ \Gamma(\overline{B}(t)\to f) &= \mathcal{N}_f \left|\langle f|\mathcal{H}|\overline{B}(t)\rangle\right|^2,\end{aligned} \tag{2.22}$$

where $\mathcal{N}_f$ is a time-independent normalisation factor containing the result of a phase-space integration to final state f.

Using the definitions of $|B(t)\rangle$ and $|\overline{B}(t)\rangle$ from Equation 2.17, we can rewrite Equation 2.22 in terms of the initial flavour state,

$$\begin{aligned}\Gamma(B(t)\to f) &= \mathcal{N}_f \left|g_+(t)\langle f|\mathcal{H}|B\rangle + \frac{q}{p}g_-(t)\langle f|\mathcal{H}|\overline{B}\rangle\right|^2, \\ \Gamma(\overline{B}(t)\to f) &= \mathcal{N}_f \left|\frac{p}{q}g_-(t)\langle f|\mathcal{H}|B\rangle + g_+(t)\langle f|\mathcal{H}|\overline{B}\rangle\right|^2.\end{aligned} \tag{2.23}$$

Substituting in A_f and $\overline{A}_f$ from Equation 2.10, expressing $\overline{A}_f$ in terms of λ_f from Equation 2.11 and defining $\left|\frac{p}{q}\right|^2 = 1-a$ we get

$$\begin{aligned}\Gamma(B(t)\to f) &= \mathcal{N}_f \left|A_f\right|^2 \left|g_+(t) + g_-(t)\lambda_f\right|^2, \\ \Gamma(\overline{B}(t)\to f) &= \mathcal{N}_f(1-a)\left|A_f\right|^2 \left|g_-(t) + g_+(t)\lambda_f\right|^2.\end{aligned} \tag{2.24}$$

From here it is possible to use Equation 2.20 to define the decay rates in terms of the observables. A similar process can also be conducted for the final state $\overline{f}$ to give all four decay rate equations:

$$\begin{aligned}\Gamma(B(t)\to f) = \frac{1}{2}\mathcal{N}_f|A_f|^2(1+|\lambda_f|^2)e^{-\Gamma t} \quad & \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_f \sinh\left(\frac{\Delta\Gamma t}{2}\right)\right. \\ & \left. + C_f\cos\left(\Delta m t\right) - S_f \sin\left(\Delta m t\right)\right],\end{aligned}$$

$$\begin{aligned}\Gamma(\overline{B}(t)\to f) = \frac{1}{2}\mathcal{N}_f|A_f|^2(1-a)(1+|\lambda_f|^2)e^{-\Gamma t} \quad & \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_f \sinh\left(\frac{\Delta\Gamma t}{2}\right)\right. \\ & \left. - C_f\cos\left(\Delta m t\right) + S_f \sin\left(\Delta m t\right)\right],\end{aligned}$$

$$\begin{aligned}\Gamma(\overline{B}(t)\to \overline{f}) = \frac{1}{2}\mathcal{N}_f|\bar{A}_{\bar{f}}|^2(1+|\bar{\lambda}_{\bar{f}}|^2)e^{-\Gamma t} \quad & \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma t}{2}\right)\right. \\ & \left. + C_{\bar{f}}\cos\left(\Delta m t\right) - S_{\bar{f}} \sin\left(\Delta m t\right)\right],\end{aligned}$$

$$\Gamma(B(t)\to \overline{f}) = \frac{1}{2}\mathcal{N}_f|\bar{A}_{\bar{f}}|^2\frac{1}{1-a}(1+|\bar{\lambda}_{\bar{f}}|^2)e^{-\Gamma t} \quad \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_{\bar{f}}\sinh\left(\frac{\Delta\Gamma t}{2}\right) - C_{\bar{f}}\cos\left(\Delta m t\right) + S_{\bar{f}}\sin\left(\Delta m t\right)\right]. \tag{2.25}$$

Here the *CP* asymmetry parameters C_f, S_f, D_f, $C_{\bar{f}}$, $S_{\bar{f}}$ and $D_{\bar{f}}$ are given by

$$\begin{aligned} C_f &= \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2} \quad , \quad S_f = \frac{2\mathcal{I}m(\lambda_f)}{1+|\lambda_f|^2} \quad , \quad D_f = \frac{2\mathcal{R}e(\lambda_f)}{1+|\lambda_f|^2} \quad , \\ C_{\bar{f}} &= \frac{1-|\bar{\lambda}_{\bar{f}}|^2}{1+|\bar{\lambda}_{\bar{f}}|^2} \quad , \quad S_{\bar{f}} = \frac{2\mathcal{I}m(\bar{\lambda}_{\bar{f}})}{1+|\bar{\lambda}_{\bar{f}}|^2} \quad , \quad D_{\bar{f}} = \frac{2\mathcal{R}e(\bar{\lambda}_{\bar{f}})}{1+|\bar{\lambda}_{\bar{f}}|^2} \quad . \end{aligned} \tag{2.26}$$

It is these *CP* parameters which the analysis described in this thesis will measure for the $B_s^0 \to D_s\pi$ decay.

2.5 The $B_s^0 \to D_s\pi$ decay

To date, it has been assumed by all particle physics experiments that the decay of $B_s^0 \to D_s^-\pi^+$ is flavour-specific, that is that the contribution of $B_s^0 \to D_s^+\pi^-$ is zero or negligible. This assumption is the basis of a number of experiments at LHCb and elsewhere. For example, the measurement of Δm_s using $B_s^0 \to D_s\pi$ decays relies on this assumption in order to calibrate its flavour tagging.

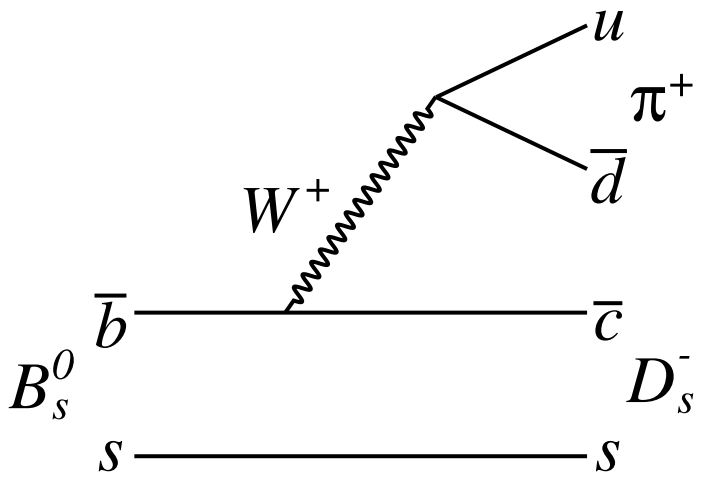


Figure 2.5: Leading order Feynman diagram for $B_s^0 \to D_s^-\pi^+$.

The leading contribution to $B_s^0 \to D_s^-\pi^+$ is shown in Figure 2.5. Comparing this to the decay of $B_s^0 \to D_s^-K^+$ shown in Figure 2.6, it can be seen that the

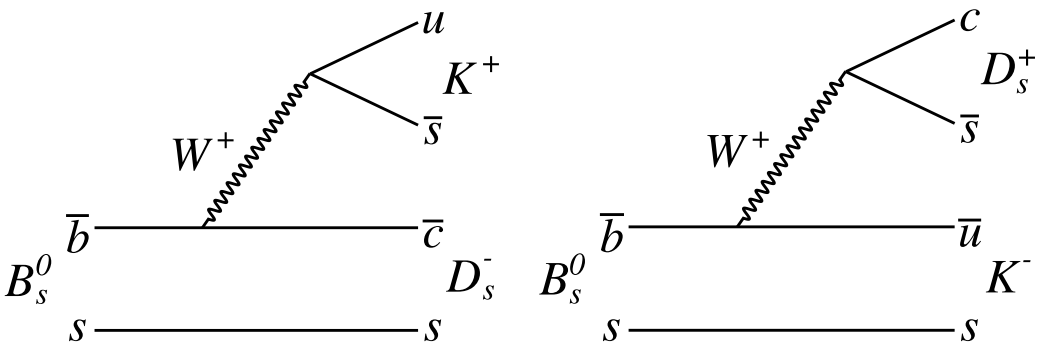


Figure 2.6: Leading order Feynman diagram for $B_s^0 \to D_s^- K^+$ and $B_s^0 \to D_s^+ K^-$.

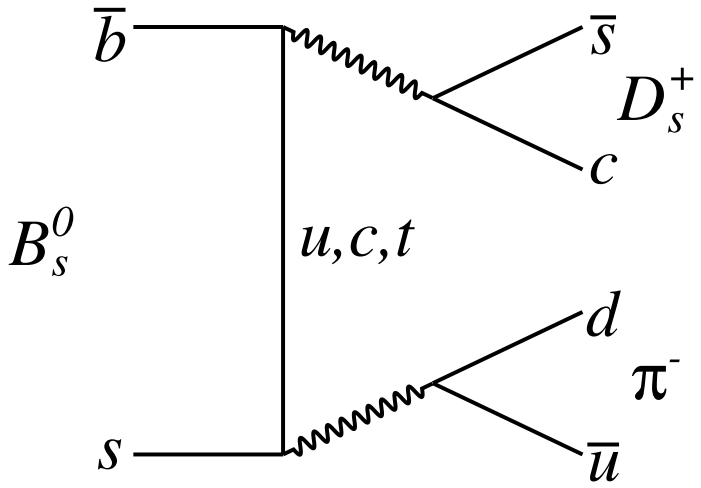


Figure 2.7: Higher order Feynman diagram for $B_s^0 \to D_s^+ \pi^-$.

two diagrams are identical except for the W^+ boson decaying to a $u\,\overline{s}$, giving the different bachelor meson.

To give the leading opposite-flavour contribution to the $B_s^0 \to D_s^+ K^-$ decay — also shown in Figure 2.6 — the $b \to c$ transition is replaced with a $b \to u$ transition to give $B_s^0 \to K^- W^+$ while the W^+ decays to the D_s^+. This simple change is not possible in $B_s^0 \to D_s \pi$ since the B_s^0 meson's s quark cannot make a pion. Instead, a higher order diagram must be used such as the one shown in Figure 2.7. It is clear that this decay will be very suppressed compared to Figure 2.5 due to it being a higher-order diagram. Other possible wrong-flavour contributions will have additional suppression through the GIM mechanism [25].

This high level of suppression yields the assumption that the $B_s^0 \to D_s \pi$ decays are flavour-specific to a very high order. It has been theorised that certain effects beyond the Standard Model could enhance the wrong-flavour contributions to a measurable degree. One route is through an exotic quark, β, with a charge of

$-4/3$ taking part in a $b \to \beta W^+$ transition [26]. This same process in semileptonic decays would provide possible $\Delta B = -\Delta Q$ transitions, in violation of the $\Delta B = \Delta Q$ selection rule.

To date there has been no experimental measurement of $\Delta B = \Delta Q$ in B meson systems or the equivalent in D meson systems while experimental limits of $\Delta S = \Delta Q$ violation in kaons have been placed at the level of 10^{-3} [27].

3

The LHCb detector

3.1 The LHC

The LHC is a proton-proton collider located at CERN, near Geneva in Switzerland. It is a circular accelerator, 27 km in circumference, consisting of straight acceleration sections and bending sections.

Protons are produced from a hydrogen duoplasmatron source and are accelerated to 50 MeV by a linear accelerator. They are fed into the Proton Synchrotron Booster where they are further accelerated up to 1.4 GeV. From here they are passed into the Proton Synchrotron (PS) where they are separated into bunches and accelerated to 25 GeV. After the PS, the protons (in bunches) are accelerated by the Super Proton Synchrotron to an energy of 450 GeV before being injected into the main LHC rings. Under design conditions, the LHC would contain 2808 bunches, each containing 1.15×10^{11} particles. The bunches move in opposite directions around the ring and are collided at 4 distinct points. The LHC was designed to run at a centre of mass energy of 14 TeV at a peak luminosity of $10^{34}\,\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$.

There are 4 main experiments located at the LHC collision points: ALICE, ATLAS, CMS and LHCb. The luminosity at the LHCb collision point is reduced to $3 - 4 \times 10^{32}\,\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ by defocusing and offsetting the beams to reduce the number of interactions per bunch crossing down to an average of 1.6.

Proton-proton collisions ran at the LHC from April 2011 until October 2011. During the 2011 run of data taking, on which the analyses reported within this thesis are performed, $1.2\,\mathrm{fb}^{-1}$ of integrated luminosity was delivered by the LHC of which $1.1\,\mathrm{fb}^{-1}$ was recorded by LHCb as shown in Figure 3.1. However, only about $1.0\,\mathrm{fb}^{-1}$ are actually available for physics analyses since some data had to be

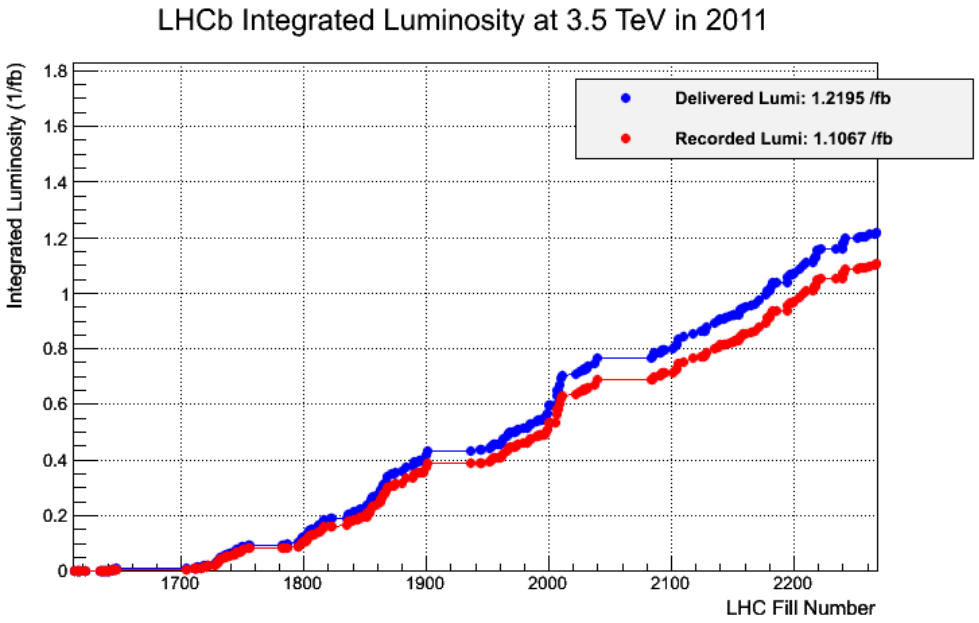


Figure 3.1: Integrated luminosity delivered to LHCb (blue) and recorded by LHCb (red) during 2011 proton-proton collisions.

discarded due to poor quality. The protons in the 2011 run were collided with a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$.

LHCb recorded an average number of interactions per bunch crossing of between 1 and 2 (Figure 3.2) which is above the design specification of 0.4. It also ran with a higher than design level instantaneous luminosity of around $3\text{–}4 \times 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$ (Figure 3.3).

3.2 LHCb

LHCb [28] is designed to focus on the investigation of *CP* violation in *B* meson decays and searches for physics beyond the standard model. There are many analyses that are being performed to investigate *CP* violation and numerous stand-alone results can hope to be found. Some of the main results being looked for by LHCb are more accurate measurement of the angles of the unitarity triangles, branching fractions of heavy meson decays and searches for new physics.

The LHCb detector (Figure 3.4) is a single arm spectrometer with angular coverage from 10 mrad to 250 mrad in the vertical plane (300 mrad in the horizontal). It is placed at one of the crossing points of the LHC (intersection point 8) where bunches of protons going in opposite directions will collide. LHCb is designed

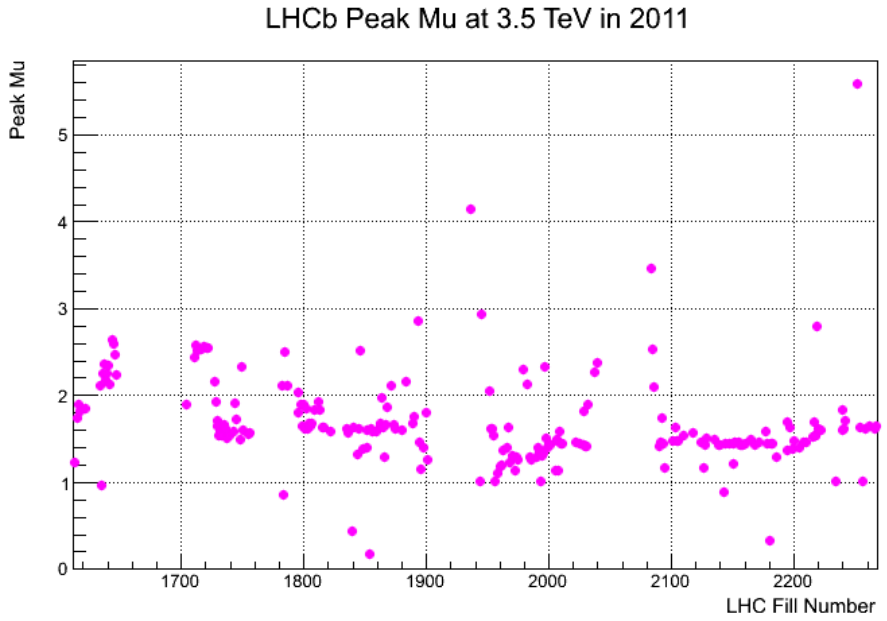


Figure 3.2: Peak interactions per bunch crossing (μ) per LHC fill at LHCb in 2011.

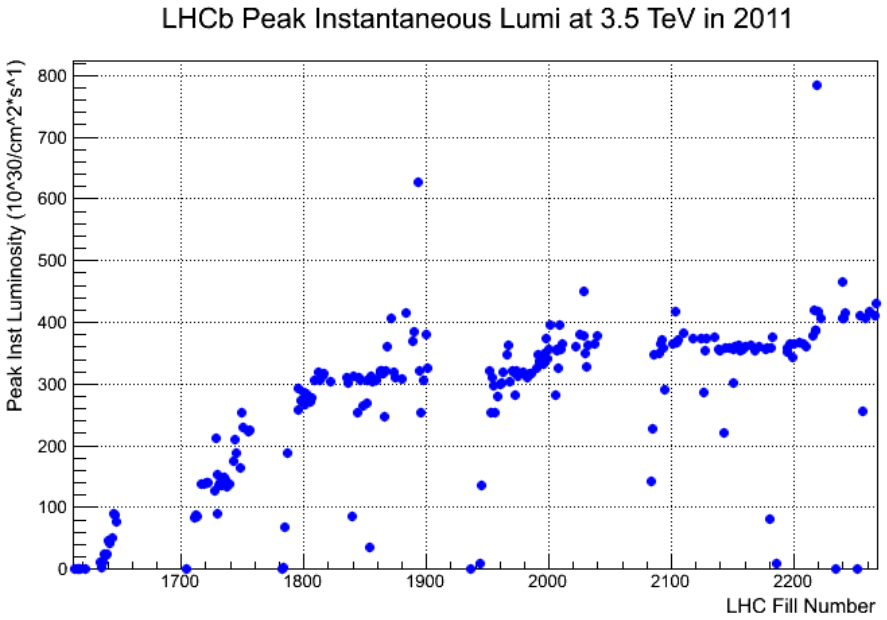


Figure 3.3: Peak instantaneous luminosity per LHC fill at LHCb in 2011.

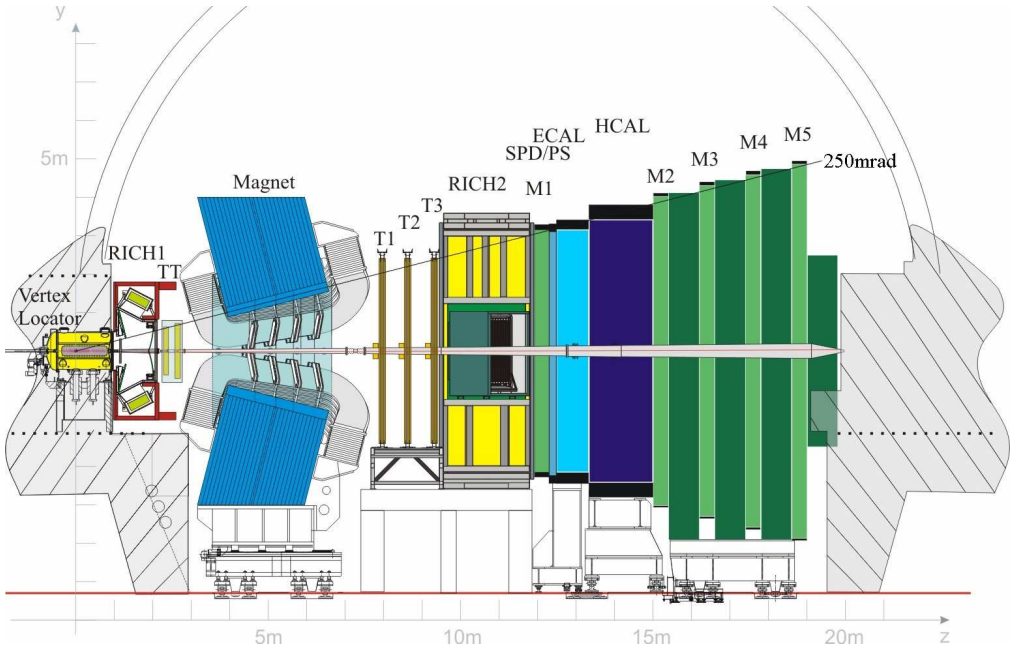


Figure 3.4: A side view of the LHCb detector showing the primary interaction point of the collider at the far left. Particles produced will travel towards the right and be detected by the various components. Reproduced from Ref [28].

for the detection of B (containing an $\overline{b}$ quark) and $\overline{B}$ (containing a b quark) mesons from this collision. At the high energies at which the LHC operates the B and $\overline{B}$ mesons will be preferentially produced in the longitudinal direction in a fairly tight cone (see Figure 3.5). That is, the direction of most of the mesons will be along the beam pipe in either direction — creating two cones with their points touching at the primary collision vertex. On average the two cones will be identical. For this reason and due to cost constraints the detector is placed to cover just one of the cones. This is different to many other particle detectors (such as ATLAS and CMS) which completely encase the interaction point so as to try to capture all particles produced.

The focus on B physics study puts many requirements on the design of the detector such as a need for excellent primary and secondary vertex resolution to allow for precise measurements of the proper decay time as well as good momentum resolution. Further to this, excellent particle identification and efficient triggering is needed. These are essential for the study of neutral B meson oscillations and decays.

The LHCb detector is designed to run at a much lower luminosity (a factor 50 lower) than the LHC's nominal value to reduce the average number of proton-

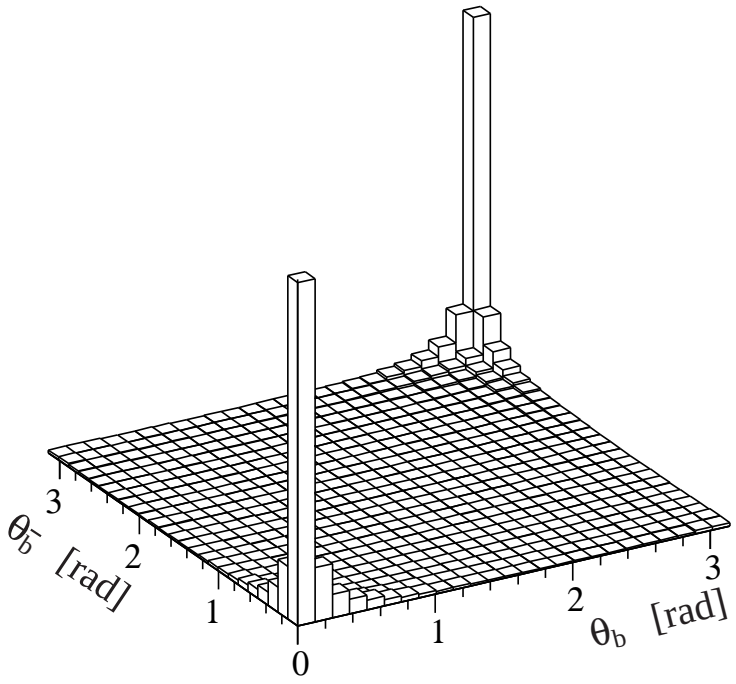


Figure 3.5: The simulated production angle of B hadrons with respect to the beam line. Reproduced from Ref [29].

proton collisions per bunch crossing to a much lower number. By design the number of collisions per bunch crossing, μ, was 0.4 while actual running conditions were closer to 1.4. A process known as *luminosity levelling* is used to both reduce the total instantaneous luminosity as well as provide a constant luminosity throughout a run. The two beams are offset from each other in the vertical direction to reduce the cross-section of the overlap of the two beams. As the run continues and the bunches are steadily depleted of protons, the beams are gradually brought back together to maintain the average number of collisions. If the beams were simply defocused then at the beginning of the run there would be too many collisions per crossing and by the end of the run there would be too few.

The detector comprises several sub-detectors, each of which serves a different purpose and can be categorised into two groups — tracking detectors and particle identification (PID) detectors.

3.2.1 The Vertex Locator (VELO)

The first sub-detector that a particle produced in the initial collision traverses is the Vertex Locator which is a tracking detector. This serves to measure the position of vertices — that is where the products of the proton-proton collision decay as

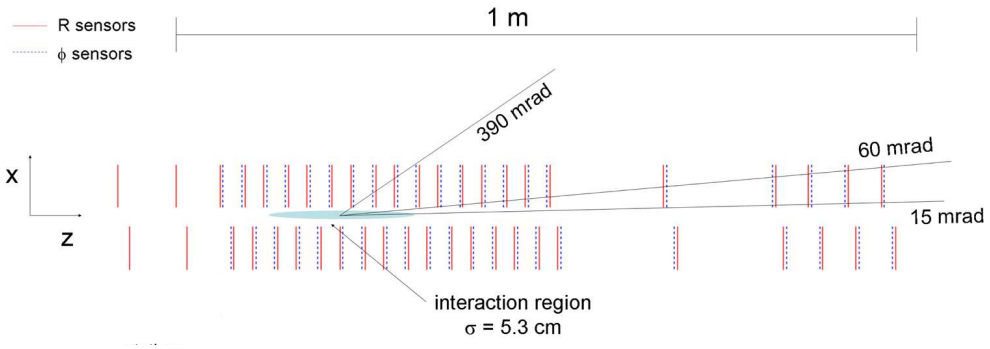


Figure 3.6: Layout of the VELO subdetector. This detector closely surrounds the primary interaction region of the collider. Shown is a slice through the VELO at $y = 0$ (top-down). The red and blue segments are the sensor modules. Reproduced from Ref [28].

well as the primary proton-proton collision vertex. Within it are a series of about 20 pairs (called stations) of sensor plates or "modules" (see Figure 3.6) each of which record the positions of particles passing through in 2 dimensions. If a particle passes through at least two of the stations then it is (in principle) possible to trace back the path to the point at which it was created and the VELO is designed such that any track which is within the angular acceptance of the rest of the LHCb detector will pass through at least 3 stations. As can be seen from the figure, the position of the modules is such that only particles with very small or very large ($15\,\text{mrad} > \phi > 300\,\text{mrad}$) angle will escape without being tracked. The total coverage of the LHCb detector is defined by $\phi < 250\,\text{mrad}$. The more stations that a particle travels though, the more accurate the measurement of the vertex will be. The resolution of the VELO is designed to be about $4\,\mu\text{m}$ for particles at $\phi = 100\,\text{mrad}$.

Separating the VELO stations from the vacuum of the beam is the RF foil which is a pair of shaped aluminium sheets each containing half the modules. The RF foil serves two purposes, firstly it is there to protect the LHC vacuum from outgassing of the VELO modules. It also acts as a shield to protect the VELO from radio-frequency pickup from the LHC beams as well as protecting the beam from wake fields which are generated as the beams pass through the VELO.

