

$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  Decays in the LHCb Detector,  
and a Study of Radiation Damage in the Vertex  
Locator



U N I V E R S I T Y O F  
LIVERPOOL

Thesis submitted in accordance with the requirements of  
the University of Liverpool for the degree of Doctor in Philosophy  
by

**Adrian Andrew Pritchard**

Department of Physics,  
Oliver Lodge Laboratory  
University of Liverpool

October 2015

# *Abstract*

The Large Hadron Collider (LHC) is currently the highest energy particle accelerator in the world, and is designed to collide protons at a centre-of-mass energy up to 14 TeV. The LHCb experiment is one of four main experiments situated on the LHC ring, and is designed for making precision measurements of the decays of particles containing a  $b$  quark.

In order to perform these measurements, it is necessary to precisely measure production and decay vertices in the collisions, and LHCb makes use of the Vertex Locator (VELO) to do this. The need for the VELO to be very close to the proton collision point requires it to be able to withstand high levels of radiation. This thesis presents studies of the damage suffered by the VELO during the full first run period of the LHC, with comparisons to the predicted levels of damage. It is observed that the VELO is performing well despite the radiation damage, and should continue to do so until the planned end of its operation.

This thesis also presents the full analysis of the measurement of the branching fraction of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decay. This measurement is of interest theoretically as it may provide information about  $B_s^0$  oscillations, a phenomenon where  $B_s^0$  mesons spontaneously change to their  $\bar{B}_s^0$  anti-matter counterpart, and back. The analysis is performed on the full LHCb 2011 data set, consisting of approximately  $1.1 \text{ fb}^{-1}$  of proton-proton collisions at centre-of-mass energy of 7 TeV. The branching fraction is measured relative to the decay  $B^0 \rightarrow D_s^\pm D^\mp$  in order to reduce the systematic uncertainty on the result, and is found to be:

$$\mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.07 \pm 0.10(stat.) \pm 0.23(sys.) \pm 0.34(norm.))\%.$$

This result is consistent with, and more precise than, all previous experimental determinations, and also the theoretical prediction.

## Declaration

This thesis is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this, or any other, university. This thesis does not exceed the word limit for the respective Degree Committee.

*Adrian Andrew Pritchard*

# Acknowledgements

*“My CPU is a neural-net processor, a learning computer. The more contact I have with humans, the more I learn”*

***Arnold Schwarzenegger - Terminator 2: Judgement Day***

I have been lucky enough to meet a wide range of people during the past 4 years, all of whom I have learned something from. The list of people I wish to thank for their contribution to the completion of this thesis is almost as long as the thesis itself, but I will try to keep it brief.

First of all, a huge thank you to my PhD supervisor, Dr David Hutchcroft, for being a fantastic teacher and for not once telling me that I was being stupid, despite the fact that at many times during the last four years I undoubtedly was.

Thanks also to the rest of the Liverpool LHCb group for always providing helpful advice and encouragement when it was needed, and to the Liverpool HEP group in general for making the last four years so enjoyable. Thanks to all of the friends I have made during this PhD, both in Liverpool and at CERN, but especially Allan and Matt for always making me laugh (and for chauffering me round the USA and Canada in that beautiful white Mustang).

Thank you to all of my friends from home and from my days as an undergraduate for keeping me (mostly) sane, and for always being there when I regularly needed a pint or ten to relieve physics related stress. There are too many of you to list individually, but I hope you all already know how important you are to me.

A massive thank you to Pratiksha; I was told by numerous people that the last few months of this PhD would be soul-destroying, but thanks to you they have been the most enjoyable of my life so far. I hope I can be equally as helpful during your write up 3 years from now!

I must also thank all of the doctors and nurses from both Alder Hey hospital and the Royal Liverpool hospital that in some way have helped drag me this far through life. There have been some very bleak times in the last 17 years, and without the incredible job they have done I would never have been able to even contemplate completing this PhD.



And lastly, and most importantly, thank you to my parents, Andrew and Jacqueline, and my brother, Darren. I would never have achieved a fraction of what I have without all of the encouragement, support, and love you have given me over the years, and for that I am eternally grateful.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Declaration</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theoretical Framework</b>	<b>3</b>
2.1 The Standard Model . . . . .	3
2.1.1 Particles and Forces . . . . .	3
2.1.2 Quantum Electrodynamics (QED) . . . . .	5
2.1.3 Quantum Chromodynamics (QCD) . . . . .	7
2.1.4 Electroweak Theory . . . . .	8
2.1.5 The Higgs Mechanism . . . . .	10
2.1.6 Problems with the Standard Model . . . . .	13
2.2 Neutral B Mesons . . . . .	14
2.2.1 $B^0$ and $B_s^0$ Oscillations . . . . .	15
2.2.2 $CP$ -Violation . . . . .	18
2.3 Quark Mixing and the Cabibbo-Kobayashi-Maskawa (CKM) Matrix . . .	19
2.3.1 Wolfenstein Parameterisation . . . . .	20
2.3.2 The Unitarity Triangle . . . . .	20
<b>3 The LHC and the LHCb detector</b>	<b>24</b>
3.1 The LHC complex . . . . .	24
3.2 The LHCb detector . . . . .	29
3.2.1 Luminosity levelling . . . . .	31
3.2.2 VELO . . . . .	33
3.2.3 Tracking system . . . . .	33
3.2.4 RICH detectors . . . . .	37
3.2.5 Dipole magnet . . . . .	41
3.2.6 Calorimeters . . . . .	41