Each module contains both an R and ϕ sensor and sits on one side of the beam line. The general layout is shown in Figure 3.7. Both modules are silicon strip detectors arranged in a semi-circular annulus with an outer radius of $41.9\,\text{mm}$ and an inner radius of $8\,\text{mm}$. The inner radius is designed to be as small as possible to get close to the interaction point but is restricted by a minimum safe distance to

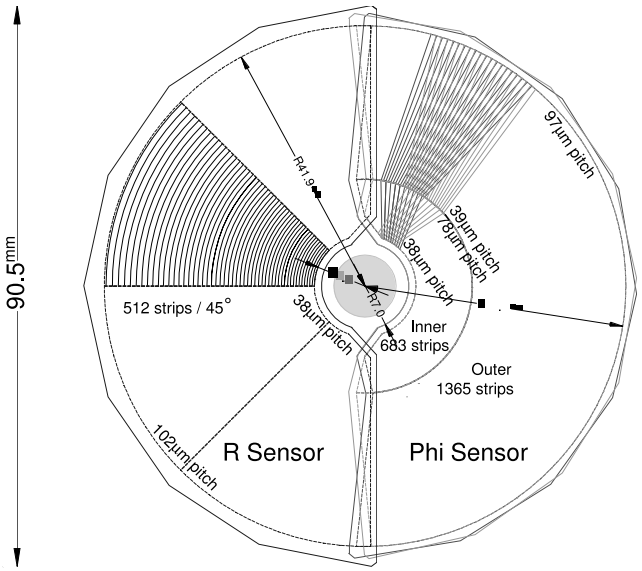


Figure 3.7: Layout of a pair of VELO modules. Shown are both radial and ϕ strips. Reproduced from Ref [28].

the beam (5 mm), 1 mm for the guard structures on the silicon and to leave space for the RF foil.

The R sensors consist of 512 strips running concentrically. Within each module the sensors are divided into 4 segments. This segmentation reduces primarily strip occupancy but also reduces the strip capacitance. The pitch of each strip increases radially to keep the occupancy per strip approximately constant as the strip length increases and the particle flux decreases. The pitch varies from 38 μm at the inner edge of the sensor out to 101.6 μm at the outer edge.

The ϕ sensors are divided into two sections radially again to keep the occupancy low and to stop the strip pitch from getting too large at high radii. The strips are not aligned perfectly in the radial direction but are skewed at an angle of 20° in the inner section and -10° in the outer section with respect to the radial direction. In alternating stations the angles are reversed to create a stereo effect. The pitch of the strips increases linearly towards the outer edge of the sensor. It is 35.5 μm at the inner edge of the inner section and increases to a value of 78.3 μm at the section border. The outer section then starts from a pitch of 39.3 μm and increases up to 97 μm.

The VELO provides excellent vertex resolution which is required for the types of analyses LHCb performs. As can be seen in Figure 3.8, the primary vertex resolutions when there are 25 tracks are: $\sigma_x = 13.1\,\mu\text{m}$, $\sigma_y = 12.5\,\mu\text{m}$, $\sigma_z = 71.1\,\mu\text{m}$ which are very close to the design values.

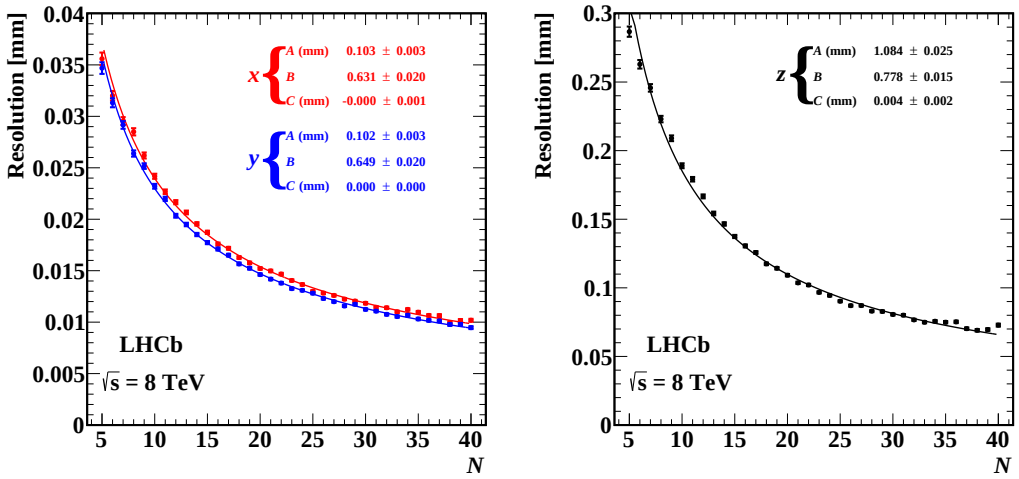


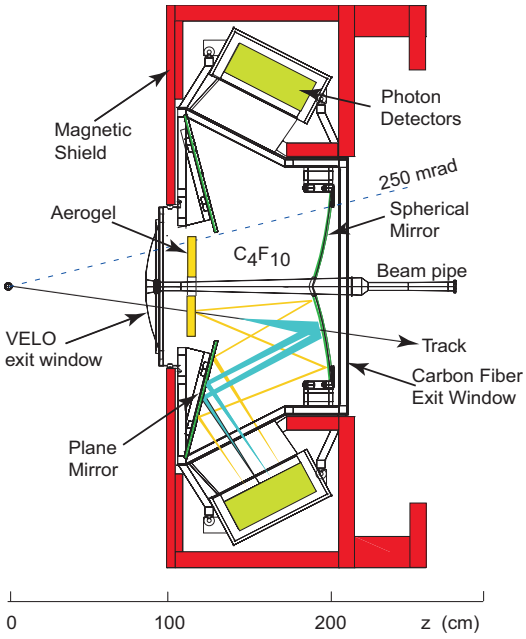
Figure 3.8: Resolution of the VELO with respect to different numbers of tracks (N) for events with one primary vertex from data collected in 2011. On the left is the x (red line) and y (blue line) resolution and on the right is the z resolution. Reproduced from Ref [30].

3.2.2 Ring Imaging Cherenkov (RICH) detector

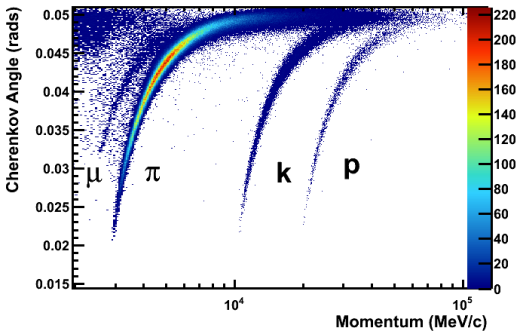
Within LHCb there are two RICH detectors, both used for particle identification — most importantly to differentiate between pions and kaons in the decay of B mesons. The first (RICH1, shown in Figure 3.9a) covers low momentum ($\sim 1 - 60\,\text{GeV}/c$) particles while the one further downstream (RICH2) covers higher momentum particles ($15\,\text{GeV}/c$ and higher).

In both of the detectors, spherical and flat mirrors are used to focus Cherenkov light produced by the particles travelling through the medium on to a set of hybrid photo-detectors (HPDs). The RICH1 detector uses a combination of aerogel and C_4F_{10} while RICH2 uses just CF_4. The difference in the radiator material and the layout of the mirrors provide good PID over a range of momenta. This is shown in Figure 3.9b where it can be seen that the combination of the aerogel and gas in RICH1 provides excellent particle differentiation — particularly for kaons and pions — across the range of momenta which is interesting for LHCb analyses as well as for leptons at lower energies.

Both RICH detectors use Hybrid Photon Detectors (HPDs) to detect the Cherenkov light. An example of one is shown in Figure 3.10. There are 196 HPDs in RICH1 and 288 HPDs in RICH2 placed in two planes in each detector. They are arranged in a hexagonal pattern to provide the best coverage given their circular profile.



(a) RICH1: Particles approach from the left after leaving the VELO and pass out the right-hand side. Reproduced from Ref [28].



(b) Reconstructed Cherenkov angle, θ, against the momentum of various tracks when passing through the C_4F_{10} radiator. Reproduced from Ref [31]

Figure 3.9: Ring Imaging Cherenkov (RICH) detectors.

An incoming photon will interact with the photocathode layer on in the inside of the quartz window and produce a photoelectron. The electron is accelerated by a 20 keV voltage onto a silicon pixel array at the back of the HPD. The spatial resolution of a HPD is 2.5 mm with a time resolution of 25 ns.

For analyses, the main output of the RICH system is a set of PID variables called *delta log-likelihoods* (DLLs). For each event, initially all tracks in the event are assumed to be pions. Based on this assumption, the probability distribution of finding photons in each HPD is calculated. This probability distribution is compared against the observed hits to calculate a PID likelihood. The hypothesis of each particle is then changed in turn to be a kaon, a proton and so on, and for each of these alternative hypotheses a global likelihood is once again calculated. The values of the likelihoods of the pion hypothesis and, for example, the kaon hypothesis is used to calculate the PID variable $\mathrm{DLL}_{K\pi}$ which is defined as

$$\mathrm{DLL}_{K\pi} = \ln \mathcal{L}(K) - \ln \mathcal{L}(\pi), \tag{3.1}$$

where $\mathcal{L}(K)$ is the global likelihood given the kaon hypothesis (and likewise for π). This means that $\mathrm{DLL}_{K\pi}$ can be used as a differentiator between kaons and pions.

Figure 3.10: Photograph of one of the HPDs used inside the RICH detector. Reproduced from Ref [28].

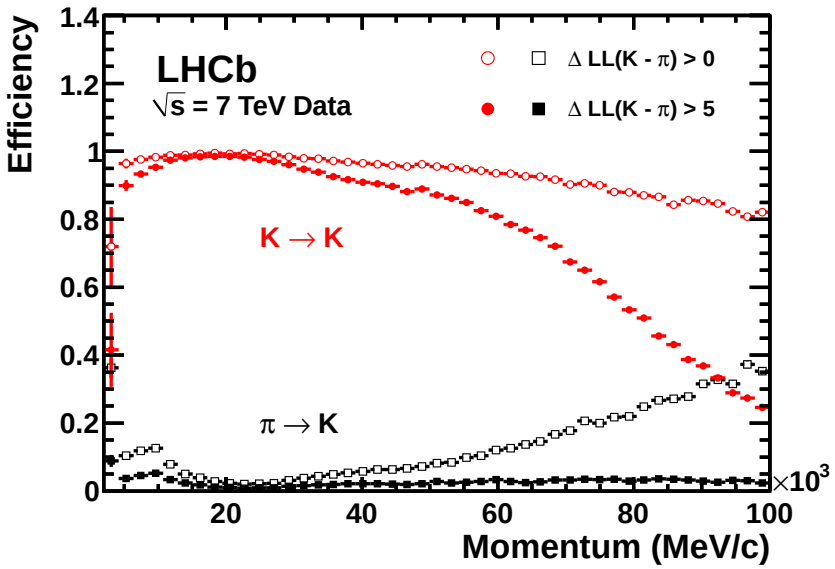


Figure 3.11: Kaon identification efficiency (in red) and pion misidentification as kaon rate (in black) requirements for data as a function of track momentum. Two different $\Delta \log \mathcal{L}(K - \pi)$ have been imposed on the samples, resulting in the open and filled markers. Reproduced from Ref [31].

The efficiency of kaon identification as a function of particle momentum is shown in Figure 3.11.

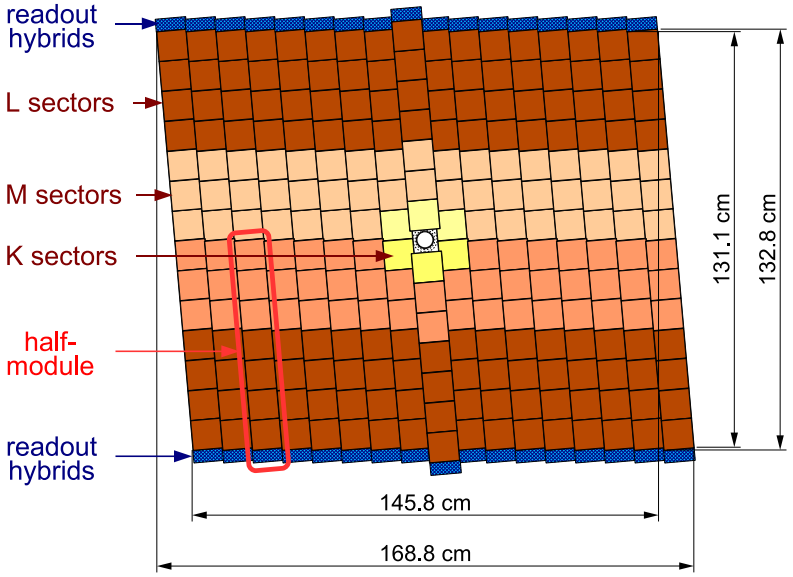
3.2.3 Tracking

The tracking subsystem comprises four subdetectors: the VELO (as discussed in Section 3.2.1), the Tracker Turicensis (TT), the Inner Tracker (IT) and the outer tracker (OT).

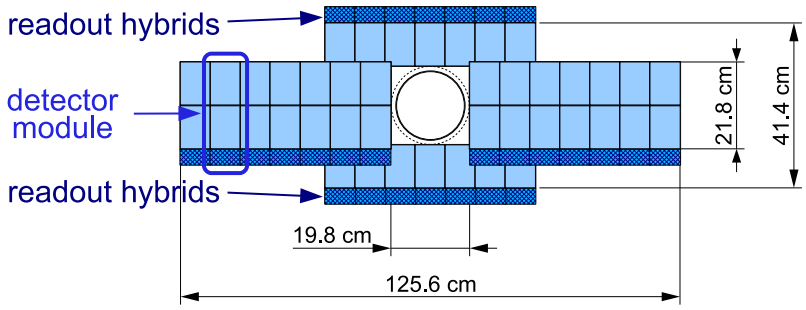
Silicon Tracker

Since the TT and the IT use the same technology, they are together referred to as the Silicon Tracker. The TT is just upstream of the main magnet and the IT is positioned downstream of the magnet. As their names suggests, they are tracking detectors, used to accurately locate the positions of particles in order to be able to reconstruct their paths.

Each detector is constructed as strips of silicon with an average pitch of 200 µm. The TT is made up of four layers with the second and third layer placed at an angle ($-5°$ and $+5°$ respectively) and each layer is placed approximately 30 cm



(a) Layout of the third TT detection layer. The different shadings indicate different readout sections.



(b) Layout of an x detection layer in the second IT layer.

Figure 3.12: Layout of the Silicon trackers. Reproduced from Ref [28].

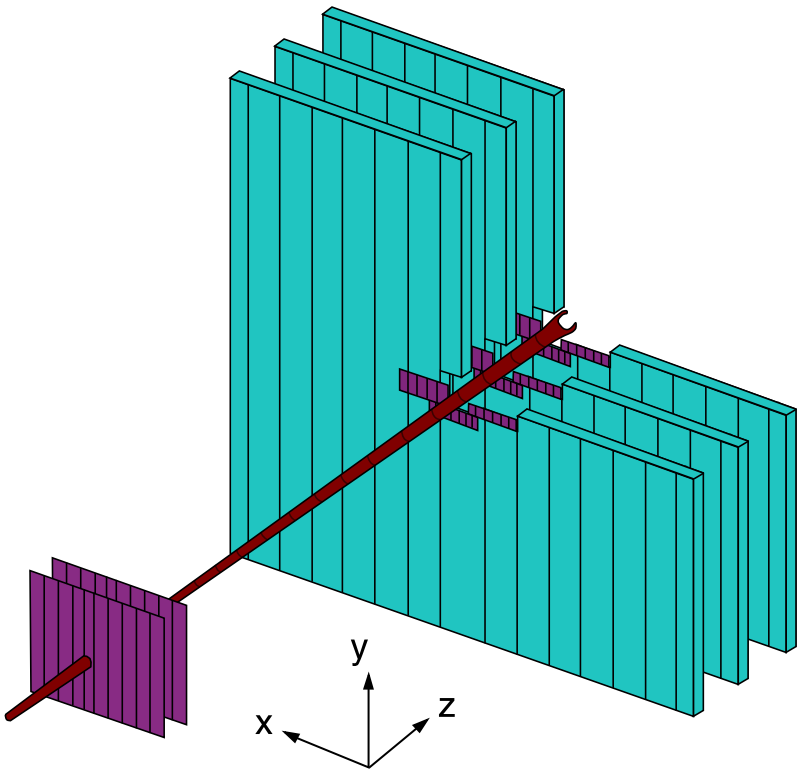


Figure 3.13: Positions of the TT and IT with respect to the Outer Tracker. The TT is on the left of the figure, upstream of the Inner and Outer Trackers. The IT (shown here in purple) is the smaller tracker near the beam pipe on the right, surrounded by the OT.

apart. A schematic view of the third layer can be seen in Figure 3.12a where the $+5°$ angle is visible.

The IT covers a small area of the LHCb acceptance near to the beam pipe. It is made up of three layers and is shown in Figure 3.13. The second and third layers are placed at a stereo angle as in the TT (the layout of the second layer is shown in Figure 3.12b).

Both trackers have a resolution of approximately $50\,\mu m$.

Outer Tracker

The outer tracker is placed at the same z positions as the IT and covers the rest of the angular acceptance (see Figure 3.13). It uses arrays of drift tubes filed with 70% Ar and 30% CO_2. It provides a spatial resolution of $200\,\mu m$ and, as with the silicon trackers, the second and third layers are placed at an angle of $-5°$ and $+5°$ with respect to the vertical axis.

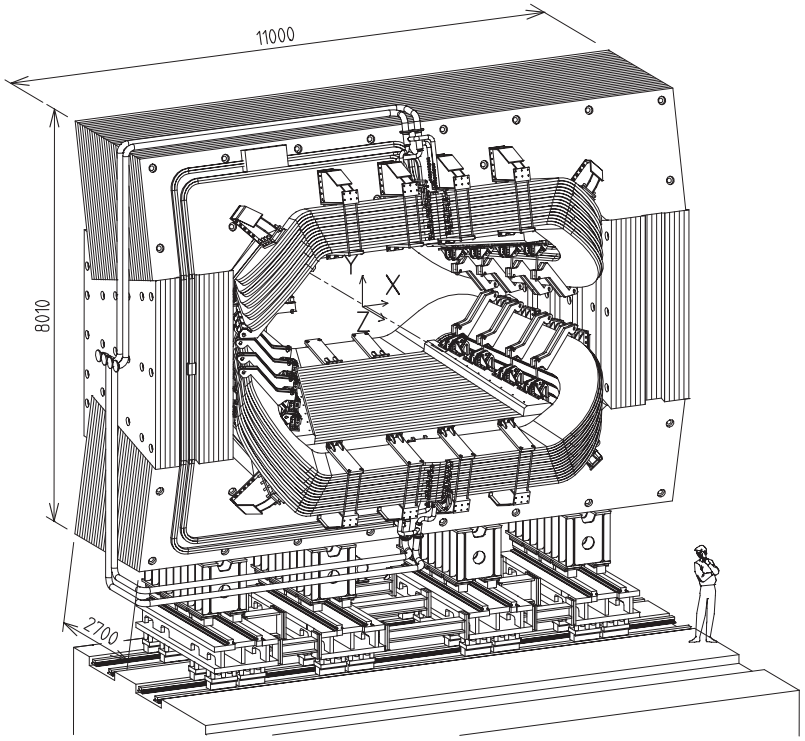


Figure 3.14: The LHCb dipole magnet. Reproduced from Ref [28].

3.2.4 Magnet

The LHCb dipole magnet (shown in Figure 3.14) is placed downstream of the TT but before the first inner and outer tracking station. It provides a vertical magnetic field to enable measurement of the momenta and charges of particles. The opening in the centre of the magnet is designed to be large enough to sit entirely outside the acceptance of the rest of the detector. In order to reduce systematic uncertainties, particularly in *CP* measurements, the magnet's polarity can be inverted. During normal running, the magnet was set to each configuration for approximately equal amounts of time.

In order to achieve the necessary momentum resolution, the magnet provides a peak field strength of about 1.1T. The strength of the field is measured throughout the interior of the magnet with a Hall probe. The strength of the field is shown in Figure 3.15.

3.2.5 Calorimeter

The calorimeter system in LHCb contains four sections: the scintillator pad detector (SPD), preshower detector (PS), an electromagnetic calorimeter (ECAL) and

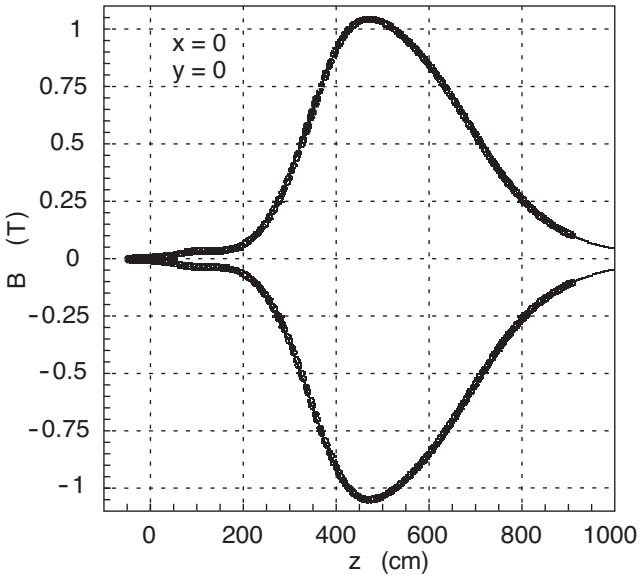


Figure 3.15: The magnetic field strength measured on-axis as a function of z. Reproduced from Ref [28].

a hadron calorimeter (HCAL). It measures the energy and position of particles by providing a heavy target to cause showers of particles. It provides particle identification for photons, electrons and hadrons. Since neutral particles will leave no trace in the tracking system, reconstructing neutral pions and prompt photons in the calorimeter is essential for studies of many decays.

The SPD/PS detector uses two layers of scintillator pads sandwiching a 15 mm Pb converter plate. Its main purpose is to validate and cross-check any signals that the ECAL receives. Any neutral particles will not interact with the scintillator pads which allows the ECAL to differentiate between high energy photons and electrons. It is also used for global event cuts to measure event activity.

The ECAL is built with layers of scintillating tiles (4 mm thick) alternating with lead (2 mm thick) acting as active material and absorber respectively. Readout is achieved with wavelength-shifting fibres, embedded in the scintillator tiles which read out into phototubes. The ECAL is split into three sections as shown in Figure 3.16 due to the fact that the track hit density varies by two orders of magnitude over the surface of the detector.

The HCAL's main purpose is to provide information for the hadron trigger and so is designed to have a very fast response time, even at the expense of good energy resolution. The design of the HCAL is similar to that of the ECAL so it uses alternating layers of scintillator and steel absorber plates. As seen in Figure 3.16, the HCAL is only split into two sections due to the wider shape of hadronic showers.

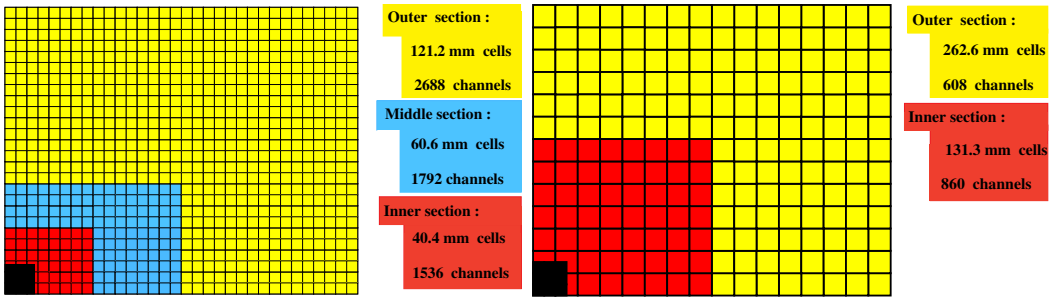


Figure 3.16: Lateral segmentation of the SPD/PS and ECAL (left) and the HCAL (right). One quarter of the detector front face is shown. In the left figure the cell dimensions are given for the ECAL. Reproduced from Ref [28].

3.2.6 Muon system

The muon system is the last stage in the LHCb detector — providing identification of muons. The precise measurement and identification of muons is essential to many of the LHCb measurements as they are present in the final state of many interesting B meson decays.

Being able to accurately distinguish muons from other particles is important for many of the key measurements that LHCb is making, such as the search for new physics in measurements of the branching fraction of $B_s^0 \to \mu^+\mu^-$. Being able to cleanly separate muons from other particles is critical to be able to measure decay channels with a very low branching fraction (a few 10^{-9}) for the decay $B_s^0 \to \mu^+\mu^-$.

The muon system provides information for the L0 high-p_{T} muon trigger as well as muon identification for both the HLT and for offline analysis.

The muon system consists of five rectangular stations as shown in Figure 3.17. The first muon station (M1) is upstream of the calorimeters and the other 4 (M2-M5) are downstream. M1 is used to improve the p_{T} measurement for the trigger. The key differentiator to identify a muon is that most particles will stop within the calorimeter system whereas only muons will manage to travel all the way through to M2-M5. Consequently any particles found in those systems are very likely to be muons. This provides a very clean muon signal.

There are two types of detector technology used in the muon system. The centre region of M1 (nearest the beam pipe) uses a triple Gas Electron Multiplier (GEM) and the rest of M1 and the entirety of M2-M5 use Multi-Wire Proportional Chambers. The detectors provide point measurements of particle tracks, giving a binary decision for whether a track was detected. Stations M1-M3 provide good spatial resolution in the x direction and are used to calculate the p_{T} of the muon (with a resolution of 20%) along with the track direction. The final two stations have

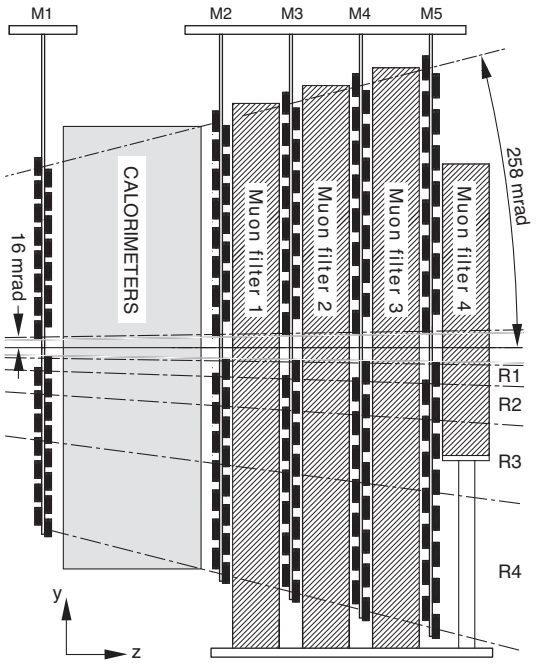


Figure 3.17: The position of the muon detectors with respect to the calorimeters. Reproduced from Ref [28].

much more limited spatial resolution being mainly used for particle identification.

3.2.7 Trigger

Almost all LHC bunch crossings will produce interesting data for physics analyses. However, due to storage space constraints, only a very small number of these can be fully recorded. To ensure that the most interesting events are stored, an efficient trigger is required.

The LHCb trigger works on a two-level system as shown in Figure 3.18 where the first trigger level (L0) is a hardware trigger and the high-level trigger (HLT) is a software trigger. The L0 uses only a limited amount of information from the detector. It is used to automatically determine whether any particular proton-proton collision event is interesting and so whether it should be subjected to further analysis — making its decision only $4\,\mu s$ after the initial collision. The input to the L0 is at a rate of up to $40\,MHz$ while it has to output at only $1\,MHz$ which is a limit set by the maximum rate of the readout electronics. There are two main parts of the L0 trigger, the calorimeter trigger and the muon trigger. The calorimeter trigger uses information from the ECAL and HCAL to select tracks with high transverse energy (E_T) and to select the photons, electrons, π^0 and hadron candidates with

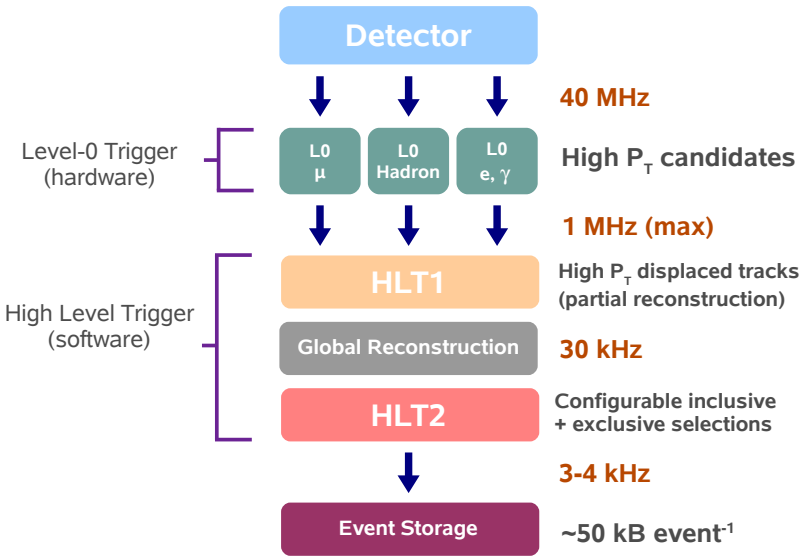


Figure 3.18: Overview of the LHCb trigger system. Reproduced from Ref [32].

the highest E_{T}. It uses information from the PS and SPD to aid with the particle identification. The muon trigger selects muon candidates with high p_{T}. To fire the trigger, a track must be present in all five of the muon stations, pointing to the interaction point.

The second level of the the trigger, called the "High-Level Trigger" (HLT), is split into two stages (HLT1 and HLT2) and is implemented in software. While the L0 trigger only uses partial information from certain subdetectors, the HLT performs partial reconstruction of the event to be able to have access to higher level information such as reconstructed tracks. HLT1 consists of a set of approximately 20 trigger selection algorithms each of which makes a decision based on the presence of one or two tracks (such as two muons or a single hadron) which match certain criteria, for example having high p_{T}. It outputs at a maximum rate of 40 kHz. The HLT2 stage then runs over those events which passed the HLT1. The HLT2 attempts to reconstruct the event in a way as similar to the full offline reconstruction as possible and triggers on inclusive decay signatures or the presence of exclusive charm and beauty hadron candidates.

3.3 LHCb software

The LHCb software framework is built on GAUDI [33] which provides an extensive framework for building HEP data analysis tools. It is an object-oriented C++ toolkit, providing all the common tasks that analysis software would need such as

data access and histogram management. Tools can be build using the framework in a modular fashion allowing the fast switching of components as the task requires.

3.3.1 Simulated event production

In order to perform in depth studies of the detector and for physics analyses, Monte-Carlo (MC) simulated events are used. Within LHCb, event generation is performed by two applications, Gauss [34] and Boole [35] and is broken into a number of steps:

Event generation Gauss uses Pythia [36] to simulate the interaction and scattering of proton-proton interactions. This produces elementary particles and hadrons which are propagated and decayed using EvtGen [37].

Event simulation Once the particles have been generated, their interaction with the LHCb detector is simulated using Geant4 [38]. Their interaction with the bulk of the detector is simulated as well as with sensitive parts of the detector such as energy deposits in the silicon detectors and Cherenkov photon creation in the RICH detectors.

Digitisation The simulated interaction with the detector results in energy being present at a specific place in the silicon and a number of photons being present in the RICH's HPDs. Boole The [35] platform is used to simulate the detector's response to these signals and convert it into virtual readout channels as the physical detector would. After this step, the data format and contents should be identical to that of real data which allows the same algorithms to be run over both simulated and real data.

The HLT software is managed by a piece of software called Moore [39]. This is usually run automatically as part of the trigger setup but it can also optionally be run "offline" on simulated events to simulate the trigger and so end up with an identical selection of events to real data.

3.3.2 Reconstruction

Reconstruction of events in LHCb is performed using Brunel [40]. It is able to run identically over simulated or real data since by this stage they should be identical in format. Reconstruction of the particle tracks based on the detector response is performed using pattern recognition algorithms. These tracks are then used in particle identification (PID) process to assign particle types to each track.

3.3.3 Analysis

The final stage in the LHCb software chain is DaVinci [41]. It is a Gaudi-based framework providing all the information needed to perform an analysis such as kinematic information about the particles, PID information and overall event information. Starting from a list of the particles in an event it can find decay chains and create a ROOT [42] ntuple containing the relevant events.

Part II

$B_s^0 \to D_s K$ analysis

4

Introduction

I was involved with the measurement of the time-dependent *CP*-violation observables in $B_s^0 \to D_s K$ decays. While that analysis focused on $B_s^0 \to D_s K$, it included $B_s^0 \to D_s \pi$ as a cross-check channel and so a full event selection, mass fit and systematic uncertainty study was performed for that channel. Since much of the work performed for the $B_s^0 \to D_s K$ study feeds directly into the main analysis topic presented in this thesis I will here give an overview of the analysis, particularly as it pertains to the main analysis on $B_s^0 \to D_s \pi$. A full internal analysis note was written up as Ref [1] with further details and was submitted as a conference paper as Ref [2].

The purpose of the analysis is to measure C, S_f, $S_{\bar{f}}$, D_f and $D_{\bar{f}}$ on $B_s^0 \to D_s K$ decays from 2011 data from LHCb over a dataset of integrated luminosity $\int \mathcal{L} = 1.0$ fb^{-1} of pp collisions recorded at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. These *CP* parameters are related to the physics parameters $r_{D_s K}$, Δ and $\gamma - 2\beta_s$ by

$$C = \frac{1 - r_{D_s K}^2}{1 + r_{D_s K}^2}, \tag{4.1}$$

$$D_f = \frac{2 r_{D_s K} \cos(\Delta - (\gamma - 2\beta_s))}{1 + r_{D_s K}^2}, \tag{4.2}$$

$$D_{\bar{f}} = \frac{2 r_{D_s K} \cos(\Delta + (\gamma - 2\beta_s))}{1 + r_{D_s K}^2}, \tag{4.3}$$

$$S_f = \frac{2 r_{D_s K} \sin(\Delta - (\gamma - 2\beta_s))}{1 + r_{D_s K}^2}, \tag{4.4}$$

$$S_{\bar{f}} = \frac{2 r_{D_s K} \sin(\Delta + (\gamma - 2\beta_s))}{1 + r_{D_s K}^2}, \tag{4.5}$$

where $r_{D_sK} = \left|A(\overline{B}^0_s \to D^-_s K^+)/A(B^0_s \to D^-_s K^+)\right|$ is the ratio of the magnitudes of the decay amplitudes and Δ is the strong phase difference. The B^0_s mixing phase, β_s is predicted by the Standard Model to be small and so from this it is possible to constrain the CKM angle γ.

5

Data selection

5.1 Data sample

This analysis uses data from the 2011 run of LHCb. This comprises an integrated luminosity $\int \mathcal{L} = 1.0\ \text{fb}^{-1}$ of pp collisions recorded at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$.

5.2 Simulated data

Several samples of simulated data were created for the analysis, primarily for the use in event selection and background studies. In each sample, a B hadron is forced to decay to a specific final state as listed in Table 5.1 along with the number of events generated for each channel.

5.3 Reconstruction

The $B_s^0 \to D_s K$ decay mode is reconstructed in two stages, first the D_s candidate is created from its daughter particles and then a K is added to make the B_s^0 meson. The D_s candidates are reconstructed in three separate final states, $D_s^+ \to K^+K^-\pi^+$, $D_s^+ \to K^+\pi^-\pi^+$ and $D_s^+ \to \pi^+\pi^-\pi^+$ each of which are selected as independent samples based on particle identification requirements. The invariant mass of the combination of the three D_s meson daughters is fixed to the nominal value of the D_s meson when reconstructing the mass of the B_s^0 meson and, conversely, when calculating the decay time of the B_s^0 meson, the mass of the D_s meson is not constrained but the momentum vector of the B_s^0 is required to point from the pp primary vertex.

Sample		**Sample size**
$B_s^0 \to D_s\pi$	$D_s^+ \to K^+K^-\pi^+$	1052495
$B_s^0 \to D_sK$	$D_s^+ \to K^+K^-\pi^+$	1887293
$B_s^0 \to D_s^*\pi$	$D_s^+ \to K^+K^-\pi^+$	524098
$B_s^0 \to D_s^{*-}K^+$	$D_s^+ \to K^+K^-\pi^+$	206000
$B_s^0 \to D_s\rho$	$D_s^+ \to K^+K^-\pi^+$	2019391
$B_s^0 \to D_s^*\rho$	$D_s^+ \to K^+K^-\pi^+$	1019191
$\Lambda_b \to D_s^-p$	$D_s^+ \to K^+K^-\pi^+$	539994
$\Lambda_b \to D_s^{*-}p$	$D_s^+ \to K^+K^-\pi^+$	630598
$\Lambda_b \to \Lambda_c\pi$	$\Lambda_c^+ \to pK^-\pi^+$	2033496
$\Lambda_b \to \Lambda_cK^-$	$\Lambda_c^+ \to pK^-\pi^+$	519495
$B^0 \to D\rho$	$D^+ \to K^-\pi^+\pi^+$	2054494
$B^0 \to D^*\pi$	$D^* \to D\pi^0$, $D^+ \to K^-\pi^+\pi^+$	1046498
$B^0 \to D\pi$	$D^+ \to K^-\pi^+\pi^+$	1016198
$B^0 \to D^-K^+$	$D^+ \to K^-\pi^+\pi^+$	958393
$B^0 \to D_s^-K^+$	$D_s^+ \to K^+K^-\pi^+$	517198

Table 5.1: Simulated samples used during the analysis for selection and background studies.

The sample is further divided based on the polarity of the magnet (*up* or *down*) as well as the flavour tagging information of the B_s^0 meson candidate (B_s^0, $\overline{B}_s^0$ or untagged). Thus there are 18 sub-samples of data with no events being present in more than one data set.

5.4 Event selection

The event selection is performed in a four-step process, defined partially by the LHCb experimental considerations. The steps of the selection process are:

1. trigger,
2. experiment-wide offline selection (*stripping*),
3. analysis-specific offline selection,
4. particle identification.

The details of which are covered in the rest of this chapter.

5.5 Trigger

The first level of selection is performed by the LHCb trigger system described in Section 3.2.7. All events for this analysis are required to be those which contained the particles which the trigger used to make its decision. This means that the track which activated the trigger is required to be used in the reconstruction of the signal candidate. For an event to be considered, two independent trigger algorithms must have fired. First, in the HLT1, a region of interest is defined by a straight line track in the VELO and then a single detached high-momentum track is required to be within that region. This trigger (internally known as `1TrackAllL0`) is detailed in a dedicated note at Ref [43]. Secondly, the HLT2 trigger is required to have fired on the detection of a single, high-momentum track, displaced from the pp collision point and to have found a single similarly displaced vertex containing the detected track and 1–3 other tracks. This trigger algorithm (called the 2-, 3- or 4-body `TopoBBDT`) is described by a public note at Ref [44].

5.6 Stripping selection

Stripping is performed centrally within the LHCb collaboration and the results of it are made available to all through the standard LHCb book-keeping processes. Its primary purpose is to provide a number of data sets, each defined by a set of relatively loose selection criteria, to be used by numerous analyses within the LHCb collaboration. While the stripping selection is performed offline, after the events have been stored to disk, it is treated as a separate step to the *offline selection*.

The selection (called a *stripping line* within LHCb) used to select the initial set of B_s^0 candidates for this analysis is the `StrippingB02DPiD2HHHBeauty2Charm-Line` and it is performed as a two-step process. First a loose pre-selection is made based on the kinematics of the particles and their displacement from the primary interaction. All charged particles which are used to make the B_s^0 meson are required to have a track $\chi^2/\text{ndof} < 4$, $p_{\text{T}} > 100\,\text{MeV}/c$ and $p > 1\,\text{GeV}/c$. Finally, each track used to reconstruct the B_s^0 meson is, in turn, artificially combined with the tracks used to create the primary vertex. If the χ^2 of this vertex combination is small ($\leqslant 4$) then the given track is not used in the B_s^0 reconstruction.

To speed up the processing, additional requirements are placed on the D_s meson candidate before its decay vertex is created: the scalar sum of the p_{T} of the particles used to create it must be $> 1.8\,\text{GeV}/c$, the largest *distance of closest approach* (DOCA) of the particles with respect to the primary vertex must be larger

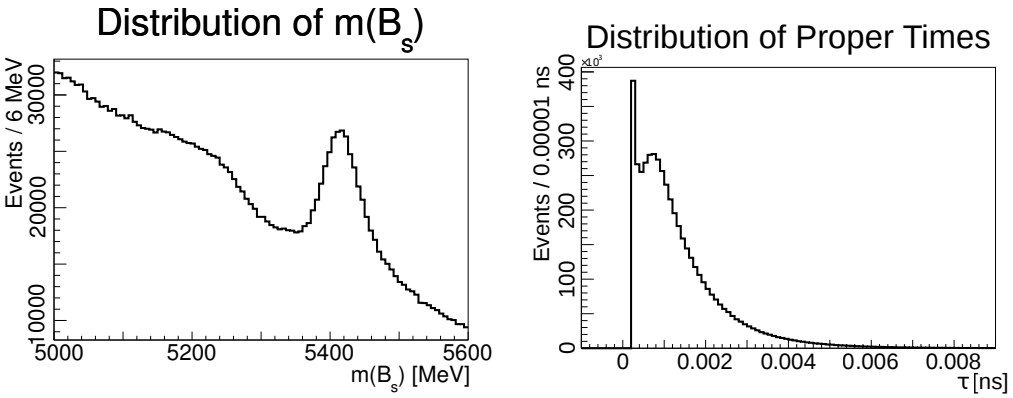


Figure 5.1: Distributions for $B_s^0 \to D_s K$ candidates in data after the stripping selection.

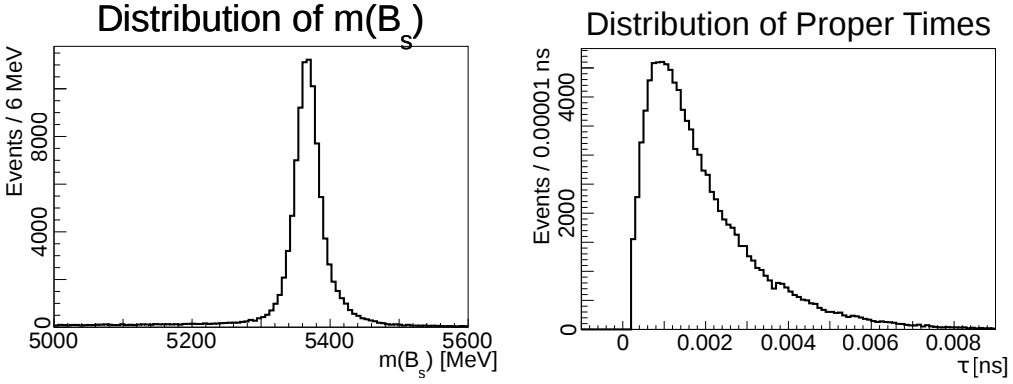


Figure 5.2: Distributions for simulated $B_s^0 \to D_s K$ decays after the stripping selection.

than 0.5 mm and the reconstructed invariant mass must be within $100\,\mathrm{MeV}/c^2$ of the nominal D^+ or D_s meson mass. After the vertex has been formed, final requirements of a vertex $\chi^2/\mathrm{ndof} < 10$ and that the vertex is well separated from the primary collision vertex are imposed.

After the initial pre-selection, remaining events are passed through a *bagged boosted decision tree* (BBDT) [45]. It is trained using the p_T of the B_s^0 meson candidate, the separation of its decay vertex from the primary vertex and a combination of the χ^2/ndof of the B_s^0 meson and D_s meson vertices. The BBDT response value is required to be > 0.05 to give the distributions shown in Figures 5.1 and 5.2.

5.7 Offline selection

The offline selection is run over the output of the stripping selection and is composed of a number of parts. First a boosted decision tree selection is used which is trained on kinematic and topological information. Then PID requirements are applied for the D_s daughters and bachelor pion and finally a set of vetoes for D, Λ_c and J/ψ decays are set.

5.7.1 Boosted decision tree training

At the core of the event selection process is a gradient boosted decision tree (BDTG) which is trained on $B_s^0 \to D_s\pi$, $D_s^+ \to K^+K^-\pi^+$ data. A BDTG is a binary tree classifier which involves making multiple yes or no decisions on an event, each based on a single variable until a certain stop condition is met. Two data sets are passed through the system to train it, one representative of the signal and one representative of any expected backgrounds. Depending on whether the majority of events ending up in a given leaf node are signal or background, all the events in that end up in that node are labelled as such. The boosting is performed by maintaining multiple trees at a time and combining them at the end to produce a continuous number. Thus, using the trained tree, each event in a data set can be assigned a value between 0 and 1 to be used as a signal discriminant. The BDTG implementation used is that from TMVA [46].

The data set is split into two equal-sized parts, one to be used for training the BDTG and the other to test its response. Each sample contains an equal amount of magnet-up and magnet-down data.

A sample of events to represent the combinatorial background in the BDTG training is taken from the upper sideband of the reconstructed B_s^0 meson, defined as $m(B_s^0) > 5445\,\mathrm{MeV}/c^2$. This data set is taken from the output of the stripping as defined above.

The signal training sample is extracted from $B_s^0 \to D_s\pi$, $D_s^+ \to K^+K^-\pi^+$ data after the stripping selection. To improve the performance of the selection, it is preferable to train on signal from data. In order to subtract the background events from the sample, the *sPlot* technique [47] is used. First, the events from the stripping selection have an additional pre-selection applied to them as given in Table 5.2 and then they are processed using the *sPlot* technique. The sPlot technique works by assigning a weight (called an *sWeight*) to each event in the sample describing how

Description	**Requirement**
Bachelor	$\mathrm{DLL}_{K\pi} < 0$
Both kaons	$\mathrm{DLL}_{K\pi} > 0$
D_s mass	$[1940, 1990]\,\mathrm{MeV}/c^2$
D^+ veto:	
$\mathrm{DLL}_{K\pi}$ of same charge K	> 10, or
D_s under D^+ hypothesis	below $1850\,\mathrm{MeV}/c^2$
Λ_c veto:	
p veto, same charge K	$\mathrm{DLL}_{K\pi}$-$\mathrm{DLL}_{p\pi} > 5$, or
D_s under Λ_c hypothesis	not in $[2250, 2320]\,\mathrm{MeV}/c^2$

Table 5.2: Additional pre-selection requirements applied to $B_s^0 \to D_s\pi$, $D_s^+ \to K^+K^-\pi^+$ used in BDTG optimisation.

signal-like it is. Based on a maximum-likelihood fit, the sWeights are given by

$$W_n(y_e) = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}, \tag{5.1}$$

where V_{nj} is the covariance matrix resulting from the likelihood maximisation, N_k is the number of events expected on the average for the k^{th} component, $f_i(y_e)$ is the value of the PDFs of the discriminating variables y for the i^{th} component and for event e, N_s denotes the number of components. In this case, the discriminating variable is the B_s^0 mass and there are two components: a Gaussian function for the signal and an exponential for the background.

The fit is performed on the full mass range of data after the selections given in Table 5.2. It is done in two steps, first with all shape and yield parameters floating to extract the shape parameters. Then, the fit is performed again with only the yields floating to avoid the correlations between the shapes and yields entering the covariance matrix. The outcome of this is an sWeight for each event in the sample which can be used in the BDTG training to subtract the background.

From all the possible kinematic and topological variables that could be used to train the BDTG, only a subset is used. They are chosen by utilising TMVA to calculate their importance as a discrimination variable in the tree. Those which are used in the final selection are given in Table 5.3 and the BDTG response distributions for the training and test samples are given in Figure 5.3.

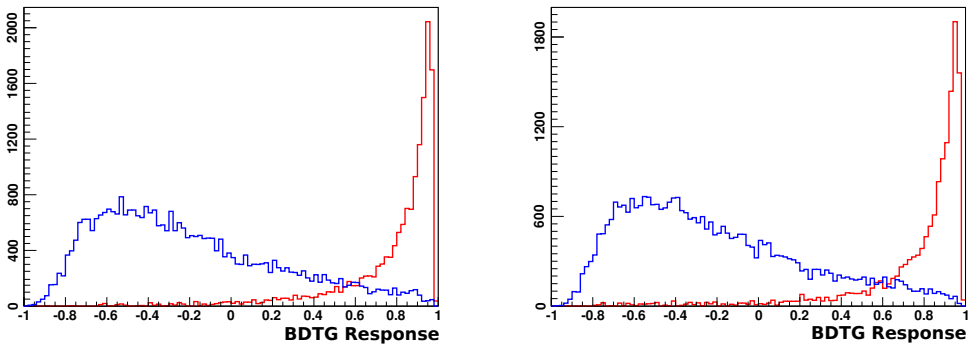


Figure 5.3: BDTG response distributions for training (left) and test (right) samples. The red histogram corresponds to signal weighted by the signal sWeights, while the blue shows background weighted by the combinatorial background sWeights.

5.7.2 Selection optimisation

The BDTG response requirement is chosen to give maximum signal significance which is defined as

$$S = \frac{N_{sig}}{\sqrt{N_{sig} + N_B}} \tag{5.2}$$

where N_{sig} and N_B are the signal and sum of all backgrounds respectively.

The yields for this significance are extracted from the data using the nominal mass fit described in Section 11.2. A full scan across the values of the BDTG response is performed with the requirement being scanned from 0.0 to 0.8 in 0.05 increments. The signal significance for these scans is shown in Figure 5.4 and Figure 5.5 shows the mass fit at two selected points in the scan.

The signal significance is observed to range between 25σ and 30σ and a final requirement of the BDTG response being larger than 0.5 is imposed.

5.8 Particle identification

Events which pass the BDTG selection are then further refined using PID requirements. These are defined using the logarithm of the likelihood of having a certain detector response given the hypothesis that the particle is a given particle, minus the same given the hypothesis of it being a pion This gives a set of *delta-log-likelihoods* (DLLs) which should discriminate between different particle types. Of most interest here is the $\mathrm{DLL}_{K\pi}$ which can be used to select between kaons and pions, though the $\mathrm{DLL}_{p\pi}$ is also used to reject protons.

It would be preferable to perform a significance scan across a range of $\mathrm{DLL}_{K\pi}$

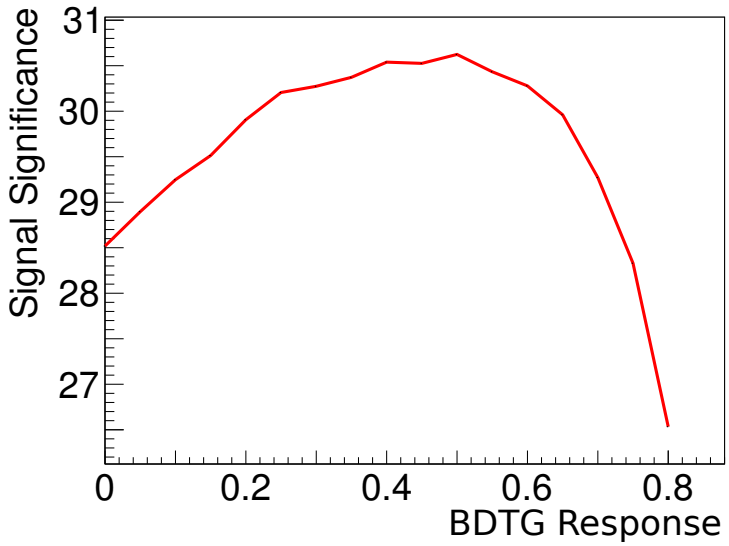


Figure 5.4: Significance scan for the BDTG response cut value.

(perhaps a range of -10 to 20) but since changing the particle identification requirement would require recomputing the shapes of the backgrounds used in the mass fit, only two values of $\mathrm{DLL}_{K\pi}$ are tested. Testing at $\mathrm{DLL}_{K\pi} > 5$ and $\mathrm{DLL}_{K\pi} > 10$ gives significances of the signal yield of 29.5σ and 30.5σ respectively. More importantly, the tighter requirement greatly reduces the cross-feed contribution of $B_s^0 \to D_s\pi$ under the signal peak. Therefore, a particle identification requirement of $\mathrm{DLL}_{K\pi} > 10$ on the bachelor track is applied.

The requirements placed on the various final-state particles are given in Table 5.4.

5.8.1 Background vetoes

A number of specific vetoes are needed in order to reduce the large number of D^+ mesons coming from $B^0 \to D\pi$ and similar decays. These also reject contributions from long-lived Λ_b decays such as $\Lambda_b \to \Lambda_c\pi$ where the proton from $\Lambda_c \to pK^-\pi^+$ is misidentified as a kaon. $J/\psi \to \mu^+\mu^-$ decays are vetoed in the case where both muons are misidentified as pions. Finally, D^0 decays such as $D^0 \to K^+K^-$ are also vetoed. All the vetoes used are given in Table 5.5.

The distribution of the key variables for both real and simulated data after all the selection requirements are shown in Figure 5.6.

Variable
B_s:
Cosine of the angle between the momentum vector and the line from the primary vertex and the B_s^0 decay vertex
The χ^2 distances of the track the primary vertices
Radial flight distance
Vertex χ^2 divided by ndof
Lifetime vertex χ^2 divided by ndof
Bachelor:
The minimum of the χ^2 distances of the track from any of the primary vertices
p_T
$\cos(\theta)$
D_s:
Cosine of the angle between the momentum vector and the line from the D_s origin vertex and the D_s decay vertex
Cosine of the angle between the momentum vector and the line from the primary vertex and the D_s decay vertex
The minimum of the χ^2 distances of the track from any of the primary vertices
Radial flight distance
Vertex χ^2 divided by ndof
D_s children:
Minimum p_T
The minimum of the χ^2 distances of the track from any of the primary vertices
Bachelor and D_s children
Maximum track ghost probability

Table 5.3: Input variables to the BDTG.

Applied to	**Description**	**Requirement**
$B_s^0 \to D_s\pi$	Bachelor pion	$\mathrm{DLL}_{K\pi} > 10$
$D_s^+ \to K^+K^-\pi^+$	Both kaons	$\mathrm{DLL}_{K\pi} > 0$
$D_s^+ \to K^+\pi^-\pi^+$	Kaon	$\mathrm{DLL}_{K\pi} > 10$
	Both pions	$\mathrm{DLL}_{K\pi} < 5$
$D_s^+ \to \pi^+\pi^-\pi^+$	All pions	$\mathrm{DLL}_{K\pi} < 10$
	All pions	$\mathrm{DLL}_{p\pi} < 10$

Table 5.4: PID selection requirements.

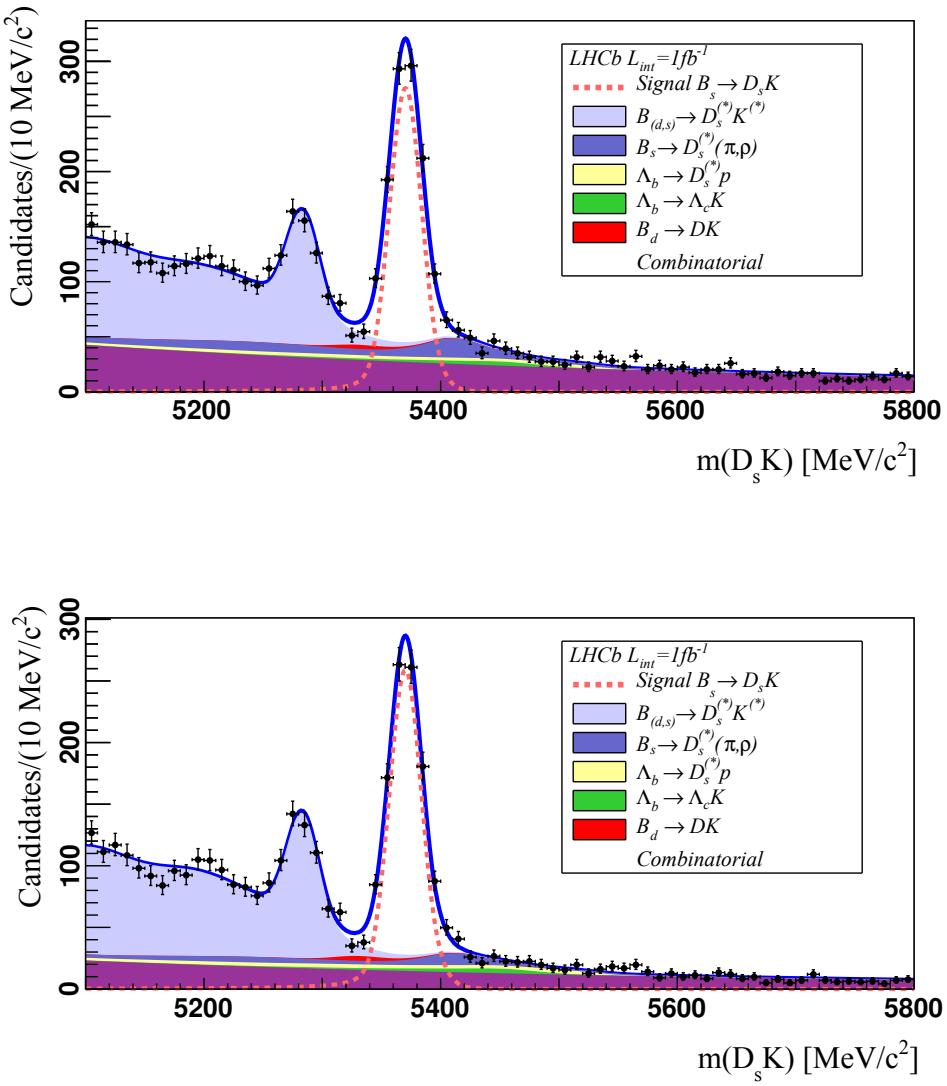


Figure 5.5: Mass fits to $B_s^0 \to D_s K$ data candidates for two selection working points. Top: BDTG response > 0.25 Bottom: BDTG response > 0.50. These fits are fully described later in Chapter 7.