3.2.7	Muon chambers . . . . .	43
3.2.8	Trigger and Stripping . . . . .	44
3.2.9	Software chain . . . . .	46
<b>4</b>	<b>Radiation Damage in the Vertex Locator</b>	<b>48</b>
4.1	Semiconductors . . . . .	48
4.2	$p$ - $n$ junctions . . . . .	50
4.2.1	$p$ - $n$ junctions as particle detectors . . . . .	50
4.3	The LHCb vertex locator . . . . .	52
4.4	Radiation damage in silicon detectors . . . . .	59
4.5	Monitoring leakage currents in the VELO . . . . .	61
4.6	Predicting leakage currents . . . . .	64
4.6.1	Sensor fluence profiles . . . . .	65
4.6.2	Annealing . . . . .	67
4.6.3	Temperature corrections . . . . .	68
4.6.4	Results . . . . .	71
4.6.5	Summary and outlook . . . . .	74
<b>5</b>	<b><math>B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}</math> branching fraction</b>	<b>75</b>
5.1	Previous measurements from $B_s^0 \rightarrow D_s^+ D_s^-$ . . . . .	77
5.1.1	$B_s^0 \rightarrow D_s^+ D_s^-$ effective lifetime . . . . .	77
5.1.2	$\phi_s$ from $B_s^0 \rightarrow D_s^+ D_s^-$ . . . . .	78
5.2	$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ . . . . .	80
5.3	Selection . . . . .	83
5.4	Backgrounds . . . . .	87
5.4.1	$B^0 \rightarrow D_s^\pm D^\mp$ . . . . .	88
5.4.2	$\Lambda_b^0 \rightarrow \Lambda_c^+ D_s^-$ . . . . .	92
5.4.3	$B_s^0 \rightarrow D_s(2460)^+ D_s^-$ . . . . .	92
5.4.4	Further potential backgrounds . . . . .	93
5.5	Efficiencies . . . . .	94
5.5.1	Signal channel efficiencies . . . . .	94
5.5.2	Normalisation channel efficiency . . . . .	97
5.5.3	Generator efficiencies . . . . .	98
5.5.4	Relative efficiencies . . . . .	98
5.6	Invariant mass distributions and fits . . . . .	99
5.6.1	$B_s^0 \rightarrow D_s^+ D_s^-$ . . . . .	100
5.6.2	$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$ . . . . .	101
5.6.3	$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$ . . . . .	103
5.6.4	$D_s^\pm$ Peaks . . . . .	104
5.6.5	$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ combinatorial background . . . . .	105
5.6.6	$B_s^0 \rightarrow D_s(2460)^+ D_s^-$ background . . . . .	107
5.6.7	$B^0 \rightarrow D_s^\pm D^\mp$ . . . . .	108
5.6.8	$B^0 \rightarrow D_s^{*\pm} D^\mp$ and $B^0 \rightarrow D_s^\pm D^{*\mp}$ . . . . .	108
5.6.9	$B^0 \rightarrow D_s^{*\pm} D^{*\mp}$ . . . . .	109
5.6.10	$D_s^\pm$ and $D^\pm$ peaks . . . . .	110
5.6.11	$B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$ combinatorial background . . . . .	111
5.7	Data Fits and Yields . . . . .	112

5.7.1	$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ . . . . .	112
5.7.2	$B^0 \rightarrow D_s^\pm D^\mp$ . . . . .	115
5.8	Systematic Uncertainties . . . . .	117
5.8.1	$D_s(2460)$ background shape . . . . .	118
5.8.2	Combinatorial background shape . . . . .	120
5.8.3	Mass PDFs . . . . .	121
5.8.4	$f_s/f_d$ and $D$ meson branching ratio . . . . .	122
5.8.5	Simulation size . . . . .	122
5.8.6	Data sample size . . . . .	123
5.8.7	PIDCalib binning . . . . .	123
5.8.8	Trigger efficiency . . . . .	123
5.8.9	Further tests . . . . .	123
5.8.10	Total systematic uncertainties . . . . .	125
5.9	Results . . . . .	126
5.10	Conclusions . . . . .	127
<b>6</b>	<b>Summary</b>	<b>131</b>
<b>A</b>	<b>BDT Variables MC vs Data</b>	<b>133</b>
<b>B</b>	<b>Toy MC Studies</b>	<b>140</b>
<b>C</b>	<b>Data Fits with Varying BDT Cut</b>	<b>142</b>
<b>D</b>	<b><math>D_s(2460)^+</math> Systematic Study Plots</b>	<b>146</b>
<b>E</b>	<b>Full Data Fit Parameters</b>	<b>148</b>
<b>F</b>	<b>Variable descriptions</b>	<b>150</b>
F.1	BDT Variables . . . . .	150
F.2	Other Variables . . . . .	151
	<b>Bibliography</b>	<b>152</b>