Applied to	Description	Requirement
$D_s^+ \to K^+K^-\pi^+$	D^0 veto:	
	$m(K^+K^-)$	$< 1840\,\mathrm{MeV}/c^2$
	D^+ veto:	
	$\mathrm{DLL}_{K\pi}$ of same charge K	> 10, or
	$\lvert m(K^+K^-) - 1020\,\mathrm{MeV}/c^2 \rvert$	$< 10\,\mathrm{MeV}/c^2$, or
	D_s under D^+ hypothesis	not in $[1840, 1900]\,\mathrm{MeV}/c^2$
	Λ_c veto:	
	p veto, same charge K	$\mathrm{DLL}_{K\pi} - \mathrm{DLL}_{p\pi} > 5$, or
	D_s under Λ_c hypothesis	not in $[2250, 2320]\,\mathrm{MeV}/c^2$
$D_s^+ \to K^+\pi^-\pi^+$	D^0 veto:	
	$m(K^+\pi^-)$	$< 1750\,\mathrm{MeV}/c^2$
	Λ_c veto:	
	p veto on kaon	$\mathrm{DLL}_{K\pi} - \mathrm{DLL}_{p\pi} > 0$, or
	D_s under Λ_c hypothesis	not in $[2250, 2320]\,\mathrm{MeV}/c^2$
	J/ψ veto:	
	Both $m(\pi^+\pi^-)$	$\lvert m - m(J/\psi) \rvert > 30\,\mathrm{MeV}/c^2$
$D_s^+ \to \pi^+\pi^-\pi^+$	D^0 veto:	
	Both $m(\pi^+\pi^-)$	$< 1700\,\mathrm{MeV}/c^2$

Table 5.5: Vetoes applied on D_s meson candidates.

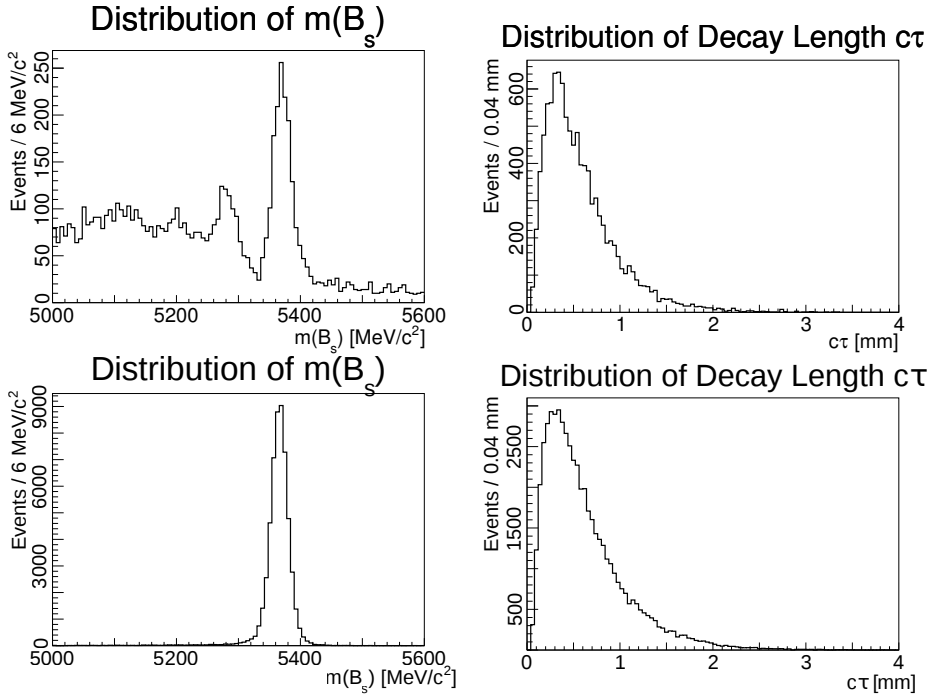


Figure 5.6: Distributions for real (top) and simulated (bottom) $B_s^0 \to D_s K$ data after the offline selection.

6

Flavour tagging

The flavour composition of the B_s^0 meson at the time of its creation is measured using a collection of algorithms. Each algorithm is known as a *tagger* and their results are combined into a final decision. The taggers used in the LHCb experiment are described fully in [48, 49].

In the proton-proton interaction, quarks are created in pairs. In the case of a $b\bar{b}$ pair being created, one of the quarks will hadronise to form a B meson (perhaps via a B^* or B^{**}) which will then decay as our measured signal. The other will also form a B hadron of some kind (labelled as the *opposite-side*), the flavour content of which will be directly related to the flavour of the 'signal B meson'. It is also possible that the partner to the signal B meson's other valence quark (an s in our case) will also hadronise and form a kaon or a pion (known as *associated production*) whose charge will be correlated with the signal B meson's flavour. The taggers fall in to two main categories: *same-side (SS) taggers* which extract information on the B meson flavour from kaons or pions emitted from the signal B meson's intermediate B^* or B^{**} state and *opposite-side (OS) taggers* which base their decision on the decay of the opposite-side B hadron. The layout of the taggers is shown in Figure 6.1. In the analysis shown in the thesis, only the opposite-side taggers are used.

There are four individual opposite-side taggers in total, each using a different facet of the decay to give an estimate of the initial flavour of the B_s^0 meson. There are two which are based on the direct semi-leptonic decay of the opposite-side hadron. In this case, the hadron decays via $b \to cW^-$ with $W^- \to \mu^- \overline{\nu}_\mu$ (or $W^- \to e^- \overline{\nu}_e$) and the muon or electron's charge is linked to the flavour of the opposite-side b hadron which in turn is related to the flavour of the signal B meson. Another tagger based on the decay of the b hadron is the *opposite-side kaon tagger* in which a kaon created

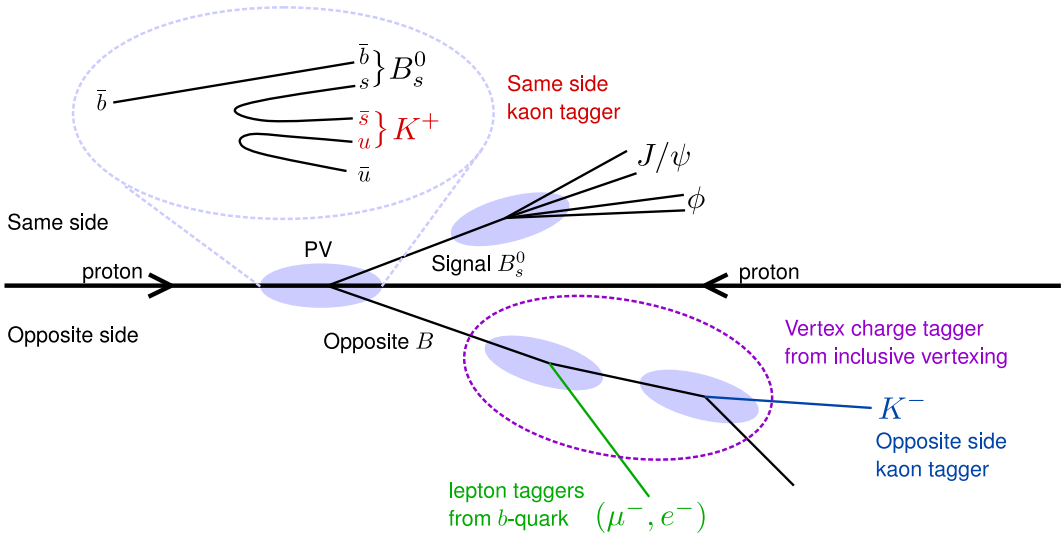


Figure 6.1: Schematic of the various taggers used in LHCb.

via a $b \to c \to s$ decay is detected and again the charge gives the flavour of the signal B meson. Finally, there is a tagger which collects together all the tracks which could be used to form an opposite-side decay chain and sums their electric charge.

Each tagger has certain sources of uncertainty associated with it. Some are irreducible and some are due to inefficiencies in the detector or selection of particles. For example, there is a chance that if the opposite-side b hadron is a B^0 or a B_s^0 then it might oscillate before decaying, causing the tagging algorithm to give the wrong answer. There is also the chance of simply selecting the wrong muon or electron from the semi-leptonic decay and instead picking up a background particle which has no relation to the system of interest at all. The ability for a tagger to get the wrong answer is called the *mistag probability* (or *mistag fraction* when talking about an ensemble of events), ω. Each tagger will have a different average value of ω.

Of course, there are also situations when a particular tagger simply can't give an answer. For example, the opposite-side b hadron might not decay semi-leptonically in which case there is no electron or muon to detect. Some taggers may even be mutually exclusive, for instance it is impossible to create both a pion and a kaon via associated production with the signal B meson in a single event. This effect is accounted for as a *tagging efficiency*, ε_{tag}, which again will potentially be different for each individual tagger.

The CP violating parameters being measured in the $B_s^0 \to D_s K$ fit are proportional to the dilution, D, which is related to the mistag probability ω as

$$D = 1 - 2\omega \tag{6.1}$$

and the statistical precision of *CP* parameters is directly related to this by an effective efficiency, ε_{eff}, given by

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2 \tag{6.2}$$

which is derived using the propagation of uncertainty [48].

6.1 Calibration

For each event, each tagger provides an initial estimate of the probability that it gave the wrong answer. This probability, η, is calculated using a neural network based on event properties (such as total number of tracks) as well as the kinematic and geometrical properties of the particles used to provide the decision. The neural network is trained on simulated data and so must be calibrated on real data to give a reliable result. This calibration is performed on a channel where the signal B meson will not oscillate (such as $B^+ \to J/\psi K^+$) so that the true flavour is known and can be compared to the estimated flavour. All the events in the sample are collected in bins of η and for each bin a mistag fraction, ω, is calculated based on how many events were tagged correctly. The resulting values of ω are plotted against η and its dependence is fitted with a linear function such as

$$\omega(\eta) = p_0 + p_1 \times (\eta - \langle\eta\rangle), \tag{6.3}$$

where p_0 and p_1 are fitted parameters and $\langle\eta\rangle$ is the average η across the whole data set. This function, along with the values of p_0 and p_1 can then be applied to any estimated η when performing an analysis.

It is possible to estimate the systematic uncertainty on the calibration parameters by performing the calibration under a number of varying conditions. By splitting the calibration data sample by magnet polarity and by the flavour of the signal meson and fitting each sample independently, a variation can be measured. Additionally, altering the model used to fit the data distribution can have an effect on the calibration. Adding these effects together in quadrature gives an estimate of the total systematic uncertainty.

6.2 Combination

Since each individual tagger is potentially giving an incorrect answer or even no answer for a given event, improved sensitivity can be gained by combining the answers from multiple, calibrated taggers. The combined probability, $\mathcal{P}(b)$ of the B

meson containing a b or $\overline{b}$ quark is given by

$$\mathcal{P}(b) = \frac{p(b)}{p(b) + p(\overline{b})}, \qquad \mathcal{P}(\overline{b}) = 1 - \mathcal{P}(b). \tag{6.4}$$

where $p(b)$ is the probability to have a b-tagged meson as a response from the combined tagger. It is defined as

$$p(b) = \prod_i \left(\frac{1 + d_i}{2} - d_i(1 - \omega_i) \right), \tag{6.5}$$

where i labels each tagger in the combination and d_i is the decision of the tagger such that $d_i = 1$ means that the signal meson contains a $\overline{b}$ and $d_i = -1$ means that the signal meson contains a b.

The final decision is made by comparing the two probabilities $\mathcal{P}(\overline{b})$ and $\mathcal{P}(b)$. If $\mathcal{P}(b) > \mathcal{P}(\overline{b})$ then the signal meson is tagged as containing a b with a mistag probability of $\eta^{comb} = \mathcal{P}(\overline{b})$ and conversely for the case where $\mathcal{P}(\overline{b}) > \mathcal{P}(b)$.

This estimated mistag probability, η^{comb}, does not take into account the correlations between the individual taggers and so it is necessary to recalibrate the combined tagger once again against data using the same method as before to give a final value of ω^{comb}. For the combination of opposite-side tagging algorithms used in this analysis, the calibration coefficients are measured on $B^+ \to J/\psi K^+$ data and are calculated to be $p_0 = 0.392 \pm 0.002 \pm 0.009$, $p_1 = 1.035 \pm 0.021 \pm 0.012$ and $\langle \eta \rangle = 0.391$ where the uncertainties are statistical followed by systematic. The fit is shown in Figure 6.2.

6.3 Per-event mistag probability

There are two ways to use the tagging information. First is to simply use an average mistag probability as a single parameter for every event in the fit. However, in order to best exploit the tagging information provided by the tagging algorithms, a *per-event* mistag probability is assigned. This mistag probability, ω, is given by the calibrated estimate of the mistag probability provided by the tagging algorithm.

In tests performed on $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ improvements in ε_{eff} of up to 60% were measured [48]. As such, for this analysis, per-event mistag probabilities are used.

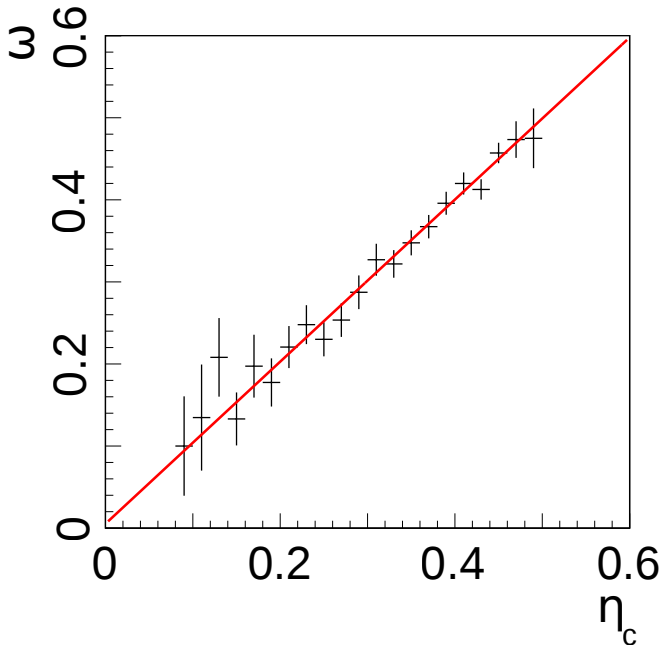


Figure 6.2: Fit for the calibration of the combined flavour taggers.

6.4 Possible optimisation of the tagger

The standard taggers in LHCb make use of simple, rectangular cuts in order to select the candidates used to determine the flavour tag. It has been seen in many physics analyses that by replacing rectangular cuts with more complex selection methods such as boosted decision trees or neural networks, it is possible to increase the event selection efficiency while keeping backgrounds low.

As part of my service work throughout my Ph.D., I performed a study into the possible improvements that could be made by using NeuroBayes [50], a neural network training package. The opposite-side muon tagger and the same-side pion taggers were the subject of the study, with the aim of improving the ε_{eff} compared with the standard LHCb taggers. The results of this work was published as an internal LHCb note as Ref [51].

For each of the two taggers, a separate neural network was trained in order to tune the tagger to the specific requirements of the candidate selection. Each neural network was trained on simulated data of the decay $B^+ \to J/\psi K^+$. In order to test the dependence of the network on the training channel, a neural network for the opposite-side muon tagger was also trained using $B \to DX$ data where X can be a pion or a kaon.

Tagger	ε_{tag} (%)	ω (%)	ε_{eff} (%)
Standard	5.15	29.3 ± 1.9	0.88 ± 0.11
$B \to DX$ Neural Network	12.44	36.0 ± 1.2	1.00 ± 0.12
$B^+ \to J/\psi K^+$ Neural Network	9.91	34.0 ± 1.2	0.99 ± 0.10

Table 6.1: The results of the neural network-based tagger compared to the existing opposite-side muon tagger

Tagger	ε_{tag} (%)	ω (%)	ε_{eff} (%)
Standard	10.34	39.4 ± 1.2	0.47 ± 0.08
Neural Network	89.87	46.0 ± 1.0	0.58 ± 0.09

Table 6.2: The results of the neural network-based tagger compared to the existing same-side pion tagger

All three resulting networks were tested on $B^+ \to J/\psi K^+$ data and the results compared with the charge of the B meson. The results for the muon tagger are shown in Table 6.1. The two neural network results differ only by the sample used to train the network. The same-side pion performance is shown in Table 6.2. This neural network was both trained and tested on the $B^+ \to J/\psi K^+$ channel.

7

Mass fit

In order to extract yields for the background contributions in the decay-time fit, a fit is performed on the distribution of reconstructed B_s^0 mass. The most important part of this is understanding the shapes of the backgrounds and the signal.

The signal is described with a double Crystal Ball function with tails pointing in opposite directions. The shapes of these functions are fixed using a fit to simulated data which has had the full selection applied to it. All shape parameters are floated freely in the fit and the results are given in Table 7.1. The resultant PDF and the simulated data it was fitted to are shown in Figure 7.1. In the final mass fit to data, the tail parameters remain fixed but the widths and means are floated.

Parameter	**Fitted value**
μ_{DOWN}	5366.5 ± 0.09 MeV/c^2
μ_{UP}	5366.6 ± 0.09 MeV/c^2
σ_1	10.88 ± 0.11 MeV/c^2
σ_2	15.71 ± 0.09 MeV/c^2
α_1	1.81 ± 0.01 MeV/c^2
α_2	-1.82 ± 0.03 MeV/c^2
n_1	1.38 ± 0.02
n_2	8.86 ± 0.09
f	0.47 ± 0.01

Table 7.1: Parameters for the sum of the two Crystal Ball functions describing the signal shapes of $B_s^0 \to D_s K$, obtained from simulated data.

The first background to consider is the combinatorial background. This was modelled as an exponential function with slope parameters fitted to data in the

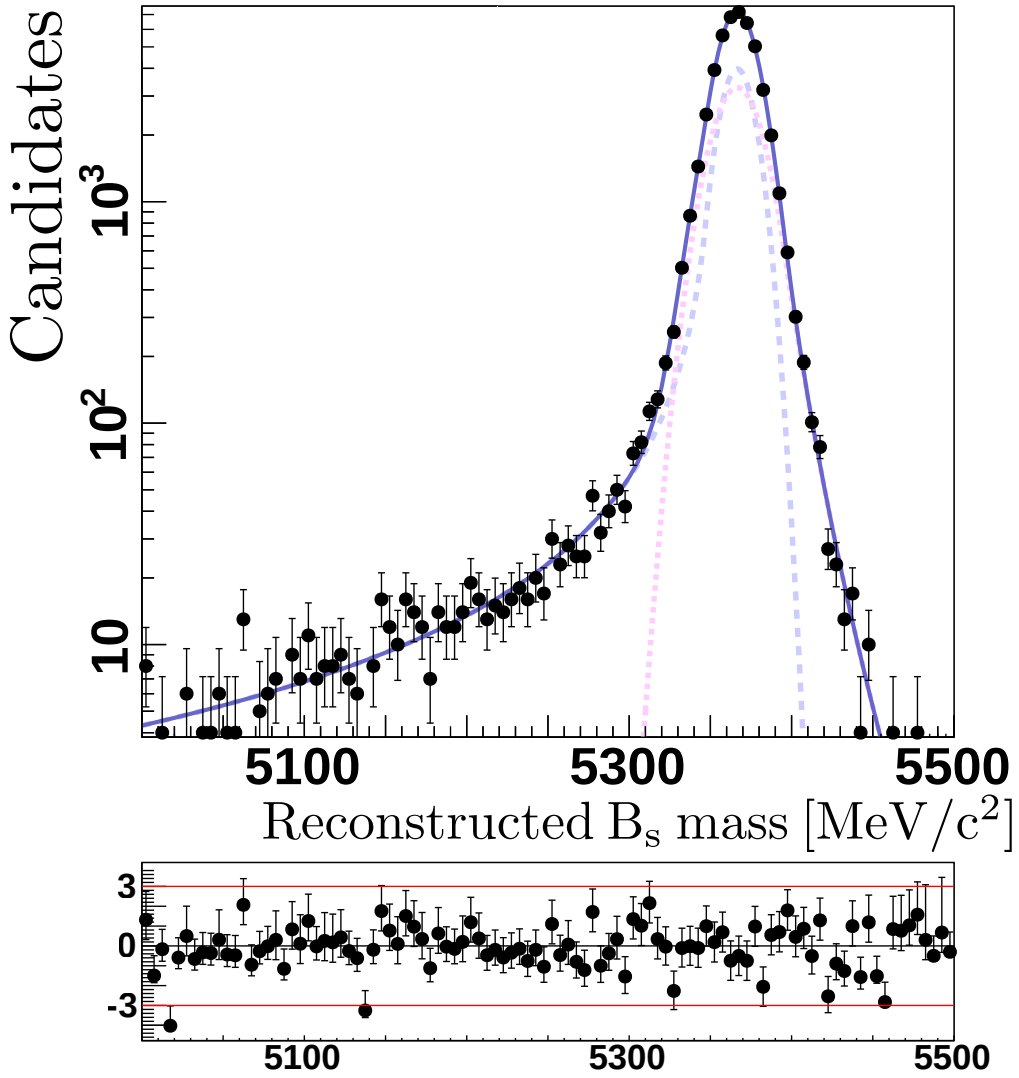


Figure 7.1: Signal mass shapes of $B_s^0 \to D_s K$ evaluated on simulated data. The solid lines correspond to the fit to the double Crystal Ball function, while the dashed lines correspond to the individual Crystal Ball components. The bottom plot show the deviation of the data from the fit line based on statistical uncertainty.

range $m(D_s) = (1868, 1948) \cup (1990, 2068)\,\text{MeV}/c^2$ and $m(B_s^0) > 5600\,\text{MeV}/c^2$ and then used in the final fit over the full mass range. The fitted slope parameters for each of the D_s modes are given in Table 7.2.

D_s **mode**	**Slope parameter** ($c^2/\,\text{MeV}$)
$D_s^+ \to K^+K^-\pi^+$	-1.58 ± 0.19
$D_s^+ \to K^+\pi^-\pi^+$	-1.09 ± 0.25
$D_s^+ \to \pi^+\pi^-\pi^+$	-0.99 ± 0.21

Table 7.2: The fitted values (in units of 10^{-3}) of the slope parameter of an exponential function describing the combinatorial background.

Further to this, there are a number of fully- and partially-reconstructed backgrounds. All these backgrounds are modelled using simulated data, forced to decay to the particular final state. The two fully-reconstructed backgrounds that factor in the fit are $B_s^0 \to D_s\pi$ and $B^0 \to D_s^- K^+$ where the former is fitted with a sum of Gaussian kernels and the latter is fitted with a double Crystal Ball function. There are a large number of partially-reconstructed backgrounds: $B^0 \to D^- K^+$, $B_s^0 \to D_s^- K^{*+}$, $B_s^0 \to D_s^{*-} K^+$, $B_s^0 \to D_s^{*-} K^{*+}$, $B_s^0 \to D_s\rho$, $B_s^0 \to D_s^*\rho$, $B_s^0 \to D_s^*\pi$, $\Lambda_b \to D_s^- p$, $\Lambda_b \to D_s^{*-} p$ and $\Lambda_b \to \Lambda_c K^-$. The shapes for each of these is taken from simulated data and corrected for the distortion caused by the momentum-dependent particle identification efficiency. Each background distribution is fitted with a sum of Gaussian kernels.

The yields of $B^0 \to D^- K^+$, $\Lambda_b \to D_s^- p$, $\Lambda_b \to D_s^{*-} p$ and $\Lambda_b \to \Lambda_c K^-$ are fixed in the mass fit. The $B^0 \to D^- K^+$ yield is fixed based on the yield of $B^0 \to D\pi$ calculated in Section 11.1.1 scaled by a factor 1/15 based on the relative branching fractions of $B_s^0 \to D_s\pi$ and $B_s^0 \to D_s K$ [52] and correcting for the relative particle identification efficiencies of the two $\text{DLL}_{K\pi} < 0$ and $\text{DLL}_{K\pi} > 10$ requirements. Again for $\Lambda_b \to \Lambda_c K^-$, the yield is based on the fitted yield from the fit to $B_s^0 \to D_s\pi$ and scaled by the same factor 1/15 as for $B^0 \to D^- K^+$. For the final two modes, a sample of events is created where the mass hypothesis used for the bachelor is changed from that of a kaon to a proton. The sample is then fitted to extract the shapes of the Λ_b backgrounds.

The remaining backgrounds are collected into two groups:

Group 1: $B^0 \to D_s^- K^+$, $B_s^0 \to D_s^- K^{*+}$, $B_s^0 \to D_s^{*-} K^+$, $B_s^0 \to D_s^{*-} K^{*+}$

Group 2: $B_s^0 \to D_s\pi$, $B_s^0 \to D_s^*\pi$, $B_s^0 \to D_s\rho$, $B_s^0 \to D_s^*\rho$

Each group is combined into a single PDF, defined as

$$\begin{aligned} f_{11}\mathrm{PDF}_{B^0\to D_s^- K^+} &+ (1-f_{11})[f_{12}\mathrm{PDF}_{B_s^0\to D_s^- K^{*+}} \\ &+ (1-f_{12})(f_{13}\mathrm{PDF}_{B_s^0\to D_s^{*-} K^+} \\ &+ (1-f_{13})\mathrm{PDF}_{B_s^0\to D_s^{*-} K^{*+}})] \end{aligned} \tag{7.1}$$

for group 1 and

$$\begin{aligned} f_{21}\mathrm{PDF}_{B_s^0\to D_s\pi} &+ f_{22}\mathrm{PDF}_{B_s^0\to D_s^*\pi} \\ &+ f_{23}\mathrm{PDF}_{B_s^0\to D_s\rho} \\ &+ (1-f_{21}-f_{22}-f_{23})\mathrm{PDF}_{B_s^0\to D_s^*\rho} \end{aligned} \tag{7.2}$$

for group 2, where $\mathrm{PDF}_{\mathrm{mode}}$ is the PDF for a given mode calculated previously and f_{NM} are relative yield fractions. The fractions, f_{2N}, for group 2 are fixed based on the yields extracted from the mass fit to $B_s^0 \to D_s\pi$ and scaled appropriately based on the particle identification efficiency. The values obtained are $f_{21} = 0.428$, $f_{22} = 0.472$, and $f_{23} = 0.052$.

The results of the final fit are plotted in Figure 7.2 and the signal shape parameters and event yields are given in Table 7.3.

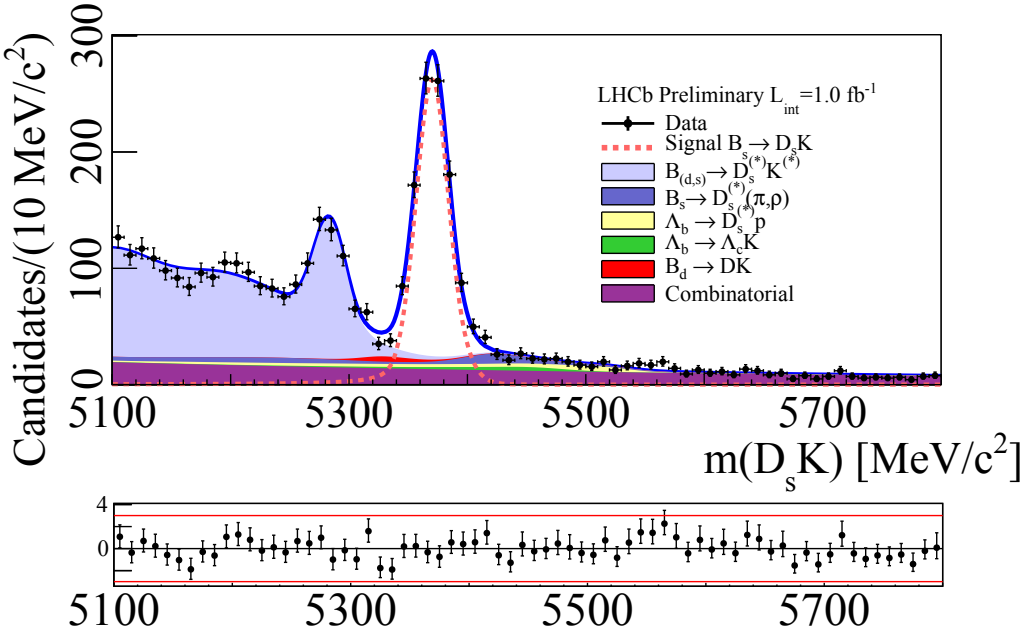


Figure 7.2: Result of the simultaneous mass fit to the $B_s^0 \to D_s K$ candidates. The pull distributions are shown in the lower part of the figure.

Parameter	**Fitted value**
Mean	$5370.30 \pm 0.47\,\mathrm{MeV}/c^2$
σ_1	$17.43 \pm 1.70\,\mathrm{MeV}/c^2$
σ_2	$12.34 \pm 0.86\,\mathrm{MeV}/c^2$
f_{11}	0.206 ± 0.022
f_{12}	0.374 ± 0.190
f_{13}	0.744 ± 0.106
$D_s^+ \to K^+K^-\pi^+$ magnet up	
$N_{B_s^0 \to D_s K}$	430 ± 25
N_{Comb}	303 ± 35
N_{Group1}	844 ± 37
N_{Group2}	51 ± 36
$D_s^+ \to K^+K^-\pi^+$ magnet down	
$N_{B_s^0 \to D_s K}$	626 ± 30
N_{Comb}	473 ± 44
N_{Group1}	1024 ± 45
N_{Group2}	100 ± 47
$D_s^+ \to K^+\pi^-\pi^+$ magnet up	
$N_{B_s^0 \to D_s K}$	51 ± 8
N_{Comb}	83 ± 17
N_{Group1}	69 ± 13
N_{Group2}	4 ± 20
$D_s^+ \to K^+\pi^-\pi^+$ magnet down	
$N_{B_s^0 \to D_s K}$	47 ± 8
N_{Comb}	115 ± 20
N_{Group1}	90 ± 15
N_{Group2}	21 ± 17
$D_s^+ \to \pi^+\pi^-\pi^+$ magnet up	
$N_{B_s^0 \to D_s K}$	101 ± 13
N_{Comb}	105 ± 18
N_{Group1}	184 ± 17
N_{Group2}	5 ± 14
$D_s^+ \to \pi^+\pi^-\pi^+$ magnet down	
$N_{B_s^0 \to D_s K}$	135 ± 14
N_{Comb}	159 ± 24
N_{Group1}	185 ± 21
N_{Group2}	61 ± 24

Table 7.3: Fitted values of parameters for $B_s^0 \to D_s K$ signal mass fit. The N_i are the yield of the signal and background contributions. Mean and sigmas are the parameters of double Crystal Ball function used to describe the signal. The parameters f_i are fractions between modes in Group 1 backgrounds: $B^0 \to D_s^- K^+$, $B_s^0 \to D_s^- K^{*+}$, $B_s^0 \to D_s^{*-} K^+$ and $B_s^0 \to D_s^{*-} K^{*+}$.

8

Decay-time fit

A conventional fit, such as the one used to fit the mass distribution uses a series of model PDFs for each expected signal and background component. They are combined together and fitted to the data using a maximum-likelihood method. However, recently an alternative method of performing time fits has started to become more common. It is referred to as the *sFit* and is based on the method given in Ref [53]. The first step of the process is identical to the conventional method discussed so far (which in contrast is referred to as the *cFit*). A fit to the mass distribution (with full descriptions of the backgrounds) is performed and two PDFs are extracted: one for the signal and another for the combination of the backgrounds. From this, a signal *sWeight* is calculated for each event, given by

$$W_s(y) = \frac{V_{ss}F_s(y) + V_{sb}F_b(y)}{N_sF_s(y) + N_bF_b(y)}, \tag{8.1}$$

where y is the variable over which the distribution is being fitted (in this case the B_s^0 reconstructed mass), N_s and N_b are the number of signal and background events, $F_s(y)$ and $F_b(y)$ are the distributions of y for the signal and background respectively and the matrix V is given by inverting

$$V_{ij}^{-1} = \sum_{e=1}^{N} \frac{F_i(y_e)F_j(y_e)}{(N_sF_s(y_e) + N_bF_b(y_e))^2}. \tag{8.2}$$

Using these weights, the time distribution of the events is plotted but with each event weighted by its signal weight. The resultant distribution gives the time distribution of the signal mode, having effectively removed the background contribution. This allows a simple fit to be performed on the time distribution without having to model

Parameter	Value
σ_1	$29.48 \pm 0.027\,\text{fs}^{-1}$
σ_2	$58.64 \pm 0.063\,\text{fs}^{-1}$
σ_3	$181.7 \pm 4.9\,\text{fs}^{-1}$
μ	$1.49 \pm 0.14\,\text{fs}^{-1}$
f_1	0.595 ± 0.011
f_2	0.386 ± 0.011

Table 8.1: Fitted parameters for the time resolution model.

the backgrounds. A full conventional fit is also performed as a cross-check during the development of the sFit but is not used to produce final results.

Due to the nature of the fit, it is not necessary to parameterise the shapes of the backgrounds since they will have been removed from the distribution. As such, the fit model is simply a signal decay function given in Equation 2.25 convolved with a resolution function and multiplied by an acceptance function.

8.1 Decay-time resolution

Any measurable oscillation is diluted by the finite decay-time resolution of the detector as well as the measurement being potentially biased. As such, it is important to account for the time resolution accurately. The time resolution is modelled by the sum of three Gaussian functions with a common mean, μ, but differing widths, σ_i, and is parameterised as

$$R(t) = f_1 G(t;\mu,\sigma_1) + f_2 G(t;\mu,\sigma_2) + (1 - f_1 - f_2) G(t;\mu,\sigma_3), \tag{8.3}$$

where f_i are the relative fractions of the first two components. The parameters for the model are calculated by fitting a sample of simulated $B_s^0 \to D_s\pi$ events giving the parameters shown in Table 8.1.

8.2 Decay-time acceptance

In addition to a finite time resolution, the events are also selected with a non-uniform acceptance which is a function of decay time. The LHCb triggers preferentially ignore events with a short decay time to avoid selecting prompt charm hadrons. The rate at which these short-lived particles are excluded can be parameterised by an envelope function by which the overall PDF is multiplied. In addition to

Parameter	Value
a	$1.420 \pm 0.204\,\mathrm{ps}^{-1}$
b	0.0230 ± 0.0364
n	1.810 ± 0.066
β	$0.0363 \pm 0.0118\,\mathrm{ps}^{-1}$

Table 8.2: Fitted parameters for the fit of the time acceptance model to simulated $B_s^0 \to D_s\pi$ data.

Parameter	a	b	n	β
a	1	0.992	0.914	−0.985
b	0.992	1	0.874	−0.973
n	0.914	0.874	1	−0.900
β	−0.985	−0.973	−0.900	1

Table 8.3: Correlations of parameters for the fit of the time acceptance model to simulated $B_s^0 \to D_s\pi$ data.

the deficit of short-lived particles, there is also an observed increasing reduction in particles as their decay-time increases. In order to model these features, the following power-law equation is used:

$$a_{trig}(t) = \begin{cases} 0 & \text{when } (at)^n - b < 0 \text{ or } t < 0.2 \text{ ps}, \\ (1 - \frac{1}{1+(at)^n - b}) \times (1 - \beta t) & \text{otherwise}, \end{cases} \tag{8.4}$$

where a models the steepness of the turn-on while the exponent n and the offset b model the position of the turn-on. The final parameter, β is used to represent the relative reduction in events at larger decay-times.

The acceptance function is calculated using a combination of simulated and real data using the formalism given in Equation 8.4. Acceptance parameters are calculated on $B_s^0 \to D_s\pi$ data and then calibrated based on a binned ratio between simulated $B_s^0 \to D_s\pi$ and $B_s^0 \to D_sK$ data. This method of calibration relies on the assumption that $B_s^0 \to D_s\pi$ is flavour-specific and that therefore the lifetime of the B_s^0 is well known.

The time acceptance is first fitted on simulated $B_s^0 \to D_s\pi$ events with a reconstructed decay-time larger than 0.2 ps and is shown in Figure 8.1. The obtained parameters are given in Table 8.2 with correlations given in Table 8.3. A second fit is then performed on simulated $B_s^0 \to D_sK$ events, the results of which are given in Table 8.4.

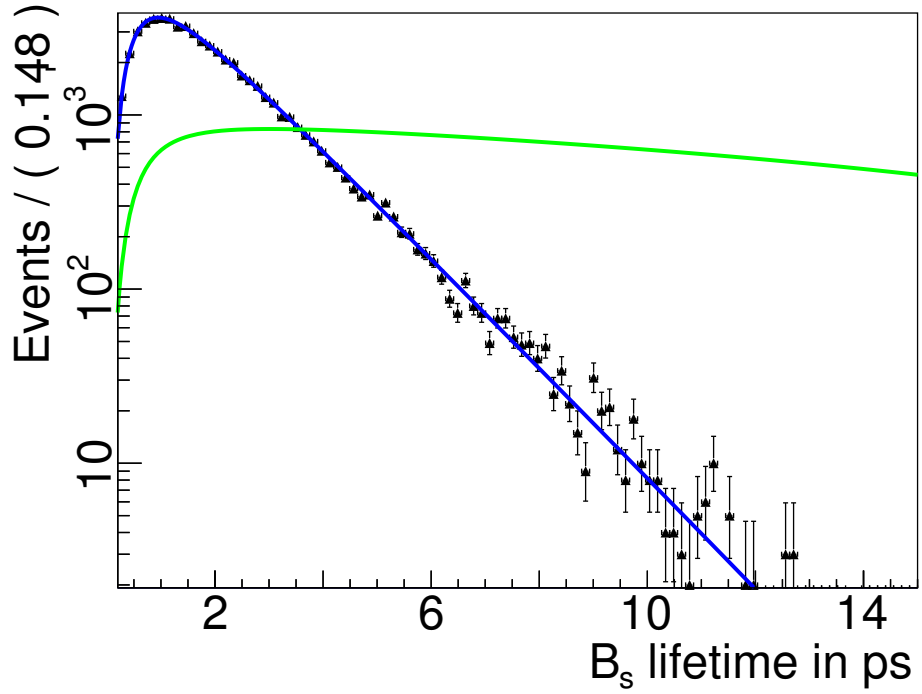


Figure 8.1: The acceptance function fitted to simulated $B_s^0 \to D_s\pi$ data. The acceptance function from Equation 8.4 is overlaid in green, scaled up for clarity.

Parameters	β	n	b	a	**Parameter value**
β	1	0.657	0.501	−0.784	$(3.94 \pm 0.28) \times 10^{-2}$ ps^{-1}
n	0.657	1	0.905	−0.574	2.00 ± 0.078
b	0.501	0.905	1	−0.276	$(-1.73 \pm 1.23) \times 10^{-2}$
a	−0.784	−0.574	−0.276	1	1.46 ± 0.025 ps^{-1}

Table 8.4: A summary of the parameter values and correlations for the acceptance function for $B_s^0 \to D_sK$.

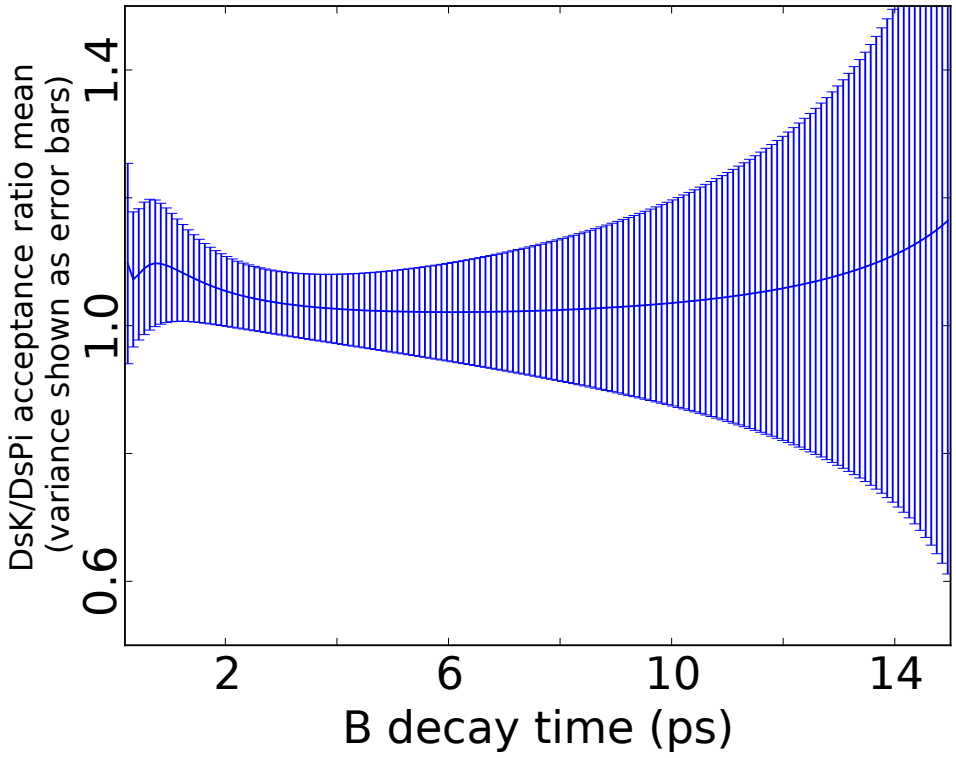


Figure 8.2: Ratio of acceptance parameters between $B_s^0 \to D_s K$ and $B_s^0 \to D_s \pi$ as a function of decay-time.

Based on these best-fit values and covariance matrices, 1000 sets of acceptance parameters are created to represent the range of uncertainties in the fit to simulated data. In 0.1 ps bins from 0.2 ps to 10 ps the mean and variance of the acceptance value is calculated for each of the decay modes. The ratio is then taken between $B_s^0 \to D_s K$ and $B_s^0 \to D_s \pi$ in each bin and is shown in Figure 8.2 to give a binned correction factor.

Finally, a fit to real $B_s^0 \to D_s \pi$ data is performed using the sFit method. In this fit the assumption is made that $C = 1$ and that $S_f = S_{\bar{f}} = D_f = D_{\bar{f}} = 0$. The physics parameters Γ_s, $\Delta\Gamma_s$ and Δm_s are also fixed to their nominal values as given in Table 12.2. The only variables floated in the fit are the four parameters of the acceptance model, the fitted values of which are shown in Table 8.5. The resultant PDF from this data fit is them scaled by the binned correction factor to give the acceptance function used in the final $B_s^0 \to D_s K$ fit.

Parameter	Fitted value
β	$0.0363 \pm 0.0068\,\mathrm{ps}^{-1}$
n	1.849 ± 0.071
b	0.0373 ± 0.0119
a	$1.215 \pm 0.053\,\mathrm{ps}^{-1}$

Table 8.5: Acceptance parameters as a result of the fit to the real data $B_s^0 \to D_s\pi$ time distribution.

Parameter	Fitted value
C	1.01 ± 0.50
S_f	-1.25 ± 0.56
$S_{\bar{f}}$	0.083 ± 0.680
D_f	-1.33 ± 0.60
$D_{\bar{f}}$	-0.81 ± 0.56

Table 8.6: CP observables fitted to the $B_s^0 \to D_sK$ decay-time distribution.

8.3 Fit to data

The distribution given by plotting the decay-time distribution of $B_s^0 \to D_sK$ candidates, weighted by their signal sWeights, is fitted using the maximum-likelihood method. The *CP* parameters C, S_f, $S_{\bar{f}}$, D_f and $D_{\bar{f}}$ are all freely floated between -3 and 3. The values of the fitted parameters are given in Table 8.6 along with their correlations in Table 8.7.

8.4 Systematic uncertainties

There are many sources of systematic uncertainty in the fit to the time-distribution of $B_s^0 \to D_sK$ data. How the various parameters are varied to estimate the mag-

Parameter	C	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
C	1	-0.155	-0.137	-0.110	0.174
D_f	-0.155	1	0.566	-0.057	-0.026
$D_{\bar{f}}$	-0.137	0.566	1	-0.025	-0.016
S_f	-0.110	-0.057	-0.025	1	-0.020
$S_{\bar{f}}$	0.174	-0.026	-0.016	-0.020	1

Table 8.7: Correlation matrix of the $B_s^0 \to D_sK$ *CP* parameter fit.

	Parameter				
	C	S_f	$S_{\bar{f}}$	D_f	$D_{\bar{f}}$
Statistical uncertainty	0.50	0.56	0.68	0.60	0.56
Systematic (σ_{stat})					
Decay-time bias	0.03	0.05	0.05	0.00	0.00
Decay-time resolution	0.11	0.08	0.09	0.00	0.00
Tagging calibration	0.23	0.17	0.16	0.00	0.00
Background yields	0.15	0.07	0.07	0.07	0.07
Physics parameters	0.15	0.22	0.20	0.40	0.42
Asymmetries	0.12	0.01	0.04	0.00	0.02
Momentum/Length Scale	0.00	0.00	0.00	0.00	0.00
k-Factors	0.27	0.27	0.27	0.08	0.08
Bias correction	0.03	0.03	0.03	0.03	0.03
Total systematic (σ_{stat})	0.46	0.50	0.35	0.43	0.46

Table 8.8: Total error budget for the decay-time fit. Systematic uncertainties are given as fractions of the statistical uncertainty. Systematic uncertainties are added in quadrature under the assumption that they are uncorrelated.

nitude of the uncertainties is explained later in Section 13.3. The calculated values for $B_s^0 \to D_s K$ are summarised in Table 8.8.

8.5 Conclusion

The final results for the $B_s^0 \to D_s K$ decay are

$$C = \quad 1.01 \pm 0.50 \pm 0.23, \tag{8.5}$$

$$S_f = -1.25 \pm 0.56 \pm 0.24, \tag{8.6}$$

$$S_{\bar{f}} = \quad 0.08 \pm 0.68 \pm 0.28, \tag{8.7}$$

$$D_f = -1.33 \pm 0.60 \pm 0.26, \tag{8.8}$$

$$D_{\bar{f}} = -0.81 \pm 0.56 \pm 0.26, \tag{8.9}$$

where the first uncertainties are statistical and the second uncertainties are systematic.

At this stage, no attempt is made to determine confidence intervals for physics parameters $r_{D_s K}$, Δ and $\gamma - 2\beta_s$ given by the relations in Equations 4.1–4.5. This is because a full treatment would require correct understanding of the statistical and systematic covariance matrices since it has been found from simula-

tion studies that they have a non-negligible effect on $\gamma - 2\beta_s$. Given future work on this matter, this analysis will provide a unique measurement of the value of γ, unavailable to other particle physics experiments.

Part III

$B_s^0 \to D_s\pi$ analysis

9

Introduction

An analysis of $B_s^0 \to D_s\pi$ decays to measure the level to which they are flavour-specific was performed. As discussed in Section 2.5, it has previously been assumed that the decay is flavour-specific but this had never been explicitly tested. The analysis builds upon the work done for the analysis of $B_s^0 \to D_sK$ and follows a similar template. The next few chapters will detail the method used to measure the *CP* parameters of the $B_s^0 \to D_s\pi$ decay, particularly focussing on differences from the method used for the analysis of $B_s^0 \to D_sK$.

The main differences occur when parameterising the backgrounds of the distribution of reconstructed B_s^0 mass and the method used for fitting the proper decay-time distribution which will be described in full in Chapter 12.

10

Data selection

10.1 Data sample

This analysis uses data from the 2011 run of LHCb. This comprises an integrated luminosity $\int \mathcal{L} = 1.0\ \text{fb}^{-1}$ of pp collisions recorded at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$.

10.2 Simulated data

Several samples of simulated data were created for the analysis, primarily for the use in event selection and background studies. In each sample, a B hadron is forced to decay to a specific final state as listed in Table 10.1 along with the number of events generated for each channel.

Sample		**Sample size**
$B_s^0 \to D_s\pi$	$D_s^+ \to K^+K^-\pi^+$	1052495
$B_s^0 \to D_s^*\pi$	$D_s^+ \to K^+K^-\pi^+$	524098
$B_s^0 \to D_s\rho$	$D_s^+ \to K^+K^-\pi^+$	2019391
$B_s^0 \to D_s^*\rho$	$D_s^+ \to K^+K^-\pi^+$	1019191
$\Lambda_b \to \Lambda_c\pi$	$\Lambda_c^+ \to pK^-\pi^+$	2033496
$B^0 \to D\rho$	$D^+ \to K^-\pi^+\pi^+$	2054494
$B^0 \to D^*\pi$	$D^* \to D\pi^0$, $D^+ \to K^-\pi^+\pi^+$	1046498
$B^0 \to D\pi$	$D^+ \to K^-\pi^+\pi^+$	1016198

Table 10.1: Simulated samples used during the analysis for selection and background studies.

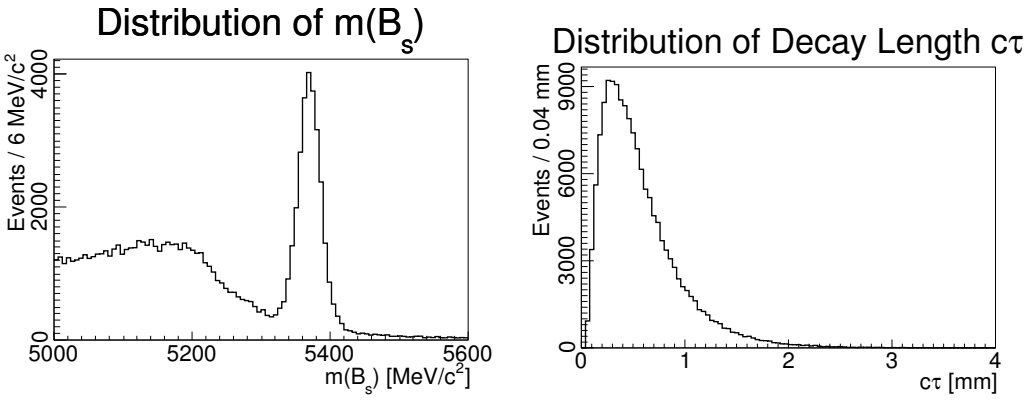


Figure 10.1: Distributions for $B_s^0 \to D_s\pi$ data after the offline selection.

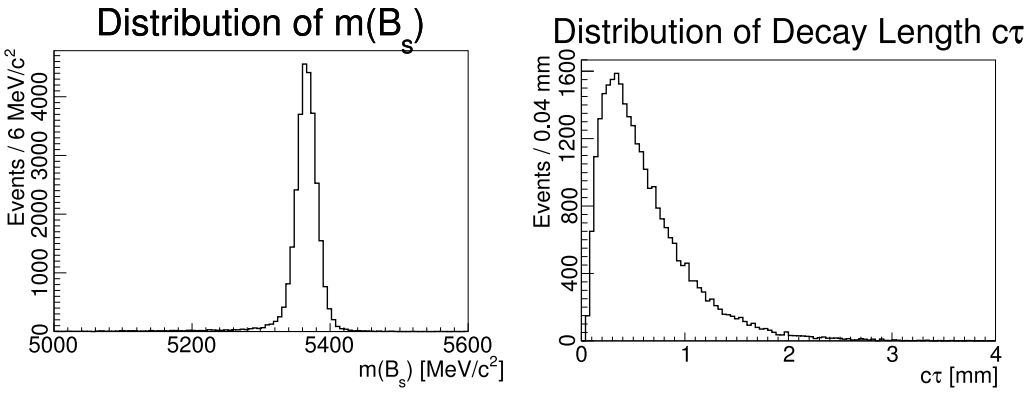


Figure 10.2: Distributions for simulated $B_s^0 \to D_s\pi$ after the offline selection.

10.3 Data selection

The reconstruction of the $B_s^0 \to D_s\pi$ signal events was performed in the same way as for the $B_s^0 \to D_sK$ analysis. The only difference is that rather than using a K mass hypothesis for the bachelor when combining the tracks, a π mass hypothesis is used.

The requirements used in order to select the $B_s^0 \to D_s\pi$ candidates are identical to those used in the $B_s^0 \to D_sK$ analysis with the exception of the particle identification requirement on the bachelor track. As both decays are kinematically similar, tuning of the BDT will perform well for selecting $B_s^0 \to D_s\pi$ since the BTD was explicitly not trained or optimised using any particle identification information about the bachelor. A relatively tight requirement of $\mathrm{DLL}_{K\pi} < 0$ is imposed on the bachelor track.

The distribution of the key variables for both real and simulated data after all the selection requirements are shown in Figures 10.1 and 10.2.

11

Data sample composition

There are a number of physics backgrounds which contaminate the data set. It is important to be able to describe and quantify these contributions in order to account for them when extracting *CP* parameters. In order to extract the *CP* parameters, a decay time fit is performed in which the yields of the signal and background channels are fixed to values extracted from a fit to the B_s^0 mass distribution.

11.1 Backgrounds to the mass fit

In order to extract yields by fitting reconstructed mass distribution of the B_s^0 meson, it is important to understand the shapes of possible background contributions.

The backgrounds fall into two main categories: partially reconstructed backgrounds, where one or more particles in the physical decay was not picked up by the reconstruction, and fully reconstructed backgrounds, where all the final state particles are detected and reconstructed into the signal candidate but one or more are misidentified as the wrong type of particle.

11.1.1 Fully reconstructed backgrounds

$B^0 \to D\pi$

There are two possible ways in which the $B^0 \to D\pi$ background can contribute to the total background. In the case where the D meson decays to a final state of $K^+\pi^-\pi^-$, if one of the π^- are misidentified as a K^- then this will create a background to $B_s^0 \to D_s\pi$ for the case when the D_s meson decays to $K^+K^-\pi^-$.

Moreover, considering the same final state of the D meson ($K^+\pi^-\pi^-$), if the K^+ is misidentified as a π^+ and also one of the π^- are misidentified as a K^- then this will be reconstructed as a D_s ($\pi^+K^-\pi^-$) which is also one of the desired final states of $B_s^0 \to D_s\pi$.

Since both of these background sources contain a D meson which has been misidentified as a D_s meson and the B_s^0 mass was reconstructed with the invariant mass of the D_s meson daughters fixed to the D_s meson mass, the background will peak underneath the signal.

It is therefore important to understand the shape and event yield for this background to avoid it being absorbed into the signal yield. A sample of $B^0 \to D\pi$ from data is selected using the correct D meson mass hypothesis. These events are then reconstructed again as $B_s^0 \to D_s\pi$ using the incorrect D_s mass hypothesis, taking into account the expected shift in momentum distribution caused by the differing particle identification requirements. The resulting distribution is fitted with a smooth PDF (a sum of Gaussian kernels) using RooFit's `RooKeysPdf` and is shown in Figure 11.1.

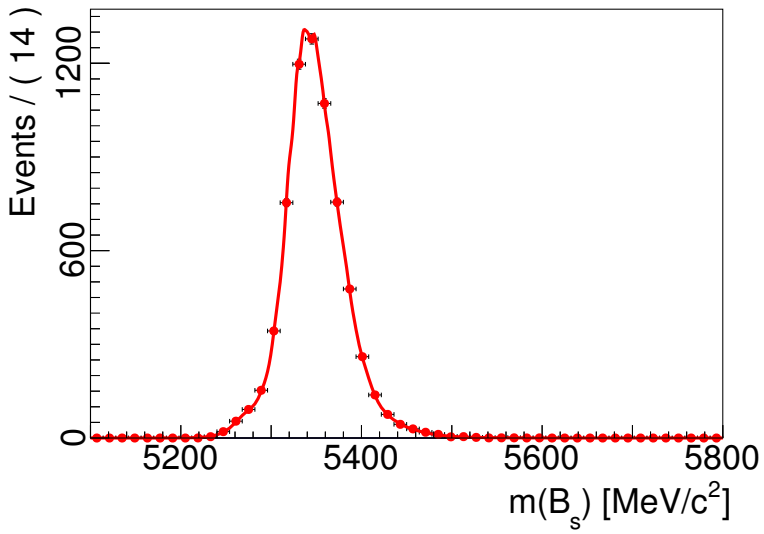


Figure 11.1: Mass distribution and the sum of Gaussian kernels for the $B^0 \to D\pi$ background from the magnet down sample.