# List of Figures

2.1	The Higgs potential . . . . .	11
2.2	Leading order Feynman diagrams for $B_s^0$ and $B^0$ mixing. . . . .	15
2.3	Experimental constraints on $B$ oscillation parameters . . . . .	17
2.4	CKM Unitarity Triangle . . . . .	21
3.1	LHC accelerator chain . . . . .	26
3.2	LHC dipole cross-section . . . . .	28
3.3	LHCb cross-section . . . . .	29
3.4	Simulation of b quark production . . . . .	31
3.5	Delivered luminosity in the LHC experiments . . . . .	32
3.6	LHCb Tracker Turicensis . . . . .	34
3.7	LHCb tracking system . . . . .	35
3.8	LHCb Inner Tracker . . . . .	36
3.9	Cherenkov angle as a funtion of momentum in RICH1 . . . . .	39
3.10	Simulated Cherenkov angle as a funtion of momentum . . . . .	39
3.11	LHCb RICH detectors . . . . .	40
3.12	LHCb Electromagnetic Calorimeter . . . . .	42
3.13	LHCb Hadronic Calorimeter . . . . .	43
3.14	Trigger flow chart for LHCb . . . . .	44
4.1	Band gap structure in different material types . . . . .	49
4.2	P-N junction as a particle detector . . . . .	51
4.3	Location of modules in the VELO . . . . .	53
4.4	Geometry of the VELO sensors . . . . .	54
4.5	A single VELO module before installation . . . . .	55
4.6	Photograph of multiple VELO modules before installation . . . . .	55
4.7	The VELO RF-Foil . . . . .	57
4.8	The VELO vacuum vessel . . . . .	57
4.9	ADC count distribution in two strip clusters . . . . .	59
4.10	NIEL scaling damage factors . . . . .	61
4.11	VELO sensor I-V characteristics before and after irradiation . . . . .	62
4.12	Leakage current evolution over time for all VELO sensors . . . . .	64
4.13	Fluence profiles in the VELO sensors . . . . .	66
4.14	Annealing results for empirical data fit . . . . .	67
4.15	Thermal camera photograph of a VELO module . . . . .	71
4.16	All VELO sensor currents with the mean predicted current overlaid . . . . .	72
4.17	One VELO sensor current evolution with prediction . . . . .	73
4.18	Current evolution in a second VELO sensor with prediction . . . . .	73

5.1	$B_s^0 \rightarrow D_s^+ D_s^-$ Feynman diagram . . . . .	76
5.2	$B_s^0 \rightarrow D_s^+ D_s^-$ penguin diagram . . . . .	76
5.3	$B_s^0 \rightarrow D_s^+ D_s^-$ effective lifetime plot . . . . .	77
5.4	$B_s^0 \rightarrow D_s^+ D_s^- \phi_s$ plot . . . . .	79
5.5	$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ previous branching ratio measurements . . . . .	83
5.6	BDT efficiency and overtraining plots . . . . .	88
5.7	TMVA background rejection vs signal efficiency . . . . .	89
5.8	BDT FOM and yields vs cut value . . . . .	90
5.9	Misidentified $D^\pm$ plots at various PID cuts . . . . .	91
5.10	$D_s^\pm$ distributions in bins of $B_s^0$ mass . . . . .	95
5.11	$B_s^0 \rightarrow D_s^+ D_s^-$ MC fit . . . . .	101
5.12	Mass fit for $B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$ MC events . . . . .	102
5.13	Double horn structure from $D_s^{*\pm}$ decays . . . . .	102
5.14	Mass fit for $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$ MC events . . . . .	103
5.15	$D_s^\pm$ mass fits for MC events . . . . .	105
5.16	Combinatorial background fit for the signal channel . . . . .	106
5.17	Combinatorial background fit in the signal channel before selection . . . . .	106
5.18	$B_s^0 \rightarrow D_s(2460)^+ D_s^-$ fit to MC events . . . . .	107
5.19	$B^0 \rightarrow D_s^\pm D^\mp$ mass fit to MC events . . . . .	108
5.20	$B^0 \rightarrow D_s^{\pm*} D^\mp$ and $B^0 \rightarrow D_s^\pm D^{\mp*}$ fits from MC events . . . . .	109
5.21	$B^0 \rightarrow D_s^{\pm*} D^{\mp*}$ mass fit from MC events . . . . .	110
5.22	$D_s^\pm$ and $D^\pm$ mass fits from MC . . . . .	111
5.23	Normalisation channel combinatorial background fit . . . . .	112
5.24	Final signal channel data fit . . . . .	113
5.25	Final signal channel data fit for $D_s^\pm$ . . . . .	114
5.26	Final normalisation channel $B^0$ data fit . . . . .	116
5.27	Final normalisation channel $D_s^\pm$ and $D^\pm$ data fits . . . . .	117
5.28	$B_s^0 \rightarrow D_s(2317) D_s^+$ distribution after full selection . . . . .	120
5.29	Branching ratios as a function of BDT cut . . . . .	124
5.30	Branching ratios for $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ with LHCb result included . . . . .	129
A.1	MC vs Data for $B_s^0$ BDT Variables . . . . .	134
A.2	$B_s^0$ impact parameter $\chi^2$ with tight BDT . . . . .	134
A.3	MC vs Data for the first $D_s^\pm$ BDT variables . . . . .	135
A.4	MC vs Data for the first $D_s^\pm$ BDT variables . . . . .	136
A.5	BDT response in data vs MC . . . . .	137
A.6	$D_s^\pm$ lifetime MC vs Data tighter BDT . . . . .	138
A.7	BDT variable signal correlation matrix . . . . .	138
A.8	BDT variable background correlation matrix . . . . .	139
B.1	Toy MC studies for the full signal channel fit . . . . .	141
C.1	Final data plots with BDT > -0.2 cut . . . . .	143
C.2	Final data plots with BDT > -0.15 cut . . . . .	143
C.3	Final data plots with BDT > -0.05 cut . . . . .	144
C.4	Final data plots with BDT > 0.0 cut . . . . .	144
C.5	Final data plots with BDT > 0.05 cut . . . . .	145