$B^0 \to D_s\pi$

It is possible for the B^0 meson to decay to the same finals state as the signal mode and is kinematically similar. Therefore, it is not possible to use particle identification

requirements to reduce it.

The mass shape for this background is taken from simulated events and fitted with a Double Crystal Ball (two Crystal Ball functions [54] sharing a common mean with tails pointing in opposite directions). The Crystal Ball function is defined as

$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases} \tag{11.1}$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \tag{11.2}$$

$$B = \frac{n}{|\alpha|} - |\alpha|\,, \tag{11.3}$$

$$N = \frac{1}{\sigma(C+D)}, \tag{11.4}$$

$$C = \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \tag{11.5}$$

$$D = \sqrt{\frac{\pi}{2}}\left(1 + \mathrm{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right), \tag{11.6}$$

N is a normalization factor, σ is the mean of the core Gaussian function and α, n and $\bar{x}$ are parameters which are fitted to the data.

$\Lambda_b \to \Lambda_c \pi$

The final fully reconstructed background considered is that of $\Lambda_b \to \Lambda_c \pi$. This background arises through the misreconstruction of the Λ_c as a D_s largely due to misidentifying the proton as a kaon in the final state and so shows as a background to the $D_s^+ \to K^+K^-\pi^+$ final state. The simulated data and the sum of Gaussian kernels fit to it is shown in Figure 11.2.

11.1.2 Partially reconstructed backgrounds

Partially reconstructed backgrounds are those where a particle is missed and there may have been one or more misidentifications. They generally peak in lower mass range (except the Λ_b) but may extend to the signal region. These backgrounds can be grouped into three main sub-categories: low-mass B_s^0, low-mass B^0 and Λ_b decays.

The shapes of all backgrounds in this category are all modelled using simu-

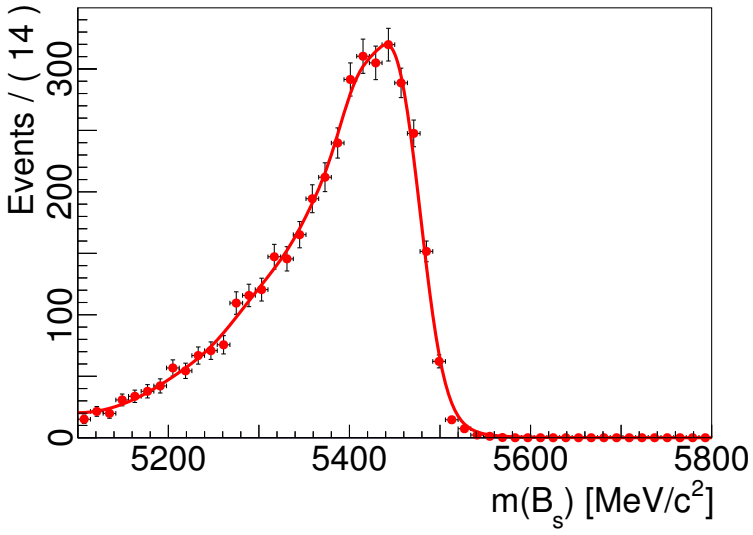


Figure 11.2: Mass distribution and the sum of Gaussian kernels for the $\Lambda_b \to \Lambda_c \pi$ background.

lated data and parameterised with a sum of Gaussian kernels.

Low-mass B_s^0

There are three backgrounds in this group, all due to the decay of a B_s^0 meson. Decays of $B_s^0 \to D_s \rho$ could show as signal if either the ρ is missed in the reconstruction with a background pion selected in its place or if the ρ decays to $\pi^+\pi^0$. If the π^0 is missed and the π^+ carries most of the momentum then the reconstructed B_s^0 will peak near the $D_s\pi$ mass. In the case where the D_s meson is produced via a D_s^*, the intermediate state of $B_s^0 \to D_s^*\pi$ could emit a photon which is then missed by the detector. In the case that both of these occur in the same event, a $B_s^0 \to D_s^*\rho$ could appear as the signal and so will provide a background. The templates used in the mass fit are shown in Figure 11.3.

Low-mass B^0

There are three backgrounds in this group: $B^0 \to D\rho$, $B^0 \to D_s^*\pi$ and $B^0 \to D^*\pi$. These backgrounds occur for similar reasons to the low-mass B_s^0 above; a combination of missed particles and misidentified particle types causing a D or D^* to be misreconstructed as a D_s. The distributions and the fitted templates are shown in Figure 11.4.

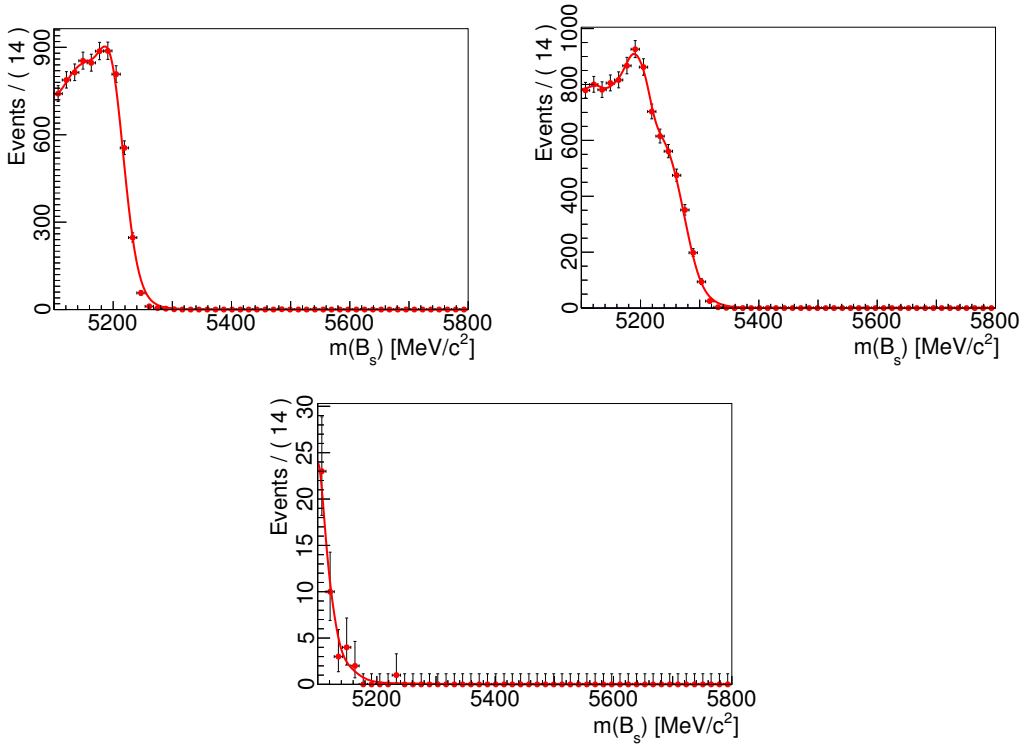


Figure 11.3: Mass distribution and the sum of Gaussian kernels. Top left: $B_s^0 \to D_s\rho$. Top right: $B_s^0 \to D_s^*\pi$. Bottom: $B_s^0 \to D_s^*\rho$.

11.1.3 Combinatorial background

The final background to consider is of the combinatorial nature. This arises due to the reconstruction algorithm choosing random background tracks and combining them while reconstructing a B_s^0 candidate. Due to the momentum distribution of the background particles, the reconstructed B_s^0 mass will follow an approximate exponential distribution, peaking at lower masses and reducing towards higher masses.

The templates for the mass fit are extracted from $B_s^0 \to D_s\pi$ data using the sidebands of the B_s^0 and D_s mass distributions. Events with $m(D_s) = (1868, 1948) \cup (1990, 2068)\,\mathrm{MeV}/c^2$ and $m(B_s^0) > 5600\,\mathrm{MeV}/c^2$ are used. A fit of an exponential function to this data is then extrapolated back over the whole B_s^0 mass range. Each D_s final state is fitted separately to produce an independent mass template. The data distributions with their fitted functions are shown in Figure 11.5 and the parameter values are displayed in Table 11.1.

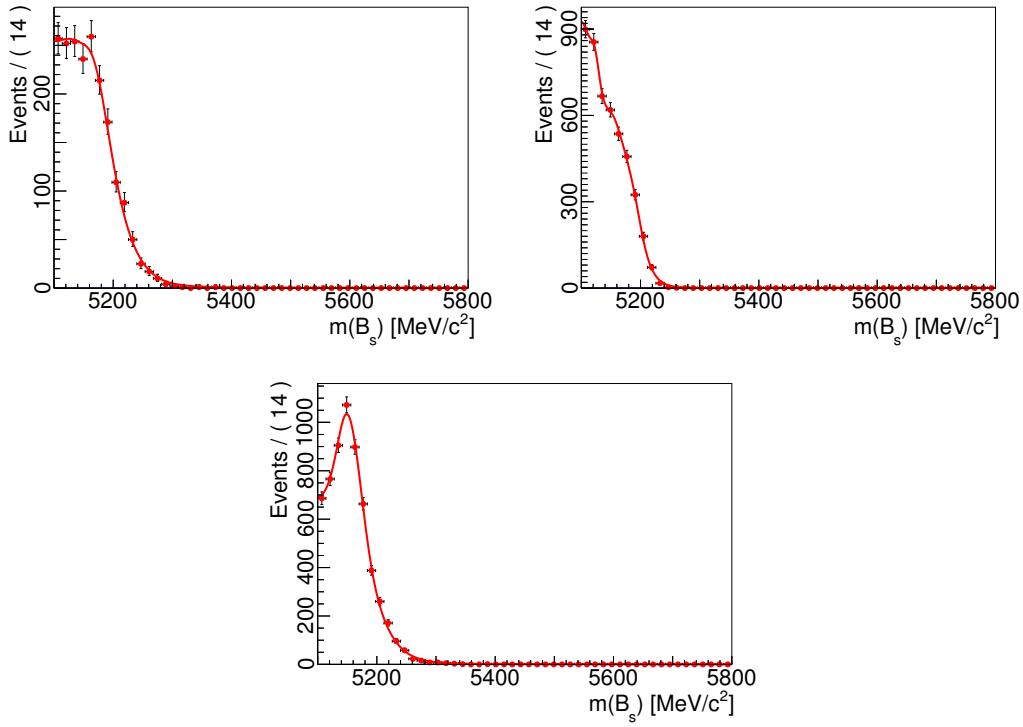


Figure 11.4: Mass distribution and the sum of Gaussian kernels. Top left: $B^0 \to D\rho$. Top right: $B^0 \to D_s^*\pi$. Bottom: $B^0 \to D^*\pi$.

11.2 Mass fit

11.2.1 Signal shape

The signal is described with a double Crystal Ball function in the same way as for the $B_s^0 \to D_s K$ analysis. The shapes of these functions are fixed using a fit to simulated data which has had the full selection applied to it. All shape parameters are floated freely in the fit and the results are given in Table 11.2. The resultant PDF and the simulated data it was fitted to are shown in Figure 11.6. In the final

D_s **mode**	**Slope parameter** ($c^2/\,\mathrm{MeV}$)
$D_s^+ \to K^+K^-\pi^+$	-1.91 ± 0.10
$D_s^+ \to K^+\pi^-\pi^+$	-1.35 ± 0.14
$D_s^+ \to \pi^+\pi^-\pi^+$	-1.26 ± 0.11

Table 11.1: The fitted values (in units of 10^{-3}) of the slope parameter of an exponential function describing the combinatorial background.

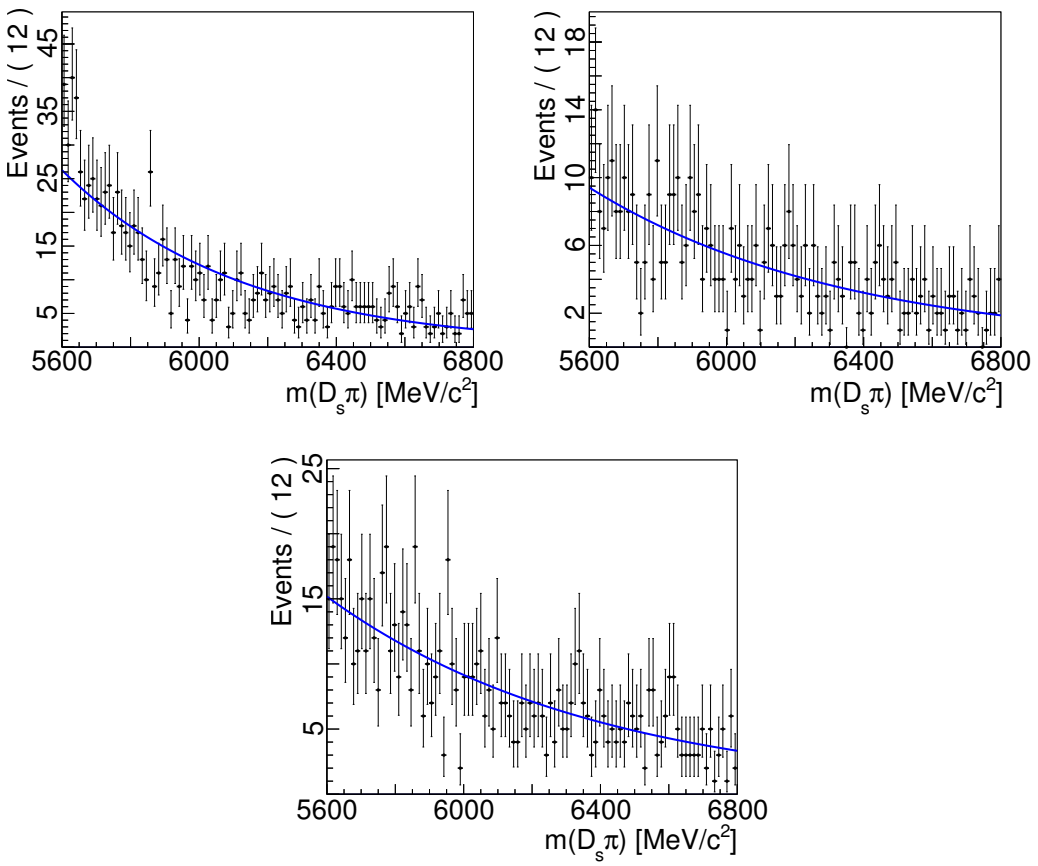


Figure 11.5: Combinatorial background slopes evaluated from the D_s and B_s^0 sidebands. Top left: $D_s^+ \to K^+K^-\pi^+$. Top right: $D_s^+ \to K^+\pi^-\pi^+$. Bottom: $D_s^+ \to \pi^+\pi^-\pi^+$.

mass fit to data, the tail parameters remain fixed but the widths and means are floated.

11.3 Event yields

The total mass fit to the data includes all the mass templates defined in the previous section. The various backgrounds are arranged into five groups.

The first group contains the low-mass partially reconstructed B_s^0 backgrounds. The backgrounds in this group are combined into a single PDF,

$$f_{11}\mathrm{PDF}_{B_s^0\to D_s^*\pi} + (1-f_{11})[f_{12}\mathrm{PDF}_{B_s^0\to D_s\rho} + (1-f_{12})(\mathrm{PDF}_{B_s^0\to D_s^*\rho})]. \qquad (11.7)$$

where f_{1N} are the relative fractions of the background yields. These relative fractions

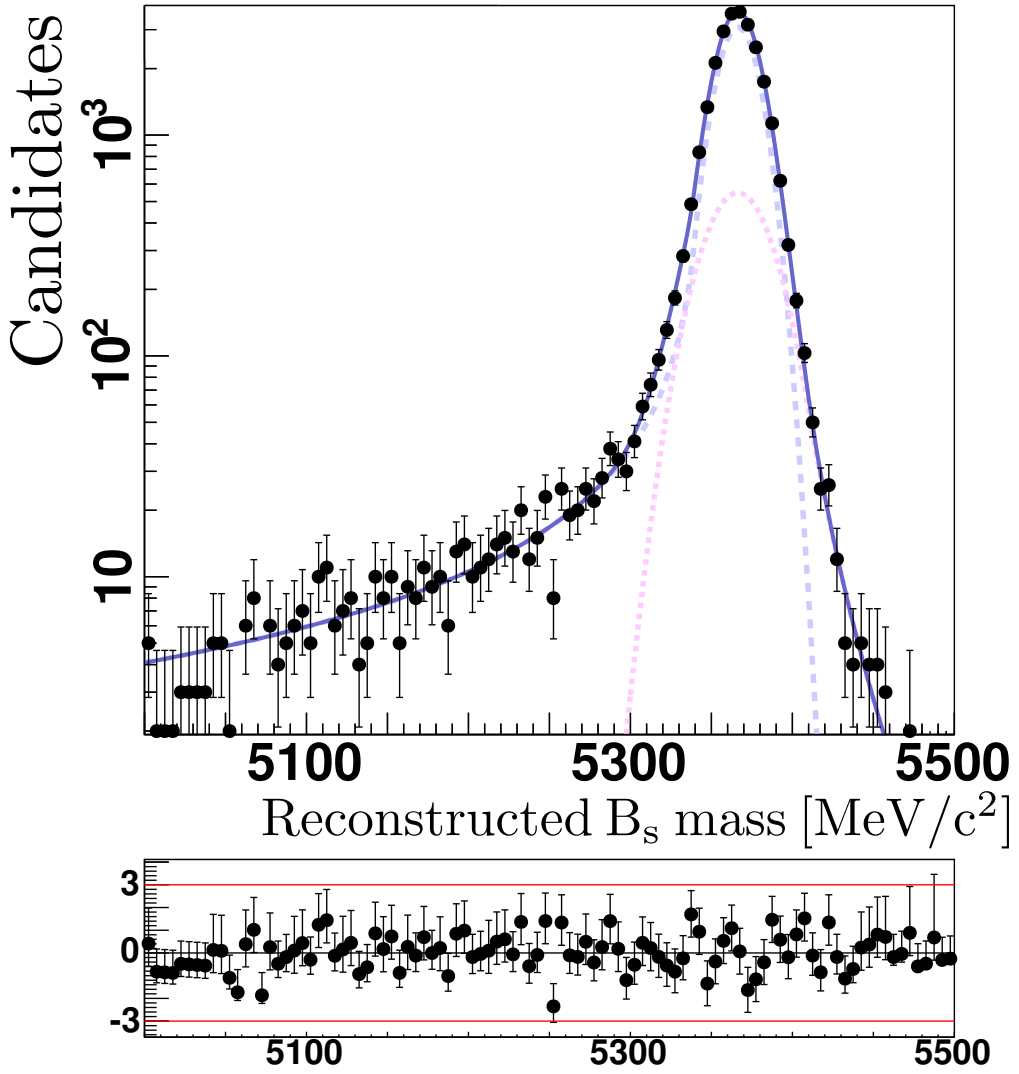


Figure 11.6: Signal mass shapes of $B_s^0 \to D_s\pi$ evaluated on simulated data. The solid lines correspond to the fit to the double Crystal Ball function, while the dashed lines correspond to the individual Crystal Ball components. The bottom plot show the deviation of the data from the fit line based on statistical uncertainty.

Parameter	Fitted value
μ_{DOWN}	5366.3 ± 0.1 MeV/c^2
μ_{UP}	5366.3 ± 0.1 MeV/c^2
σ_1	12.69 ± 0.09 MeV/c^2
σ_2	20.5 ± 0.3 MeV/c^2
α_1	2.13 ± 0.02 MeV/c^2
α_2	-2.06 ± 0.08 MeV/c^2
n_1	1.10 ± 0.02
n_2	5.8 ± 1.3
f	0.78 ± 0.01

Table 11.2: Parameters for the sum of the two Crystal Ball functions describing the signal shapes of $B_s^0 \to D_s\pi$, obtained from simulated data.

are freely floated and shared across all magnet polarities and D_s final states.

The second group contains the low-mass partially reconstructed B^0 backgrounds as well as the fully reconstructed $B^0 \to D_s\pi$ background. The event yields of $B^0 \to D\rho$ and $B^0 \to D^*\pi$ are fixed to be $1/3.5$ and $1/4$ of the $B^0 \to D\pi$ yield respectively based on their relative experimental branching fractions [3]. The $B^0 \to D_s\pi$ and $B^0 \to D_s^*\pi$ yields are both constrained to be equal to $1/30$ of $B_s^0 \to D_s\pi$ [3]. The fixed yields for these two groups are shown in Table 11.3.

The third group consists of just the $B^0 \to D\pi$ background mode. The shape and yield of this background is fixed on data as described in Section 11.1.1 and is shown in Table 11.3.

Then the $\Lambda_b \to \Lambda_c\pi$ misidentified background is considered. Since the branching fraction for this decay is very large, the yield is not fixed. Instead it is floated freely in the fit to the $D_s^+ \to K^+K^-\pi^+$ mode and set to zero for the other two since for it be a background in those modes would require a double misidentification.

The final group contains the combinatorial background and is considered separately for each D_s final state as described in Section 11.1.3.

The fitted values of the parameters in the fit are shown in Table 11.4 and plots of the fit are shown in Figure 11.7.

Mode	**Magnet up**	**Magnet down**
$B^0 \to D\pi$		
$D_s^+ \to K^+K^-\pi^+$	260	363
$D_s^+ \to K^+\pi^-\pi^+$	28	38
$D_s^+ \to \pi^+\pi^-\pi^+$	0	0
$B^0 \to D\rho$		
$D_s^+ \to K^+K^-\pi^+$	74	104
$D_s^+ \to K^+\pi^-\pi^+$	8	11
$D_s^+ \to \pi^+\pi^-\pi^+$	0	0
$B^0 \to D^*\pi$		
$D_s^+ \to K^+K^-\pi^+$	65	91
$D_s^+ \to K^+\pi^-\pi^+$	7	10
$D_s^+ \to \pi^+\pi^-\pi^+$	0	0
$\Lambda_b \to \Lambda_c\pi$		
$D_s^+ \to K^+\pi^-\pi^+$	0	0
$D_s^+ \to \pi^+\pi^-\pi^+$	0	0

Table 11.3: Constrained yields in the $B_s^0 \to D_s\pi$ mass fit.

Parameter	**Fitted value**
mean	$5370.00 \pm 0.19\,\mathrm{MeV}/c^2$
σ_1	$14.95 \pm 0.13\,\mathrm{MeV}/c^2$
σ_2	$27.20 \pm 0.85\,\mathrm{MeV}/c^2$
f_{11}	0.904 ± 0.015
f_{12}	0.861 ± 0.055
$D_s^+ \to K^+K^-\pi^+$ Magnet up	
$N_{B_s^0 \to D_s\pi}$	9180 ± 108
N_{Comb}	1460 ± 75
N_{Group1}	$11\,291 \pm 118$
$N_{\Lambda_b \to \Lambda_c\pi}$	408 ± 65
$D_s^+ \to K^+K^-\pi^+$ magnet down	
$N_{B_s^0 \to D_s\pi}$	$13\,007 \pm 129$
N_{Comb}	1963 ± 87
N_{Group1}	$15\,472 \pm 139$
$N_{\Lambda_b \to \Lambda_c\pi}$	669 ± 79
$D_s^+ \to K^+\pi^-\pi^+$ magnet up	
$N_{B_s^0 \to D_s\pi}$	727 ± 29
N_{Comb}	260 ± 25
N_{Group1}	891 ± 34
$D_s^+ \to K^+\pi^-\pi^+$ magnet down	
$N_{B_s^0 \to D_s\pi}$	1058 ± 35
N_{Comb}	361 ± 30
N_{Group1}	1268 ± 41
$D_s^+ \to \pi^+\pi^-\pi^+$ magnet up	
$N_{B_s^0 \to D_s\pi}$	1679 ± 43
N_{Comb}	520 ± 35
N_{Group1}	1917 ± 50
$D_s^+ \to \pi^+\pi^-\pi^+$ magnet down	
$N_{B_s^0 \to D_s\pi}$	2314 ± 51
N_{Comb}	734 ± 42
N_{Group1}	2581 ± 58

Table 11.4: Fitted values of the parameters for the $B_s^0 \to D_s\pi$ signal mass fit. The N_i are the yields of the signal and background contributions. Mean and width are the parameters of the double Crystal Ball used to describe the signal. The parameters f_i are fractions between modes in the group 1 backgrounds: $B_s^0 \to D_s\rho$, $B_s^0 \to D_s^*\rho$ and $B_s^0 \to D_s^*\pi$.

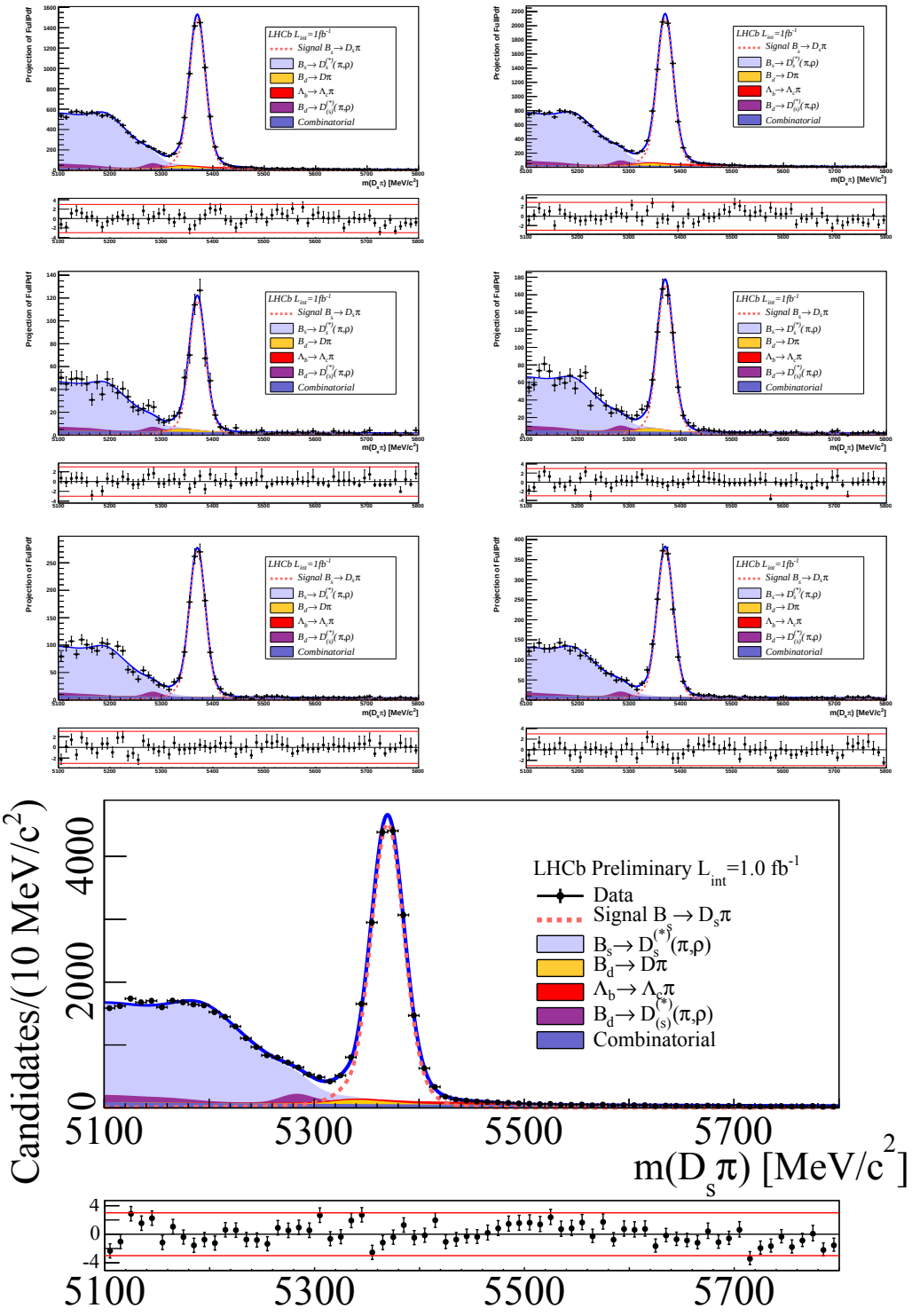


Figure 11.7: Result of the simultaneous fit to the $B_s^0 \to D_s\pi$ candidates, magnet up left, magnet down right, for $D_s^+ \to K^+K^-\pi^+$ (top), $D_s^+ \to K^+\pi^-\pi^+$ (middle), and $D_s^+ \to \pi^+\pi^-\pi^+$ (bottom). The simultaneous fit to the $B_s^0 \to D_s\pi$ candidates for both polarities and all D_s final states combined is shown at the very bottom.

12

Decay-time fit

In order to extract the *CP* parameters, a fit to the reconstructed proper decay-time distribution of the B meson candidates is performed. The yields from the mass fit are used to fix the yields in the decay-time fit.

In contrast to the $B_s^0 \to D_s K$ analysis, the fit to the time distribution of $B_s^0 \to D_s \pi$ candidates is performed using a conventional fit rather than the sFit method. This is because it was found to be difficult to control the correlations between the parameters in the sFit, leading to large statistical uncertainties.

12.1 Background classification

The decay time behaviour of the various backgrounds to the $B_s^0 \to D_s \pi$ decay fall into two main categories:

- Some do not oscillate at all and so can be modelled as a simple exponential decay.
- Others are neutral B^0 or B_s^0 mesons which oscillate while propagating through the detector. Of these types, there are two sub-categories:
  - Those which only decay to a single, flavour-specific, final state. For example, B^0 can only decay into a single final state, f, while $\overline{B}^0$ can only decay into a final state $\bar{f}$.
  - Those which oscillate in flight but also decay with a contribution from a wrong-flavour final state. This means that B^0 can also decay into $\bar{f}$ and $\overline{B}^0$ can decay into f. These are referred to as $B^0 \to D\pi$-*like*.

Oscillating with wrong-flavour	Flavour specific	Non-oscillating
$B^0 \to D\pi$	$B^0 \to D_s\pi$	$\Lambda_b \to \Lambda_c\pi$
$B^0 \to D\rho$	$B^0 \to D_s^*\pi$	Combinatorial
$B^0 \to D^*\pi$	$B_s^0 \to D_s\rho$	
	$B_s^0 \to D_s^*\pi$	
	$B_s^0 \to D_s^*\rho$	

Table 12.1: Categories for the physics backgrounds in the time fit.

The backgrounds that were discussed in Chapter 11 considered in this fit are grouped according to Table 12.1.

For the backgrounds that are flavour-specific, the *CP* parameters from Equation 2.25 are fixed to be $C_f = C_{\bar{f}} = 1$ and $S_f = S_{\bar{f}} = D_f = D_{\bar{f}} = 0$. Meanwhile, the background modes with possible wrong-flavour contributions are fitted with those values fixed to their measured physical values. For the signal mode they are freely floated without constraint. In the standard parameterisation of Equation 2.25, the four S_f, $S_{\bar{f}}$, D_f and $D_{\bar{f}}$ parameters are strongly correlated with each other. In order to remove these correlations in the signal mode, the *CP* variables are instead parameterised as

$$\begin{aligned} \overline{S} &= \frac{S_f + S_{\bar{f}}}{2} \quad & \Delta S &= \frac{S_f - S_{\bar{f}}}{2} \\ \overline{D} &= \frac{D_f + D_{\bar{f}}}{2} \quad & \Delta D &= \frac{D_f - D_{\bar{f}}}{2}. \end{aligned} \tag{12.1}$$

12.2 PDFs and fit setup

The signal mode is treated specially and independently of the other modes and is modelled using the time-dependent equations in Section 2.4. The *CP* parameters are treated as in Equation 12.1 and are floated freely. From Equation 2.26 it can be seen that while S_f, $S_{\bar{f}}$, D_f and $D_{\bar{f}}$ are mostly linearly dependent on $|\lambda_f|$, C depends quadratically on it. Since the expected value of $|\lambda_f|$ is small (in the case of a flavour-specific decay it will be 0), $C - 1$ is expected to be very small and so C is fixed to be 1. In addition to this, Γ_s and $\Delta\Gamma_s$ (as defined in Equation 2.19) are also fixed based on an LHCb measurement of $B_s^0 \to J/\psi\,\phi$ [23] to the values given in Table 12.2. They are fixed in order to reduce the number of free parameters in the fit to speed it up and also to help stability given the strong correlation between $\Delta\Gamma_s$ and ΔD.

The only other physical parameter floated is Δm_s. While allowing this pa-

Parameter	Value
Γ_d	$0.656\,\text{ps}^{-1}$
$\Delta\Gamma_d$	$0\,\text{ps}^{-1}$
Γ_s	$0.658\,\text{ps}^{-1}$
$\Delta\Gamma_s$	$-0.116\,\text{ps}^{-1}$
Δm_d	$0.507\,\text{ps}^{-1}$
Γ_{Λ_b}	$0.719\,\text{ps}^{-1}$
Γ_{comb}	$0.800\,\text{ps}^{-1}$

Table 12.2: Parameters which are fixed in the fit and their values.

Mode	Strong phase	Weak phase	$\lvert\lambda_f\rvert$
$B^0 \to D\pi$	0.0	$-\pi/2$	0.012
$B^0 \to D^*\pi$	0.0	$-\pi/2$	0.015
$B^0 \to D\rho$	0.0	$-\pi/2$	0.038

Table 12.3: Input *CP* parameters for non-flavour-specific backgrounds. Calculated from HFAG [59].

rameter to freely float will negatively affect the determination of some of the *CP* parameters due to strong correlations, it was decided that fixing Δm_s to a published value would potentially bias any measurement of the *CP* parameters of the signal mode. This is because the best measurements from LHCb were measured on our signal mode and many historical measurements were made under the assumption which this analysis is testing. Therefore, Δm_s is allowed to float freely between $5.0\,\text{ps}^{-1}$ and $30.0\,\text{ps}^{-1}$ with an initial value of $17.719\,\text{ps}^{-1}$.

The three non-flavour-specific backgrounds ($B^0 \to D\pi$, $B^0 \to D\rho$ and $B^0 \to D^*\pi$) are again modelled using the time-dependent equations in Section 2.4. The *CP* parameters for these three backgrounds are known and published by previous experiments at BABAR and BELLE [55, 56, 57, 58]. The world-averages of these combined measurements are published by the Heavy Flavour Averaging Group (HFAG) and are given in Table 12.3. These *CP* parameters are fixed in the fit as are Γ_d, $\Delta\Gamma_d$ and Δm_d as given in Table 12.2.

The flavour-specific oscillating modes are grouped into two categories, those coming from a B^0 and those coming from a B_s^0. The same base PDF model is used as in the signal mode but since it is known that the flavour-specific modes have *CP* parameters of $S_f = S_{\bar{f}} = D_f = D_{\bar{f}} = 0$, they are fixed to these values in the fit. As with the non-flavour-specific backgrounds, the B^0 modes have Γ_d, $\Delta\Gamma_d$ and Δm_d fixed while the B_s^0 modes have only Γ_s and $\Delta\Gamma_s$ fixed with Δm_s being floated and

shared with the signal mode.

The two non-oscillating modes are each modelled with a single exponential decay function, the lifetimes for which are fixed. The values of Γ used are based on the mean Λ_b lifetime for the Λ_b mode and $0.8\,\mathrm{ps}^{-1}$ for the combinatorial background.

As for modelling the tagging behaviour, the modes are grouped and are each given an independent tagging efficiency. The signal mode has its own tagging efficiency while the three non-flavour-specific B^0 modes are grouped together as are the two flavour-specific B^0 modes. The three B_s^0 backgrounds are also grouped. The combinatorial and the Λ_b backgrounds each have their own individual tagging efficiency. The tagging efficiency for each group is floated freely in the fit between 0.0 and 1.0. The fit is performed with per-event mistag probabilities rather than using a single average mistag across all events.

12.2.1 Decay-time acceptance

Previously, in the $B_s^0 \to D_s K$ analysis, the acceptance function was calculated using a multi-step process as described in Section 8.2. First a fit to simulated $B_s^0 \to D_s \pi$ data was performed which was then scaled by a ratio between simulated $B_s^0 \to D_s \pi$ and $B_s^0 \to D_s K$. Finally, this model was fitted to real $B_s^0 \to D_s \pi$ data in order to match the properties of the real data distribution.

Since that process relied on the fact that $B_s^0 \to D_s \pi$ was flavour-specific, it is not possible to use that result in this analysis and so the result after just the first step of the fit to the simulated $B_s^0 \to D_s \pi$ data is used as given in Table 8.2.

12.2.2 Blind fit to data

In order to test that the time fit converges correctly, a fit to data is performed. To avoid the results of this fit biasing the analysis, the true central values of the fit are kept hidden. The statistical uncertainties are kept intact however and are shown in Table 12.4.

Parameter	**Statistical uncertainty**
$\overline{S}$	± 0.150
$\overline{D}$	± 0.098
ΔS	± 0.083
ΔD	± 0.050

Table 12.4: Statistical uncertainties from the blind fit to data.

13

Systematic uncertainties

In order to estimate the systematic uncertainties due to various effects, a large number of simulated experiments are performed. Each of the experiments output data in a format which matches the output of the standard LHCb processing so that the full fit procedure can be tested as if it were fitting real data. This allows one to test that the fit converges well given the number of events that we have.

13.1 Experiment generation

The experiments are generated to emulate the real data that we expect as closely as possible. The first step is to create a total PDF of the mass distribution of the events. This total PDF is based on the expected background and signal shapes and their respective expected yields. Events are generated based on this PDF and these events are then fitted using the data mass fitter, which produces yields for each of the channels of interest. A similar process is performed for the time distribution; each channel has an expected distribution of events. However, for the generation of the time distribution, the relative yield for each channel is taken from the output of the mass fit. For each experiment, the number of events generated is set to be equal to the expected yield from the full data sample and for most tests, an ensemble of 500 experiments is created. If any test uses a different number of experiments in the ensemble it is noted in its corresponding place in the text.

The time distribution model for each mode during generation is identical to that used in the fit. Any parameter which is floated in the fit has a specific value used during generation. These values are given in Table 13.1. The *CP* parameters for the three non-flavour-specific backgrounds are grouped together in the fit and

Parameter	Value
Δm_s	$17.719\,\mathrm{ps}^{-1}$
$\overline{S}$	0.0
$\overline{D}$	0.0
ΔS	0.0
ΔD	0.0
$\varepsilon_{\mathrm{tag}}$	0.403

Table 13.1: Values for parameters used to generate experiments.

floated, but during the generation they are calculated separately. Table 12.3 shows the values used as inputs to the process.

Each time an experiment is run it is given a different seed for its random number generator. Events are generated based on the total PDF using an accept/ reject method.

Given an ensemble of experiments, it is possible to check whether a certain fit method is biasing certain observables and also whether the reported uncertainty is realistic. For a given experiment and observable, a value called the pull is calculated. For an observable x, the pull, g, is defined as

$$g = \frac{x_{\mathrm{fit}} - x_{\mathrm{gen}}}{\sigma_{\mathrm{fit}}}, \tag{13.1}$$

where x_{fit} is the measured value given by the fit, x_{gen} is the value that was used to generate the experiment and is usually the expected physical value and σ_{fit} is the uncertainty on the fitted value output from the fit. If g is calculated for each experiment in the ensemble, a pull distribution is created and should exhibit two properties:

- The mean of the distribution should equal to 0 and any deviation from that indicates a bias in the fit.
- The pulls should be Gaussian distributed with a width of 1.0. If it is wider or narrower then it indicates that the uncertainties on the fitted values are being over- or underestimated.

In addition to checking the pull distribution of the experiments it is also important to have a handle on which observables are correlated with each other. If two correlated variables are allowed to float in the fit the sensitivity is lost due to variations in either of them. In the case that one of them is an observable from which one wants to precisely measure a physical quantity, the lost sensitivity

Parameter	Pull mean	Pull width
ΔS $(B_s^0 \to D_s\pi)$	0.04 ± 0.03	1.03 ± 0.02
ΔD $(B_s^0 \to D_s\pi)$	0.02 ± 0.03	0.99 ± 0.02
$\overline{S}$ $(B_s^0 \to D_s\pi)$	0.02 ± 0.04	1.04 ± 0.03
$\overline{D}$ $(B_s^0 \to D_s\pi)$	-0.06 ± 0.02	1.04 ± 0.03
Δm_s	-0.01 ± 0.03	1.01 ± 0.02
$B_s^0 \to D_s\pi$ ε_{tag}	0.01 ± 0.03	0.98 ± 0.02
$B^0 \to D\pi$-like ε_{tag}	-0.05 ± 0.03	0.95 ± 0.02
$B^0 \to D_s^{(*)}\pi$ ε_{tag}	-0.07 ± 0.02	0.61 ± 0.01
Flavour-specific B_s^0 ε_{tag}	0.01 ± 0.03	0.91 ± 0.02
Λ_b ε_{tag}	-0.02 ± 0.03	0.96 ± 0.02
Combinatorial ε_{tag}	0.03 ± 0.03	0.98 ± 0.02

Table 13.2: Means and widths of the pulls of the floated parameters.

can be disastrous. Again, in any one experiment, two variables may appear to be strongly correlated but this may be due to some statistical fluctuation. Drawn as a distribution over a large ensemble however, a single peak should appear and give an accurate representation of the correlation.

13.2 Fitter validation

To validate that the time fit is performing correctly and that the parameters of interest are unbiased, an ensemble of experiments are generated and fitted. Once the events have been generated, they are fitted as if they were real data.

An ensemble of 500 experiments are created of which 460 have their fit converge correctly. A pull distribution for each floated parameter is fitted with a single Gaussian function. Distributions of fitted values, their uncertainty and pulls for $\overline{S}$, $\overline{D}$, ΔS and ΔD are plotted in Figures 13.1 and 13.2. Corresponding distributions for Δm_s are shown in Figure 13.3.

The means and widths of the Gaussian function used to fit the pull distributions of the floated parameters are shown in Table 13.2. It can be seen that all the pulls and widths are consistent with zero. The only exception is the $B^0 \to D_s^{(*)}\pi$ tagging efficiency which has a pull width which is too small. This suggests that the fit is overestimating the uncertainty on this parameter. This is not expected to have a large effect on the result as the yield of that background is approximately 0.1% of the signal yield.

Finally, the correlations between floated parameters across the ensemble is investigated. For each pair of variables in each experiment, a correlation is calculated

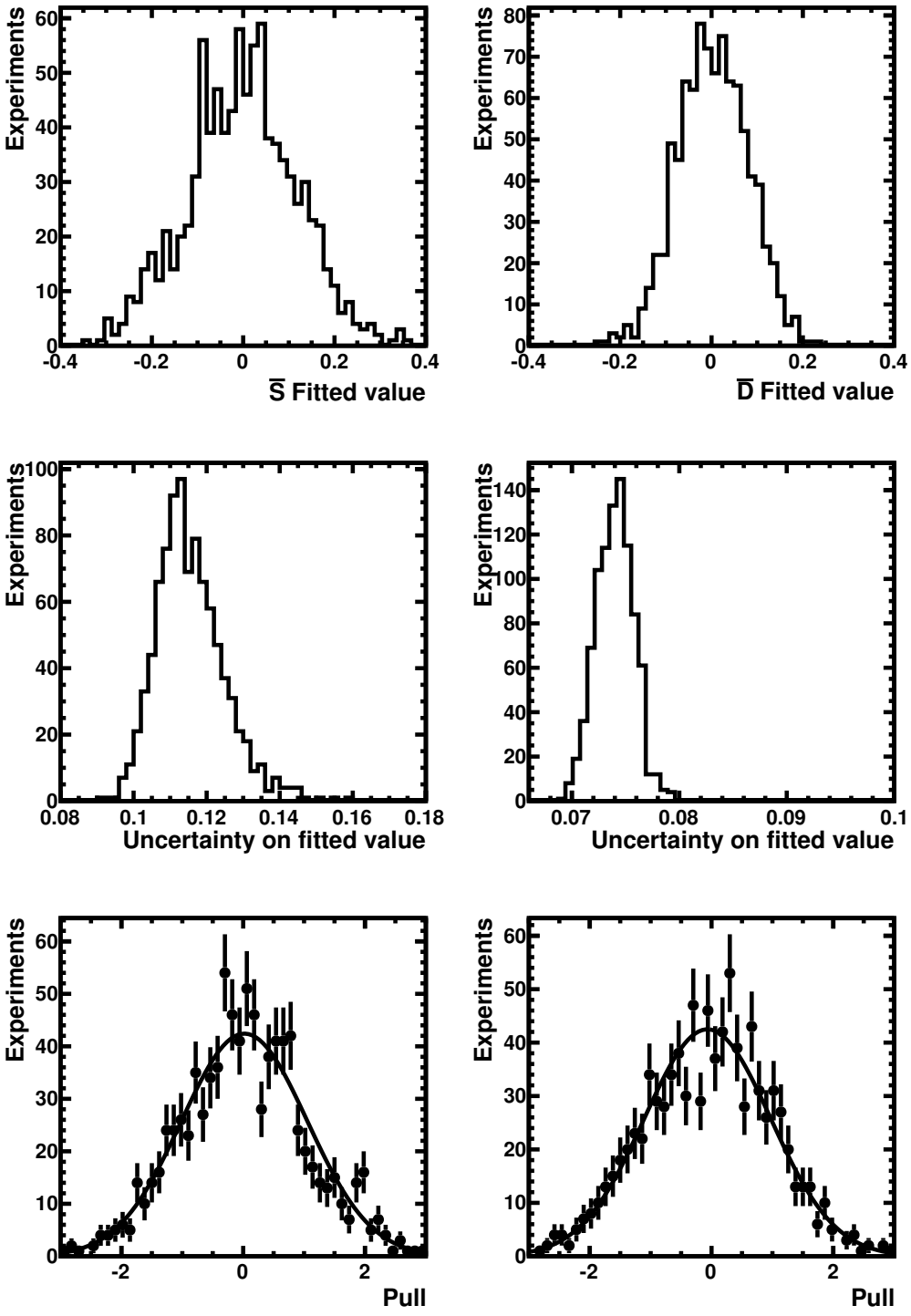


Figure 13.1: Distributions for $\overline{S}$ and $\overline{D}$. From top to bottom in each plot, fitted value, uncertainty and pull distribution. The pull distribution is fitted with a Gaussian function.

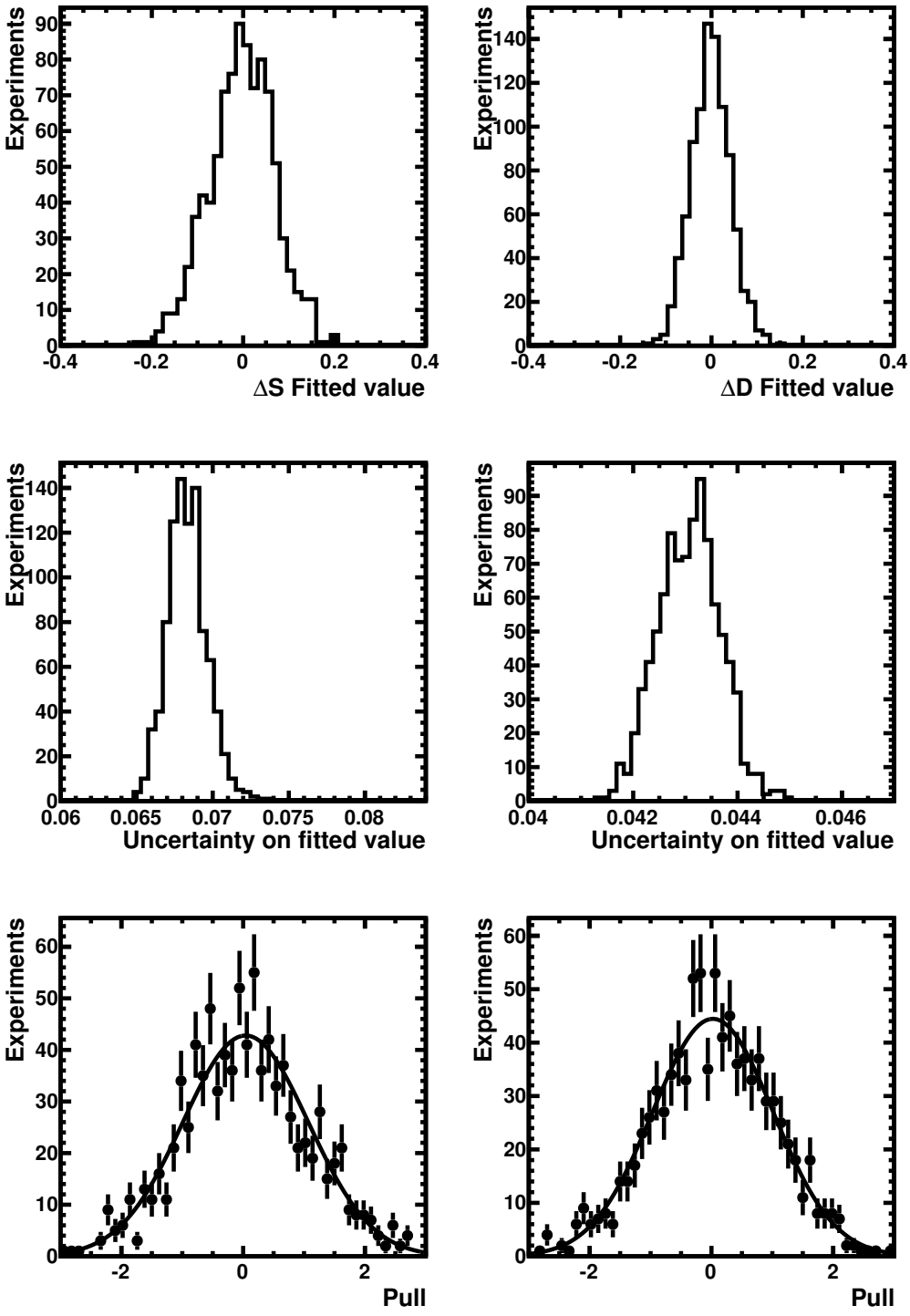


Figure 13.2: Distributions for ΔS and ΔD. From top to bottom in each plot, fitted value, uncertainty and pull distribution. The pull distribution is fitted with a Gaussian function.

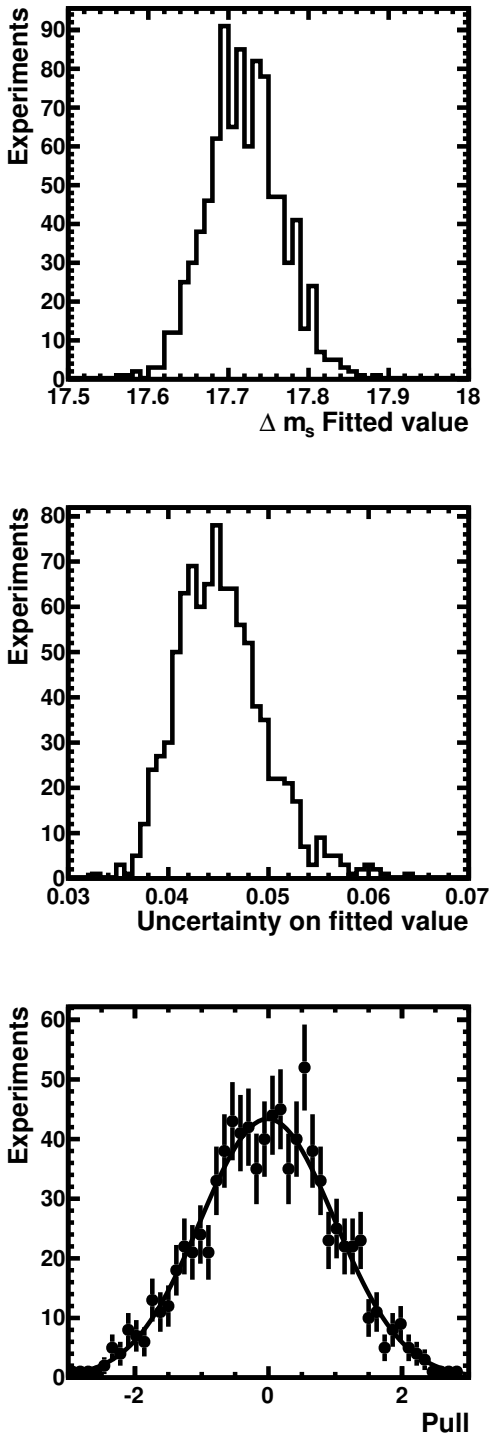


Figure 13.3: From top to bottom, fitted value of Δm_s, uncertainty on Δm_s and pull distribution. The pull distribution is fitted with a Gaussian function.

by the software performing the fit. Since the precise value of this correlation is subject to the statistical limits of the number of events in each fit, an average is calculated across the ensemble. The average correlation between the parameters are given in Table 13.3.

As can be seen, the four *CP* parameters of interest are completely uncorrelated with each other and generally have very small correlations with other parameters in the fit. The only large correlation is between $\overline{S}$ and Δm_s at the level of -0.81. A large correlation between these variables is expected given the mathematical formulation used in the PDFs.

13.3 Sources of systematic uncertainties

The process of establishing the systematic uncertainties is covered in the following sections. The general procedure is that for each parameter for which we want to calculate a systematic uncertainty, we generate a large ensemble of experiments and then fit it twice, once with the particular parameter fixed at a given nominal value and another with the parameter varied by some fixed amount. The amount by which the parameter is varied is usually 1σ of its experimentally measured value but in cases where this isn't possible, it is varied by a fixed percentage amount.

To extract the effect of the systematic uncertainty, the distribution of fitted values for both the nominally fitted ensemble and the varied one are fitted with a Gaussian function. The amount by which the mean of the Gaussian function shifts is taken as the systematic uncertainty and in the case that the variation up and variation down result in differing measured shifts, the larger of the two is used for the total contribution.

13.3.1 Decay-time resolution

To study the effect of the decay-time resolution model, the experiments are fitted with a different resolution to that used in the generation. The width of all three Gaussian functions constituting the resolution model is scaled both up and down by 20% and the effect of this as a systematic uncertainty is given in Table 13.4.

The modelling of the time resolution has a very small effect on the measured values. In particular, ΔS and ΔD are negligibly affected while $\overline{S}$ and $\overline{D}$ have a measurable shift.

Parameter	$B_s^0 \to D_s\pi\ \overline{S}$	$B_s^0 \to D_s\pi\ \overline{D}$	$B_s^0 \to D_s\pi\ \Delta S$	$B_s^0 \to D_s\pi\ \Delta D$	Δm_s	$B_s^0 \to D_s\pi\ \varepsilon_{\rm tag}$	$B^0 \to D\pi$-like $\ \varepsilon_{\rm tag}$	$B^0 \to D_s^{(*)}\pi\ \ \varepsilon_{\rm tag}$	Flavour-specific $B_s^0\ \varepsilon_{\rm tag}$	$\Lambda_b\ \varepsilon_{\rm tag}$	Combinatorial $\ \varepsilon_{\rm tag}$
$B_s^0 \to D_s\pi\ \overline{S}$	1	0.00	−0.01	0.00	−0.81	0.00	0.00	0.00	0.00	0.00	0.00
$B_s^0 \to D_s\pi\ \overline{D}$	0.00	1	0.00	0.00	0.00	−0.01	0.02	0.01	−0.02	0.00	0.01
$B_s^0 \to D_s\pi\ \Delta S$	−0.01	0.00	1	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
$B_s^0 \to D_s\pi\ \Delta D$	0.00	0.00	0.00	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Δm_s	−0.81	0.00	0.01	0.00	1	0.00	0.00	0.00	0.00	0.00	0.00
$B_s^0 \to D_s\pi\ \varepsilon_{\rm tag}$	0.00	−0.01	0.00	0.00	0.00	1	−0.42	0.01	0.28	−0.03	−0.39
$B^0 \to D\pi$-like $\ \varepsilon_{\rm tag}$	0.00	0.02	0.00	0.00	0.00	−0.42	1	0.13	−0.29	−0.32	0.18
$B^0 \to D_s^{(*)}\pi\ \ \varepsilon_{\rm tag}$	0.00	0.01	0.00	0.00	0.00	0.01	0.13	1	−0.15	−0.02	0.04
Flavour-specific $B_s^0\ \varepsilon_{\rm tag}$	0.00	−0.02	0.00	0.00	0.00	0.28	−0.29	−0.15	1	0.24	−0.42
$\Lambda_b\ \varepsilon_{\rm tag}$	0.00	0.00	0.00	0.00	0.00	−0.03	−0.32	−0.02	0.24	1	−0.45
Combinatorial $\ \varepsilon_{\rm tag}$	0.00	0.01	0.00	0.00	0.00	−0.39	0.18	0.04	−0.42	−0.45	1

Table 13.3: Mean values of correlations between floated parameters across the whole ensemble.