---

C.6	Final data plots with $\text{BDT} > 0.1$ cut . . . . .	145
D.1	$D_s(2460)^+$ systematic uncertainty plots . . . . .	147

# List of Tables

2.1	Standard Model fermions . . . . .	4
2.2	Standard Model gauge bosons . . . . .	5
4.1	Parameters used in annealing calculations . . . . .	68
5.1	BDT Variables . . . . .	86
5.2	BDT Variable Ranking . . . . .	87
5.3	Background rejection in multibody channels . . . . .	94
5.4	$D_s^\pm$ fit parameters in bins of $B_s^0$ mass . . . . .	95
5.5	Selection efficiencies for the signal channels . . . . .	96
5.6	Selection efficiencies for the normalisation channel . . . . .	97
5.7	Generator level efficiencies . . . . .	98
5.8	Relative efficiencies for signal to normalisation channels . . . . .	99
5.9	$B_s^0 \rightarrow D_s^{*+} D_s^{*\mp}$ helicity dependence . . . . .	104
5.10	Final signal channel data yields . . . . .	114
5.11	Final normalisation channel data yields . . . . .	117
5.12	$D_s(2460)$ fit parameters . . . . .	118
5.13	$D_s(2460)$ systematic variation . . . . .	119
5.14	$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ yields with Argus parameter variation . . . . .	120
5.15	Combinatorial background systematic using exponential fit . . . . .	121
5.16	Systematic resulting from the fit model . . . . .	122
5.17	$f_s/f_d$ and $D$ meson branching ratio uncertainties . . . . .	122
5.18	Multiple candidate events . . . . .	125
5.19	Total systematic uncertainties in all signal channels . . . . .	126
E.1	Signal channel data fit parameters . . . . .	148
E.2	Normalisation channel data fit parameters . . . . .	149



# Chapter 1

## Introduction

Particle physics is the study of the fundamental particles that constitute the Universe, and the forces that govern how these particles interact with one another. The Standard Model currently provides the most successful and complete theory of nature, having made a number of key predictions that have subsequently been verified experimentally. This thesis begins with a brief overview of the Standard Model and the current status of some key aspects of high energy physics in chapter 2.

Despite being an extremely successful theory, there are still a number of important observations, in both nature and experimentally, that cannot be explained within the current realm of the Standard Model. One of the most important omissions is an explanation of the level of  $CP$ -violation observed in the Universe; there is no understanding of why matter seems to be so dominant over anti-matter. One of the key goals in experimental particle physics is to determine exactly why this is the case, and one of the favoured methods for investigating this is the study of  $B$  hadrons (particles containing a  $b$  or  $\bar{b}$  quark). Chapter 3 examines how the LHC particle accelerator creates the necessary conditions for the creation of  $B$  hadrons, and how the LHCb detector has been designed and optimised to measure these particles.

A characteristic feature of  $B$  hadrons is the relatively long time they exist for before decaying into lighter, more stable particles. This long lifetime means a  $B$  hadron can travel a measurable distance after creation, and this provides a key way of identifying these particles experimentally. Therefore it is vitally important for the efficient performance of LHCb that the points at which particles are created and where they later

decay, known as vertices, can be well measured; this is performed with the Vertex Locator (VELO) detector. Chapter 4 outlines how this detector is designed to help identify the production of these  $B$  hadrons, and also presents a study of the radiation damage suffered by the detector due to its proximity to the particle collisions in the LHC.

LHCb allows the measurement of many different  $B$  hadron decays that could offer some insight into the matter and anti-matter imbalance in the Universe. Chapter 5 presents the measurement of the branching fraction of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decay, and explains why this is an important measurement for LHCb to make.

The thesis concludes with a summary of the two analyses presented in chapter 6, and also further details of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  analysis in the Appendix.

## Chapter 2

# Theoretical Framework

### 2.1 The Standard Model

The Standard Model is currently the most complete theory explaining the fundamental particles that make up the Universe, and also the forces that determine their interactions. It is a Quantum Field Theory using a product of the groups  $SU(3) \times SU(2) \times U(1)$  to simultaneously contain information on the QCD, electroweak, and QED interactions. The remainder of this section firstly outlines the particles and forces involved in the Standard Model, before covering each of the forces in more depth.

#### 2.1.1 Particles and Forces

In the Standard Model, all matter is built from 12 fundamental spin- $\frac{1}{2}$  particles known as fermions, along with their antimatter counterparts. These fermions are separated into two distinct groups with different characteristics, the quarks and the leptons. The quarks are further split into up-type quarks, which carry a charge of  $+\frac{2}{3}e$ , and the down-type quarks, which carry a charge of  $-\frac{1}{3}e$ , where  $e$  is the fundamental charge on the positron. Stable particles and bound states with fractional charge are never experimentally observed, and quarks can only exist in bound states of three quarks or three anti-quarks, known as baryons, or quark and anti-quark pairs, known as mesons. The quarks are also grouped in three generations, with each generation consisting of an up-type and a down-type quark, and the masses increasing with the generation. This

generational structure is also important in weak transitions between quarks, as will be further discussed in section 2.3.

Leptons exist in three pairs, with each pair consisting of a massive particle carrying a charge of  $-1e$  and a massless<sup>1</sup> neutrino with the same flavour that carries no charge. The anti-fermions carry the opposite charge to their fermion partner. The fermions that are comprised within the Standard Model are summarised in table 2.1.

Generation	Quarks		Leptons	
I	up (u)	down (d)	electron (e)	electron neutrino ( $\nu_e$ )
II	charm (c)	strange (s)	muon ( $\mu$ )	muon neutrino ( $\nu_\mu$ )
III	top (t)	bottom (b)	tau ( $\tau$ )	tau neutrino ( $\nu_\tau$ )
Charge	+2/3	-1/3	-1	0

TABLE 2.1: The 12 fundamental fermions in the Standard Model. The charges given are in units of elementary charge.

The fundamental particles detailed in table 2.1 interact in four main ways: via the electromagnetic force, the strong nuclear force, the weak nuclear force, or the gravitational force. The Standard Model describes the first three of these forces through the exchange of particles with integer spin, called bosons. The electromagnetic force acts on any particle carrying charge, and is mediated via the photon. The strong nuclear force acts on all particles carrying a colour charge, which holds together baryons and mesons, and proceeds through the exchange of gluons. The weak nuclear force is responsible for the decay of heavy leptons into lighter ones, and is mediated by either the  $W^\pm$  or  $Z$  bosons. The weak force also governs quark flavour changing processes mediated by a  $W^\pm$  boson, as discussed in more detail in section 2.2.1. No theory of the gravitational force exists in the model, although a spin-2 particle known as a graviton is postulated to mediate this force. The forces of the Standard Model, along with the gauge bosons that are involved in each force, are summarised in table 2.2 and detailed further in the rest of this section.

In the Standard Model, particles acquire mass through interaction with the Higgs field. The Higgs mechanism introduces another boson, the scalar Higgs boson, which has

<sup>1</sup>In the Standard Model neutrinos are massless, although this is known to be incorrect. See section 2.1.6 for further details.

Force	Gauge Boson	Charge
Strong	8 gluons (g)	0
Electromagnetic	photon ( $\gamma$ )	0
Weak	$Z$	0
	$W^\pm$	$\pm 1$

TABLE 2.2: The fundamental interactions in the Standard Model, along with the mediator gauge bosons for each interaction. The charges given are in units of elementary charge.

not been included with the spin-1 gauge bosons in table 2.2. The Higgs mechanism is discussed in more detail in section 2.1.5.

### 2.1.2 Quantum Electrodynamics (QED)

Quantum Electrodynamics (QED) is the field theory that describes electromagnetic interactions, and is mathematically the simplest of the quantum field theories. The theory starts with a free Dirac field,  $\psi(x)$ , which describes the behaviour of a fermion with charge  $Q$  and mass  $m$ , with dynamics described by the Dirac equation [1], shown in 2.1:

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0. \quad (2.1)$$

Here,  $\gamma^\mu$  are the four Dirac gamma matrices. The Dirac equation is derived by applying the Euler-Lagrange equations [2] to the Lagrangian density shown in equation 2.2, where  $\bar{\psi}$  is the complex conjugate field:

$$\mathcal{L}_{Dirac} = \bar{\psi}i\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \quad (2.2)$$

Equation 2.2 is invariant under a global  $U(1)$  gauge transformation of the form shown in 2.3, where  $\alpha$  is a global phase independent of spacetime position  $x$ :

$$\psi(x) \rightarrow e^{i\alpha}\psi(x). \quad (2.3)$$

To transform this to a local gauge invariance requires that  $\alpha$  is allowed to be a function of the spacetime position  $x$ , such that the phases are instead  $\alpha(x)$ . Equation 2.2 does not exhibit local gauge invariance under this transformation; to enable this, the derivative  $\partial_\mu$  must be switched for the covariant derivative  $D_\mu$  shown in 2.4:

$$D_\mu \equiv \partial_\mu + ieA_\mu. \quad (2.4)$$

Here,  $A_\mu$  is a new gauge vector boson field that transforms as shown in 2.5 under a local gauge transformation:

$$A_\mu \rightarrow A'_\mu = A_\mu - \frac{1}{e}\partial_\mu\alpha(x). \quad (2.5)$$

Physically,  $A_\mu$  represents the photon field. The final step is to introduce a gauge invariant kinetic term describing the energy stored within the interacting fields, called the electromagnetic field strength tensor, as shown in 2.6:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (2.6)$$

Replacing the derivative in equation 2.2 with the covariant derivative from 2.4, and adding in the term containing  $F_{\mu\nu}$ , leads to the QED Lagrangian shown in 2.7:

$$\mathcal{L}_{QED} = \bar{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (2.7)$$

The photon field does not introduce any new mass terms to the Lagrangian, and doing so is forbidden as this would violate the gauge invariance of the Lagrangian. This is consistent with experimental observation, where the photon is indeed seen to be massless.

The QED coupling constant depends on the energy scale of the interaction. At a very low energy, the coupling is  $\alpha^{-1} \approx 137$ . However at the  $Z^0$  mass the coupling is instead  $\alpha^{-1} = 127.944 \pm 0.014$  [3].