Parameter	Scaled up	Scaled down
$\overline{S}$	-0.005	0.022
$\overline{D}$	-0.020	0.018
ΔS	-0.001	-0.001
ΔD	0.000	0.000

Table 13.4: Systematic uncertainties due to fitting with a scaled decay-time resolution.

Parameter	p_0	p_1
$\overline{S}$	0.004	0.002
$\overline{D}$	0.007	0.003
ΔS	0.001	0.001
ΔD	0.000	0.000

Table 13.5: Systematic uncertainties due to fitting with varied flavour tagging calibration parameters.

13.3.2 Flavour tagging calibration parameters

The flavour tagging calibration parameters used in the fit are given with experimental uncertainties. To measure the systematic uncertainty due to this in the fit, both p_0 and p_1 are varied within their uncertainties which for p_0 is ± 0.009 and for p_1 it is ± 0.024. The systematic uncertainty due to the flavour tagging calibration is given in Table 13.5.

As with the decay time resolution, the shifts in ΔS and ΔD are very small but here the variations in $\overline{S}$ and $\overline{D}$ are also both very small.

13.3.3 Background yields

The fixed values of the background yields are varied individually and the fit rerun for each. A fit is run with the background yield of each increased to 150% of its nominal value and again with it reduced to 50%. The combinatorial background yield is varied within its uncertainty.

The result of performing this study is given in Table 13.6.

Since the total yields of the backgrounds is quite small under the $B_s^0 \to D_s\pi$ signal mass peak, the effect of the backgrounds yields is small. In all four of the *CP* parameters, the combined effect is small with only $\overline{S}$ and $\overline{D}$ having any significant uncertainty.

Background	$\overline{S}$	$\overline{D}$	ΔS	ΔD
$B^0 \to D\pi$	0.003	0.003	0.001	0.001
$B^0 \to D_s\pi$	0.004	0.002	0.001	0.001
$B^0_s \to D^*_s\pi$	0.002	0.003	0.001	0.001
$B^0_s \to D_s\rho$	0.001	0.002	0.000	0.001
$B^0_s \to D^*_s\rho$	0.001	0.002	0.000	0.001
$B^0 \to D^*\pi$	0.001	0.002	0.000	0.000
$B^0 \to D\rho$	0.002	0.002	0.000	0.001
$B^0 \to D^*_s\pi$	0.001	0.001	0.001	0.001
$\Lambda_b \to \Lambda_c\pi$	0.002	0.007	0.001	0.001
Combinatorial	0.001	0.004	0.000	0.001
Total	0.007	0.010	0.002	0.003

Table 13.6: Change in measured value of *CP* parameter due to varying fixed background yields.

Parameter	**Shift**
$\overline{S}$	0.005
$\overline{D}$	0.008
ΔS	0.002
ΔD	0.001

Table 13.7: Systematic uncertainties due to fitting with the sFit method to test the PDF parameterisation of the backgrounds.

13.3.4 Decay-time PDFs of backgrounds

The parameterisation of the various backgrounds used in the time fit is tested by comparing the standard fit with the sFit as described in Section 8. Since the sFit does not contain any modelling of the decay-time distribution of the backgrounds, an sFit cross-check will determine the effect of any mismodelling of the background distributions. As much as is possible, all other configuration for the fit such as which parameters are fixed or floated as well as the decay-time resolution and acceptance functions are kept the same.

In order to reduce any unknown differences between the two fit methods, a single ensemble is generated using the same method as for the rest of the tests. This ensemble is fitted with each of the two fit methods and as with the other systematic sources, the fitted value distributions are fitted with a Gaussian function and the means are compared. The systematic shifts are shown in Table 13.7.

As with the background yields, the effect of the modelling of the backgrounds

Parameter	Variation	Systematic uncertainty			
		$\overline{S}$	$\overline{D}$	ΔS	ΔD
Γ_s	± 0.0054	0.001	0.097	0.001	0.001
$\Delta\Gamma_s$	± 0.018	0.001	0.065	0.001	0.001
$\Delta\Gamma_d$	± 0.009	0.002	0.004	0.001	0.001
Δm_d	± 0.004	0.002	0.001	0.000	0.000
Total		0.003	0.117	0.002	0.002

Table 13.8: Measured systematic uncertainties in *CP* parameters due to fitting with varied fixed physics parameters. The amount by which the fixed parameters are varied are also given.

is small.

13.3.5 Physics parameters which are fixed in the fit

There are a number of parameters which are fixed in the fit to the central values of published measurements. To measure the effect of these assumptions, the fit is run with each of them fixed to a value which has been varied by $\pm 1\sigma$. The estimated systematic uncertainty given by each of the fixed parameters is given in Table 13.8.

The variation in the measured values of the *CP* parameters is largely unaffected by the uncertainty in the fixed physics parameters except in the case of $\overline{D}$. Here there is a large uncertainty contribution from both Γ_s and $\Delta\Gamma_s$ which is expected given the formulation given in Equation 2.25. The parameterisation given in Equation 12.1 also allows the systematic effect to be constrained to only affect $\overline{D}$ without affecting ΔD.

13.3.6 Production, detection and flavour tagging asymmetries

In the fit it is possible that there are some production or flavour tagging asymmetries which are unaccounted for. To study how the fit is influenced by these, a set of experiments with asymmetries present is produced and fitted. To account for differing flavour tagging efficiencies, the combinatorial background and the signal are generated with different values: 60% for the combinatorial and 40% for the signal as this is what is seen in data.

Potential production asymmetries are considered by generating events with an imbalance with respect to how many B mesons are created compared to $\overline{B}$ mesons. The B, B_s^0 and Λ modes are generated with a 3% asymmetry while the combinatorial is generated with no asymmetry at all. A small 1% flavour tagging

Parameter	Shift
$\overline{S}$	0.006
$\overline{D}$	0.009
ΔS	0.001
ΔD	0.169

Table 13.9: Systematic uncertainties due to fitting an ensemble with asymmetries compared to the nominal ensemble.

efficiency asymmetry between B and $\overline{B}$ candidates is also included.

Since there would be very large variations between a single experiment with this configuration and the nominal experiment, a toy-by-toy comparison is not possible. Instead, the pull distribution of each ensemble is compared and a single systematic shift value is extracted for each parameter. The results of these are summarised in Table 13.9.

The effect of the asymmetries on the fit are negligible for $\overline{S}$, $\overline{D}$ and ΔS. There is a significant contribution in ΔD however of an amount larger then the statistical uncertainty.

13.3.7 Decay time acceptance function

Since the acceptance function parameters are only measured on simulated events and have large relative uncertainties, measuring their systematic effects is important, particularly on parameters which depend strongly on accurate measurements of the lifetime of the particles.

Each of the four *CP* parameters are varied both up and down by 1σ based on their systematic uncertainty from the fit to simulated events. Due to the strong correlations between the variables in the acceptance model (shown in Table 8.3), the uncertainties are not simply added in quadrature but also take into account the correlations between the variables. The effect on the fitted *CP* parameters is shown in Table 13.10.

Both ΔS and ΔD as well as $\overline{S}$ are only negligibly affected by the decay-time acceptance function. However, $\overline{D}$ is affected strongly by the acceptance function. The large effect on $\overline{D}$ is expected given the large uncertainties on the parameters of the acceptance function and since $\overline{D}$ is directly related to the average effective lifetime of the B_s^0. Since these modelling parameters for the acceptance function came from a fit to simulated data, the acceptance function on real data is not known well. In the analysis on $B_s^0 \to D_s K$ described in Part II, a fit to real $B_s^0 \to D_s \pi$ data

Parameter	Variation	Systematic uncertainty			
		$\overline{S}$	$\overline{D}$	ΔS	ΔD
a (turn-on)	± 0.204	0.002	0.972	0.001	0.001
b (offset)	± 0.0364	0.002	-0.148	0.001	0.001
n (exponent)	± 0.066	0.002	-0.038	0.001	0.001
β	± 0.0118	0.003	0.268	0.001	0.001
Total		0.003	0.528	0.002	0.002

Table 13.10: Measured systematic uncertainties in *CP* parameters due to fitting with varied model parameters for the decay time acceptance function. The amount by which the fixed parameters are varied are also given.

was performed to extract the acceptance parameters. In that fit, the acceptance parameter a was found to be 1.215 ± 0.053 — a full standard deviation away from the value used in the fit shown here.

In order to give a precise determination of $\overline{D}$, an analysis would have to give much greater care to the exact shape of the acceptance function, for example using a data-driven method. A common example of this is a technique known as *swimming* whereby real data events are altered to give a different meson decay time and then measure how that affects whether or not the event would be accepted. This is performed over a large sample of events to extract the ensemble's acceptance function. While this method provides a good description of the acceptance model, it requires a good understanding of how the decay time is experimentally correlated with other parameters which would have taken too much time.

13.3.8 Total systematic uncertainties

Each source of systematic uncertainty is assumed to be uncorrelated with the others and so the contributions are added in quadrature to give the total estimate of the systematic uncertainty shown in Table 13.11.

The systematic uncertainties on $\overline{S}$ and ΔS are both small compared to their statistical uncertainty (17% and 5% respectively). The systematic uncertainty on ΔD is sizeable (just over 3 times the statistical uncertainty) and is almost entirely due to assumptions about asymmetries in the fit. Finally, $\overline{D}$ has a very large systematic uncertainty over five times more than that due to statistics. There is a sizeable contribution to this from Γ_s and $\Delta\Gamma_s$ but by far the largest contribution is from the modelling of the decay-time acceptance.

Source	Parameter $\overline{S}$	$\overline{D}$	ΔS	ΔD
Decay-time resolution	0.022	0.020	0.001	0.000
Flavour tagging calibration	0.004	0.008	0.001	0.000
Background yields	0.007	0.010	0.002	0.003
Background parametrisation	0.005	0.008	0.002	0.001
Physics parameters	0.003	0.117	0.002	0.002
Asymmetries	0.006	0.009	0.001	0.169
Decay time acceptance	0.003	0.528	0.002	0.002
Total systematic uncertainty	0.025	0.541	0.004	0.169

Table 13.11: Summary of the sources of systematic uncertainty in the time fit.

14

Summary

The results of the fit to $B_s^0 \to D_s\pi$ data are

$$\overline{S} = \quad 0.197 \pm 0.150 \pm 0.025, \tag{14.1}$$

$$\overline{D} = -0.888 \pm 0.098 \pm 0.541, \tag{14.2}$$

$$\Delta S = \quad 0.066 \pm 0.083 \pm 0.004, \tag{14.3}$$

$$\Delta D = -0.062 \pm 0.050 \pm 0.169, \tag{14.4}$$

where the first uncertainty is statistical and the second is systematic.

All the parameters are within 2σ of zero and ΔS and ΔD are less than 1σ. At this level of uncertainty, there is no evidence of non-flavour-specific decays of $B_s^0 \to D_s\pi$.

Furthermore, ΔD is related to the difference in the effective lifetimes of the $D_s^-\pi^+$ and $D_s^+\pi^-$ final states. If the above result is argued to state that $|\Delta D| < 0.1$ then it is possible to put a constraint on the difference in effective lifetimes between these two final states. The effective lifetime is given by

$$\tau_{\text{eff}} = \tau_{B_s^0}\left(1 + D_f \times y_s\right), \tag{14.5}$$

where

$$y_s = \frac{\Delta\Gamma_s}{2\Gamma_s} = 0.09. \tag{14.6}$$

Therefore, taking the difference between the effective lifetimes,

$$\frac{\Delta\tau_{\text{eff}}}{\tau_{B_s^0}} = 2\Delta D \times y_s, \tag{14.7}$$

and using the measured value of ΔD we find

$$\left|\frac{\Delta\tau_{\text{eff}}}{\tau_{B_s^0}}\right| \lesssim 0.02. \tag{14.8}$$

This shows that the $D_s^-\pi^+$ and the $D_s^+\pi^-$ have the same lifetime to better than 2%. In the case that the $B_s^0 \to D_s\pi$ decay is assumed to be flavour-specific, this provides a test of CPT invariance which predicts that particles and anti-particles have equal lifetimes.

Comparing these results to those obtained from the $B_s^0 \to D_sK$ analysis, there are a few notable differences in the results, particularly with respect to the uncertainties. Firstly, the statistical uncertainties are generally higher across all parameters in $B_s^0 \to D_sK$ and this is simply due to the reduced number of events being fitted. However, they are also much more consistent, with all four statistical uncertainties (and systematic uncertainties) being very similar to each other.

In the $B_s^0 \to D_s\pi$ results shown above, there is clearly a much larger variation between the uncertainty on each parameter and this variation is driven largely by the correlations between the underlying D_f, $D_{\bar{f}}$, S_f and $S_{\bar{f}}$. This is explored in more detail in Appendix A.

Using the same method described in the appendix and applying it to the results of the $B_s^0 \to D_sK$ fit (using the correlations given in Table 8.7), it is found that the statistical uncertainties of a reparameterised $B_s^0 \to D_sK$ fit would be those given in Table 14.1. Here it is clear that the statistical uncertainties of $\overline{D}$ and ΔD have been altered considerably, entirely due to the correlation between D_f and $D_{\bar{f}}$.

Parameters	**Statistical uncertainty**
$\overline{S}$	0.44
$\overline{D}$	0.51
ΔS	0.45
ΔD	0.27

Table 14.1: Estimated statistical uncertainties of the reparameterised $B_s^0 \to D_sK$ CP parameters.

It is also of note that in the altered uncertainties given in Table 14.1, $\overline{S}$ and ΔS are still effectively the same as each other, in contrast to the difference between $\overline{S}$ and ΔS from $B_s^0 \to D_s\pi$ shown at the beginning of this section. This is due to a difference in analysis method leading to a difference in the correlation matrix between $B_s^0 \to D_s\pi$ and $B_s^0 \to D_sK$. In the $B_s^0 \to D_s\pi$ analysis, S_f and $S_{\bar{f}}$ are highly correlated (see Table A.2) with each other but in the $B_s^0 \to D_sK$ analysis

they are not. This is caused by the difference in the treatment of Δm_s. When fitting $B_s^0 \to D_s K$, Δm_s was fixed but when fitting $B_s^0 \to D_s \pi$ it was floated, allowing it to carry information between S_f and $S_{\bar{f}}$, increasing the correlation.

It is worth noting that it can be seen that performing the *CP* reparameterisation has reduced the overall uncertainty on the result. Looking at the $B_s^0 \to D_s K$ statistical uncertainties, all four statistical uncertainties are lower than they were compared to the conventional approach while at the same time, removing all correlations between the parameters. It might be possible that while the statistical uncertainties have shrunk, the systematic uncertainties may have grown. However, while the latter are large on the $B_s^0 \to D_s \pi$ results, these are largely due the fact that Δm_s was not fixed and the parameterisation of the decay-time acceptance rather than the structure of the *CP* parameters.

Since the value of Δm_s was not fixed in the fit to $B_s^0 \to D_s \pi$ data, it is possible to make a measurement of it. A value of $\Delta m_s = 17.71 \pm 0.06\,\text{ps}^{-1}$ was obtained where the uncertainty is statistical. Comparing this with the currently published world-average of $17.69 \pm 0.08\,\text{ps}^{-1}$ [3] shows that they are in agreement. The latest published results from LHCb give values of $17.93 \pm 0.22 \pm 0.15\,\text{ps}^{-1}$ [60] where the first uncertainty is statistical and the second is systematic. While the measurement from $B_s^0 \to D_s \pi$ may appear to be very competitive with both of these, no attempt has been made here to quantify the systematic uncertainties on Δm_s.

A

Propagation of uncertainties in *CP* parameterisation

The results of the measurement of the *CP* parameters in $B_s^0 \to D_s K$ found that S_f, $S_{\bar{f}}$, D_f and $D_{\bar{f}}$ have approximately equal statistical uncertainties. This is what would be expected given the formulation of the decay rates given in Equation 2.25.

During the $B_s^0 \to D_s \pi$ analysis, these four *CP* parameters were reparameterised as

$$\begin{aligned} \overline{S} &= \frac{S_f + S_{\bar{f}}}{2}, \quad & \Delta S &= \frac{S_f - S_{\bar{f}}}{2}, \\ \overline{D} &= \frac{D_f + D_{\bar{f}}}{2}, \quad & \Delta D &= \frac{D_f - D_{\bar{f}}}{2}. \end{aligned} \tag{A.1}$$

with the aim of reducing the correlation between the parameters. This was successful as can be seen in Table A.1 where all the correlations were found to be negligible.

Naïvely, one might expect that since the original *CP* parameters had similar statistical uncertainties and since they are being combined linearly with identical

Parameters	$\overline{S}$	$\overline{D}$	ΔS	ΔD	**Statistical uncertainty**
$\overline{S}$	1	0.00	0.00	0.00	0.150
$\overline{D}$	0.00	1	0.00	0.00	0.098
ΔS	0.00	0.00	1	0.00	0.083
ΔD	0.00	0.00	0.00	1	0.050

Table A.1: The correlations and statistical uncertainties for the *CP* parameters in the fit to $B_s^0 \to D_s \pi$ data.

Parameters	S_f	$S_{\bar{f}}$	D_f	$D_{\bar{f}}$	**Statistical uncertainty**
S_f	1	0.535	-0.053	-0.039	0.17937
$S_{\bar{f}}$	0.535	1	-0.036	-0.032	0.16368
D_f	-0.053	-0.036	1	0.584	0.1113
$D_{\bar{f}}$	-0.039	-0.032	0.584	1	0.10862

Table A.2: The correlations and statistical uncertainties for the *CP* parameters in the fit to $B_s^0 \to D_s\pi$ data.

coefficients that the resultant uncertainties of $\overline{S}$, $\overline{D}$, ΔS and ΔD would be consistent with each other.

However, this does not take into account the correlations between the original *CP* parameters which were found in the $B_s^0 \to D_s K$ to be non-negligible. A test fit is run over the $B_s^0 \to D_s\pi$ data with the decay rate equations parameterised as S_f, $S_{\bar{f}}$, D_f and $D_{\bar{f}}$ in order to find the correlations and statistical uncertainties of the parameters. The results of this fit are shown in Table A.2.

The uncertainties of the reparameterised *CP* parameters, taking into account the correlations between the progenitor parameters are given by

$$\begin{aligned}
\sigma^2_{\overline{S}} &= \frac{1}{4}\sigma^2_{S_f} + \frac{1}{4}\sigma^2_{S_{\bar{f}}} + \frac{1}{2}\rho_{S_f S_{\bar{f}}}\sigma_{S_f}\sigma_{S_{\bar{f}}}, \\
\sigma^2_{\overline{D}} &= \frac{1}{4}\sigma^2_{D_f} + \frac{1}{4}\sigma^2_{D_{\bar{f}}} + \frac{1}{2}\rho_{D_f D_{\bar{f}}}\sigma_{D_f}\sigma_{D_{\bar{f}}}, \\
\sigma^2_{\Delta S} &= \frac{1}{4}\sigma^2_{S_f} + \frac{1}{4}\sigma^2_{S_{\bar{f}}} - \frac{1}{2}\rho_{S_f S_{\bar{f}}}\sigma_{S_f}\sigma_{S_{\bar{f}}}, \\
\sigma^2_{\Delta D} &= \frac{1}{4}\sigma^2_{D_f} + \frac{1}{4}\sigma^2_{D_{\bar{f}}} - \frac{1}{2}\rho_{D_f D_{\bar{f}}}\sigma_{D_f}\sigma_{D_{\bar{f}}},
\end{aligned} \tag{A.2}$$

where σ is the statistical uncertainty of a given parameter and ρ is the correlation coefficient between two parameters. Using these equations and the values from Table A.2, the expected statistical uncertainties are

$$\begin{aligned}
\sigma_{\overline{S}} = 0.150, \quad \sigma_{\overline{D}} &= 0.098, \\
\sigma_{\Delta S} = 0.083, \quad \sigma_{\Delta D} &= 0.050.
\end{aligned} \tag{A.3}$$

which are exactly what were found when fitting with the $\overline{S}$, $\overline{D}$, ΔS, ΔD parameterisation as shown in Table A.1.

Bibliography

[1] S. Ali *et al.*, *Measurement of time-dependent CP-violation observables in* $B_s^0 \to D_s^{\mp} K^{\pm}$, LHCb Note (2012), LHCb-ANA-2012-068.

[2] LHCb Collaboration, *Measurement of time-dependent* CP*-violation observables in* $B_s^0 \to D_s^{\mp} K^{\pm}$, in *7th International Workshop on the CKM Unitarity Triangle*, Oct, 2012. LHCb-CONF-2012-029.

[3] J. Beringer *et al.*, *2012 Review of Particle Physics*, Physics Review **D 86** (2012).

[4] 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** (1995) 2626.

[5] K. Kodama *et al.*, *Observation of tau neutrino interactions*, Physics Letters B **504** (2001) 218.

[6] T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254.

[7] C. S. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Physical Review Letters **105** (1957) 1413.

[8] I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics*, vol. 2, CRC Press, 3rd ed., 2003.

[9] L. Landau, *On the conservation laws for weak interactions*, Nuclear Physics **3** (1957) 127.

[10] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2 *Meson*, Physical Review Letters **13** (1964) 138.

[11] NA48 Collaboration, *A New Measurement of Direct CP Violation in Two Pion Decays of the Neutral Kaon*, Physical Review Letters **B465** (1999) , CERN-EP/99-114.

[12] BaBar Collaboration, *Observation of CP Violation in the B Meson System*, Phys. Rev. Lett. **87** (2001) 091801, arXiv:hep-ex/0107013v1.

[13] Belle Collaboration, K. Abe *et al.*, *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87** (2001) 091802.

[14] BaBar Collaboration, *Direct CP Violating Asymmetry in $B^0 \to K^+\pi^-$ Decays*, Physical Review Letters **93** (2004) , arXiv:hep-ex/0407057v2.

[15] Belle Collaboration, Y. Chao *et al.*, *Evidence for Direct CP Violation in $B^0 \to K^+\pi^-$ Decays*, Phys. Rev. Lett. **93** (2004) 191802.

[16] M. Adinolfi, Y. Amhis, T. Bauer *et al.*, *The tree-level determination of gamma at LHCb*, LHCb Note (2009), LHCB-ROADMAP5-002.

[17] A. D. Sakharov, *Violation of CP invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz **5** (1967).

[18] A. Riotto, *Theories of Baryogenesis*, tech. rep., CERN, July, 1998. CERN-TH/98-204.

[19] H. E. Haber and G. L. Kane, *The search for supersymmetry: Probing physics beyond the standard model*, Physics Reports **117** (1985) 75.

[20] E. Akhmedov, V. Rubakov, and A. Smirnov, *Baryogenesis via Neutrino Oscillations*, Physical Review Letters **81** (1998) 1359, IC/98/22.

[21] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[22] A. Hocker *et al.*, *CP Violation and the CKM Matrix: Assessing the Impact of the Asymmetric B Factories*, The European Physical Journal C – Particles and Fields **41** (2005) .

[23] LHCb Collaboration, *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays at LHCb*, in *47th Rencontres de Moriond on Electroweak Interactions and Unified Theories*, Mar, 2012. LHCb-CONF-2012-002.

[24] Y. Grossman, A. Soffer, and J. Zupan, *Effect of $D - \bar{D}$ mixing on the measurement of γ in $B \to DK$ decays*, Phys. Rev. D **72** (2005) 031501.

[25] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.

[26] G. V. Dass and K. V. L. Sarma, *Effects of possible $\Delta B = -\Delta Q$ transitions in neutral B meson decays*, Phys. Rev. Lett. **72** (1994) 191.

[27] F. Ambrosino *et al.*, *Study of the branching ratio and charge asymmetry for the decay $K_s \to \pi e \nu$ with the KLOE detector*, Physics Letters B **636** (2006) 173.

[28] The LHCb Collaboration *et al.*, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) S08005.

[29] CERN, *LHCb Technical Proposal*, tech. rep., 1998. CERN-LHCC-98-004.

[30] T. Latham, *Performance of the LHCb Vertex Locator*, in *2nd International Conference on Technology and Instrumentation in Particle Physics*, Nov, 2012. LHCb-PROC-2012-054.

[31] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, tech. rep., CERN, Geneva, Nov, 2012. LHCb-DP-2012-003.

[32] D. Dossett, *Hadronic charmless B decays*, Phenomenology 2012 Symposium: LHC Lights the Way to New Physics, LHCb-TALK-2012-089, 2012.

[33] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Computer Physics Communications **140** (2001) 45.

[34] *GAUSS*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/`.

[35] *BOOLE*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`.

[36] *PYTHIA*, `http://home.thep.lu.se/~torbjorn/Pythia.html`.

[37] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A **462** (2001) 152.

[38] J. Allison *et al.*, *Geant4 developments and applications*, Nuclear Science, IEEE Transactions on **53** (2006) 270.

[39] *MOORE*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/`.

[40] *Brunel*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`.

[41] *DaVinci*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/`.

[42] R. Brun and F. Rademakers, *ROOT – An Object Oriented Data Analysis Framework*, Nuclear Instruments and Methods in Physics Research Section A **389** (1996) 81.

[43] V. V. Gligorov, *A single track HLT1 trigger*, tech. rep., CERN, Geneva, Jan, 2011. LHCb-PUB-2011-003. LHCb-INT-2010-053.

[44] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, tech. rep., CERN, Geneva, Sep, 2011. LHCb-PUB-2011-016. LHCb-INT-2011-030.

[45] L. Breiman, *Bagging predictors*, Machine Learning **24** (1996) 123.

[46] *TMVA*, `http://tmva.sourceforge.net/`.

[47] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research A **555** (2005) 356.

[48] M. Grabalosa and M. Musy, *Flavour Tagging developments within the LHCb experiment*, PhD thesis, Barcelona U., Mar, 2012, CERN-THESIS-2012-075.

[49] LHCb Collaboration, *Optimization and calibration of the LHCb flavour tagging performance using 2010 data*, LHCb Note (2011), LHCb-CONF-2011-003.

[50] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, in *Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, vol. 559, pp. 190–194, April, 2006. doi: 10.1016/j.nima.2005.11.166.

[51] M. Williams, T. Gershon, and S. Easo, *An experimental neural network-based tagger*, tech. rep., CERN, Geneva, Oct, 2011. LHCb-INT-2011-046.

[52] R. Aaij *et al.*, *Measurements of the branching fractions of the decays* $B_s^0 \to D_s^{\mp} K^{\pm}$ *and* $B_s^0 \to D_s^- \pi^+$, Journal of High Energy Physics **2012** (2012) 1, LHCB-PAPER-2011-022.

[53] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, ArXiv e-prints (2009) `arXiv:0905.0724`.

[54] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, Poland, Apr, 1986, DESY F31-86-02(1986).

[55] BABAR Collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP-violating asymmetries and constraints on* $\sin(2\beta+\gamma)$ *with partial reconstruction of* $B \to D^{*\mp}\pi^{\pm}$ *decays*, Phys. Rev. D **71** (2005) 112003.

[56] BABAR Collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP asymmetries in* $B^0 \to D^{(*)\pm}\pi^{\mp}$ *and* $B^0 \to D^{\pm}\rho^{\mp}$ *decays*, Phys. Rev. D **73** (2006) 111101.

[57] The Belle Collaboration, S. Bahinipati *et al.*, *Measurements of time-dependent CP asymmetries in* $B \to D^{*\mp}\pi^{\pm}$ *decays using a partial reconstruction technique*, Phys. Rev. D **84** (2011) 021101.

[58] Belle Collaboration, F. J. Ronga *et al.*, *Measurement of CP violation in* $B^0 \to D^{*-}\pi^{+}$ *and* $B^0 \to D^{-}\pi^{+}$ *decays*, Phys. Rev. D **73** (2006) 092003.

[59] 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`.

[60] LHCb collaboration, R. Aaij *et al.*, *Observation of* B_s^0-$\bar{B}_s^0$ *mixing and measurement of mixing frequencies using semileptonic B decays*, Eur. Phys. J. **C73** (2013) 2655, `arXiv:1308.1302`, LHCB-PAPER-2013-036.