### 2.1.3 Quantum Chromodynamics (QCD)

The strong force governs the interaction of quarks and gluons, and is described mathematically in a similar way to QED. Quantum Chromodynamics (QCD) is a more complicated theory however, as the field is instead a triplet of Dirac fields as shown in 2.8 (representing the three separate colours), as opposed to the single field in QED:

$$\psi(x) = \begin{pmatrix} \psi_R(x) \\ \psi_G(x) \\ \psi_B(x) \end{pmatrix}. \quad (2.8)$$

QCD is a non-abelian field theory based on the  $SU(N_C)$  gauge group, where  $N_C$  is the number of colours, which has experimentally been shown to be 3. The group has  $N_C^2 - 1$  generators, and these correspond to the 8 different gluons. The QCD Lagrangian is shown in 2.9:

$$\mathcal{L}_{QCD} = \sum_f \bar{\psi}_f(x)(i\gamma_\mu D^\mu - m_f)\psi_f(x) - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}. \quad (2.9)$$

Here, the summation  $f$  refers to the quark flavour, and  $m_f$  is the mass of that particular quark flavour. The covariant derivative,  $D^\mu$ , is now defined as shown in 2.10:

$$D^\mu = \left( \partial_\mu - ig_s \left( \frac{\lambda_a}{2} \right) A_\mu \right). \quad (2.10)$$

The field strength tensor is defined similarly to that for QED, but includes an extra term as shown in 2.11:

$$G_{\mu\nu}^a(x) = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f^{abc} A_\mu^b A_\nu^c. \quad (2.11)$$

In the above equations,  $g_s$  is the strong coupling constant, and  $\lambda_a$  are the 8 Gell-Mann  $\lambda$ -matrices corresponding to the 8 generators of the group. The term  $f^{abc}$  represents the structure constants of the  $SU(3)$  group; the extra terms in the field strength tensor result from the non-abelian nature of the group, and represent the gluon self-couplings. This is a result of the gluons themselves carrying colour charge (the photon in QED

carries no electric charge). There are two important consequences from the non-abelian gauge theory that is used in QCD: *confinement* and *asymptotic freedom*.

$$\alpha_s(Q^2) = \frac{g_s^2}{4\pi}. \quad (2.12)$$

The strong coupling constant varies with the energy scale of the interaction,  $Q^2$ , and as such is generally quoted at specific energy scales, as shown in equation 2.12. For example, at the  $Z^0$  mass the coupling constant has a value of  $\alpha_s(M_Z) = 0.1196 \pm 0.0017$  [3]. At a low  $Q^2$ , the strong force becomes very large. The force between two quarks increases as they are separated, and it becomes energetically favourable to create a quark-antiquark pair before the two original quarks are fully separated. Therefore isolated quarks are never observed, and exist only in bound hadronic states; this is known as *confinement*. Conversely, *asymptotic freedom* is a result of the strong coupling constant becoming very small in high  $Q^2$  interactions. At this scale, quarks and gluons interact very weakly. This ultimately means that at very small distances within hadrons, the constituent quarks are essentially free [4].

#### 2.1.4 Electroweak Theory

Electroweak theory is a unified field theory describing both the weak and electromagnetic interactions, and is produced by requiring that the Lagrangian is locally gauge invariant under transformations relating to the  $SU(2)_L \otimes U(1)_Y$  gauge group. The generators of the group result in weak isospin,  $I$ , and weak hypercharge,  $Y$ , where the weak hypercharge and the third component of weak isospin,  $I_3$ , are related to the charge,  $Q$ , as shown in equation 2.13:

$$Q = I_3 + \frac{Y}{2}. \quad (2.13)$$

Left-handed fermions have  $I = \frac{1}{2}$ , and form doublets with  $I_3 = \pm\frac{1}{2}$ . Up-type quarks and neutrinos take the value of  $I_3 = +\frac{1}{2}$ , whereas down-type quarks and leptons take  $I_3 = -\frac{1}{2}$ . Quarks cannot weakly decay into other quarks with the same value of  $I_3$ , meaning up-type quarks only decay into down-type quarks, and vice versa. Right-handed fermions instead have  $I = 0$ , and therefore form singlet states with  $I_3 = 0$ . It should



be noted that, in the Standard Model, neutrinos are exclusively left-handed (with anti-neutrinos being exclusively right-handed) and are also massless. This is inconsistent with experimental observations, and is one of a number of issues with the Standard Model that will be discussed in brief in section 2.1.6.

The  $SU(2) \times U(1)$  group describing Electroweak interactions introduces 4 separate gauge fields:  $W_\mu^1$ ,  $W_\mu^2$ ,  $W_\mu^3$ , and  $B_\mu$ . The first 3 of these result from the  $SU(2)$  group and are required to maintain invariance under weak transformations, and the other is required to maintain invariance under electromagnetic interactions. These gauge fields are related to the physically observable vector bosons as shown in 2.14:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2)$$

$$Z_\mu = \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \tag{2.14}$$

$$A_\mu = \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu.$$

Here,  $\theta_W$  is the weak mixing angle, also known as the Weinberg Angle. More commonly quoted as  $\sin^2 \theta_W$ , the combined value from D0 and CDF as measured at the  $Z^0$  mass is  $\sin^2 \theta_W = 0.23200 \pm 0.00076$  [3]. The Weinberg Angle is related to the weak isospin coupling constant,  $g$ , and the weak hypercharge coupling constant,  $g'$ , as shown in equation 2.15:

$$\tan \theta_W = \frac{g'}{g}. \tag{2.15}$$

Ensuring the local gauge invariance of the Electroweak Lagrangian requires the gauge bosons to be massless. As with the QED and QCD Lagrangians, adding a mass term to the gauge fields would break the local gauge invariance of the Lagrangian. Experimentally it has been observed that, whilst the photon and the gluons are indeed massless, the weak gauge bosons  $W^+$ ,  $W^-$ , and  $Z$ , are actually massive. In order to unite the

theoretical picture with these experimental observations, it is necessary to have a mechanism to generate mass whilst preserving the gauge invariance of the Lagrangian; in the Standard Model, this is achieved through the Higgs Mechanism.

### 2.1.5 The Higgs Mechanism

The Higgs Mechanism is introduced in order to generate the masses of the weak gauge bosons and the fermions without breaking the gauge invariance of the Lagrangian [5, 6]. This involves a scalar field, known as the Higgs field, that interacts with the Electroweak boson fields and breaks the symmetry of the  $SU(2) \otimes U(1)$  group. The Higgs field has an associated boson called the Higgs Boson; this was observed experimentally by both the CMS [7] and ATLAS [8] experiments.

The Higgs field is an isospin  $SU(2)$  doublet,  $\Phi$ , comprised of two complex scalar fields, as shown in 2.16:

$$\Phi = \frac{1}{2} \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.16)$$

The fields have an associated potential,  $V(\Phi)$ , as shown in 2.17:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \quad (2.17)$$

Here,  $\mu$  and  $\lambda$  are constants. For the case where  $\mu^2 > 0$ , the potential is symmetric about a minimum at  $\Phi = 0$ . However, for the case where  $\mu^2 < 0$ , the potential instead takes the form shown in figure 2.1, with an infinite number of minima as given by equation 2.18:

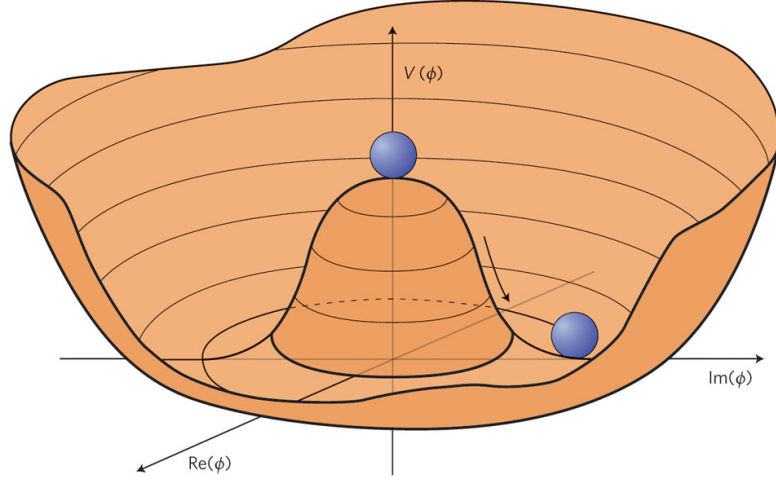


FIGURE 2.1: Shape of the Higgs potential for  $\mu^2 < 0$  [9]. The motion of the blue ball illustrates the Spontaneous Symmetry Breaking process, resulting in a non-zero Vacuum Expectation Value for the Higgs field.

$$|\Phi| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}. \quad (2.18)$$

In equation 2.18,  $v$  is called the Vacuum Expectation Value (VEV). The solutions for the minima of the potential are given by any point on the circle described by equation 2.19:

$$|\Phi_0|^2 = \frac{\phi_1 + \phi_2 + \phi_3 + \phi_4}{2} = \frac{v^2}{2}. \quad (2.19)$$

The vacuum state can be chosen by selecting a particular point on the circle, in a process called Spontaneous Symmetry Breaking (SSB). The vacuum is selected so that the VEVs of  $\phi_1$ ,  $\phi_2$ , and  $\phi_4$  equal 0. The Goldstone Theorem [10] results in each of the four fields having an associated Goldstone boson, three of which are massless and unphysical, with the fourth being the Higgs Boson. It is possible to select a gauge, called the Unitary Gauge, where the three unphysical fields,  $\phi_1$ ,  $\phi_2$ , and  $\phi_4$ , are eliminated. This transforms equation 2.16 into the equation shown in 2.20, where  $H$  is the neutral scalar Higgs field. This also causes the three bosons  $W^+$ ,  $W^-$ , and  $Z^0$  to acquire mass terms:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \quad (2.20)$$

The Lagrangian for the four Electroweak fields  $W_\mu^{1,2,3}$  and  $B_\mu$  introduced in section 2.1.4 can be written to include the interaction with the Higgs field, as shown in 2.21:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2. \quad (2.21)$$

In the Unitary Gauge, this can be written as shown in 2.22:

$$\begin{aligned} \mathcal{L}_{Higgs} = & \frac{1}{2} \partial_\mu H \partial^\mu H + \frac{1}{4} g^2 (H^2 + 2vH + v^2) W_\mu^+ W^{-\mu} \\ & + \frac{1}{8} (g^2 + g'^2) (H^2 + 2vH + v^2) Z_\mu Z^\mu \\ & - \mu^2 H^2 - \frac{\lambda}{4} (H^4 + 4vH^3). \end{aligned} \quad (2.22)$$

Identifying the terms  $M_W^2 W_\mu^+ W^{-\mu}$  and  $\frac{1}{2} M_Z^2 Z_\mu Z^\mu$ , the mass terms corresponding to the  $W^\pm$  and  $Z^0$  bosons are therefore:

$$\begin{aligned} M_W &= \sqrt{\frac{1}{4} g^2 v^2} = \frac{1}{2} g v \\ M_Z &= \sqrt{\frac{1}{4} (g^2 + g'^2) v^2} = \frac{1}{2} (g^2 + g'^2)^{1/2} v \\ &= \frac{1}{2} \frac{g v}{\cos \theta_w} \\ &= \frac{M_W}{\cos \theta_w} \end{aligned} \quad (2.23)$$

where the relations  $(\cos^2 \theta_w + \sin^2 \theta_w) = 1$  and  $g' = g \tan \theta_w$  have been used to express the  $Z^0$  mass in terms of the  $W^\pm$  mass and the Weinberg angle.

In order for the fermions to also acquire mass, a further Yukawa term,  $\mathcal{L}_Y$ , must be added to the Standard Model Lagrangian to describe a coupling between the fermions and the Higgs field. The Lagrangian contribution for each fermion is of the form shown in 2.24:

$$\mathcal{L}_Y(f) = -g_f \bar{\Psi}_f \Psi_f \Phi. \quad (2.24)$$

Here,  $g_f$  represents the Yukawa coupling for a particular fermion. After SSB, this becomes:

$$\mathcal{L}_Y(f) = -\frac{1}{\sqrt{2}}g_f v \bar{\Psi}_f \Psi_f - \frac{1}{\sqrt{2}}g_f \bar{\Psi}_f \Psi_f H. \quad (2.25)$$

The first term gives the fermion mass,  $m_f = \frac{g_f v}{\sqrt{2}}$ . The Yukawa coupling constants are not predicted by the Standard Model, and can be set to match the fermion masses observed in experiment. The second term gives the coupling of the fermion field to the Higgs field, which equals  $\frac{m_f}{v}$ . This means that the coupling to the Higgs field is proportional to the mass of the fermion.

### 2.1.6 Problems with the Standard Model

Although the Standard Model has been hugely successful in both describing and predicting experimental observations, there are still a number of issues preventing the model from being a complete theory of the Universe, a few of which will be discussed here.

*Neutrino masses:* As noted in section 2.1.4, the Standard Model does not include mass terms for any of the neutrino generations. Experimentally, however, neutrinos have been observed to oscillate between flavours [11]. This means that the flavour and mass eigenstates must be different, and means the neutrinos must have a non-zero mass. Two possible mechanisms for generating neutrino masses have been proposed theoretically. Introducing a right-handed neutrino field that interacts only with the Yukawa potential discussed in section 2.1.5, therefore resulting in a sterile neutrino, is one option. The alternative is that neutrinos and anti-neutrinos are the same particle, known as Majorana neutrinos, and that mass terms are produced through the combination of the left handed neutrino field with its complex conjugate. Experimentally neither of these theories has so far been confirmed or disproven.

*Gravity:* The Standard Model does not include any theory of gravitation. It has been postulated that the graviton should be the fundamental boson responsible for mediating the gravitational force, but no experimental signature of the graviton has been observed. To date it has not been possible to mathematically incorporate a field theory describing gravity into the Standard Model.

*Dark matter/Dark energy:* Observations from cosmology indicate that the particles described by the Standard Model account for only approximately 5% of the energy in the Universe. The rest is thought to be made up of dark matter, which would interact very weakly with the forces in the Standard Model, and also dark energy. There are no particles in the Standard Model that fit the characteristics required for dark matter, and there has been no successful theory explaining dark energy in the realm of the Standard Model.

*Matter and anti-matter imbalance:* The Standard Model does not account for the scale of matter dominance over anti-matter in the Universe. It is generally accepted that the Big Bang should have created equal amounts of matter and anti-matter, and the existence of  $CP$ -violation subsequently caused the excess of matter currently observed. Investigating potential sources of  $CP$ -violation is a key goal in experimental particle physics; currently the only observed source of  $CP$ -violation is the Cabibbo-Kobayashi-Maskawa matrix that governs the weak transitions of quarks. Studying the decays of neutral particles containing  $b$  quarks is a key experimental technique for assessing the levels of  $CP$ -violation in the quark sector.

## 2.2 Neutral B Mesons

Neutral  $B$  mesons are formed from the pairing of a  $b$  quark with either an  $s$  anti-quark or a  $d$  anti-quark (or a  $b$  anti-quark with an  $s$  or  $d$  quark). The study of  $B$  mesons is of great interest experimentally, as the  $b$  quark is the heaviest quark that has been observed to hadronise. This means that  $B$  hadrons have a wider number of decays than any other type of observed hadron, thus providing the largest number of channels for testing the Standard Model. Numerous ways exist in which New Physics could be observed in the decay of  $B$  mesons. The Cabibbo favoured transition for a decaying  $b$  quark is to a  $c$  quark, but it is also possible, albeit much less likely, for the  $b$  quark to instead decay to a first generation quark (see section 2.3 for more details about quark transitions). Some of these decays proceed via loop diagrams and are therefore interesting to study, as these loops give access to possible undiscovered particles. Comparing decay rates in these channels to Standard Model predictions, or to other tree-level  $B$  meson decays, is a key way of searching for potential new particles [12].

The decay of  $B$  mesons is also of interest for a number of other reasons; these will be the focus of the remainder of this chapter.

### 2.2.1 $B^0$ and $B_s^0$ Oscillations

The spontaneous oscillation of  $B^0$  mesons to their  $\bar{B}^0$  antiparticle counterpart (and similarly for  $B_s^0$  and  $\bar{B}_s^0$  mesons), and vice versa, was a fundamental prediction from the Standard Model. This phenomenon was first observed in  $B^0$  mesons by the ARGUS collaboration in 1987 [13], and, in 2006, the CDF experiment was also able to determine the rate of oscillation in the  $B_s^0$  system [14]. The leading order Feynman diagrams for this process are shown in figure 2.2.

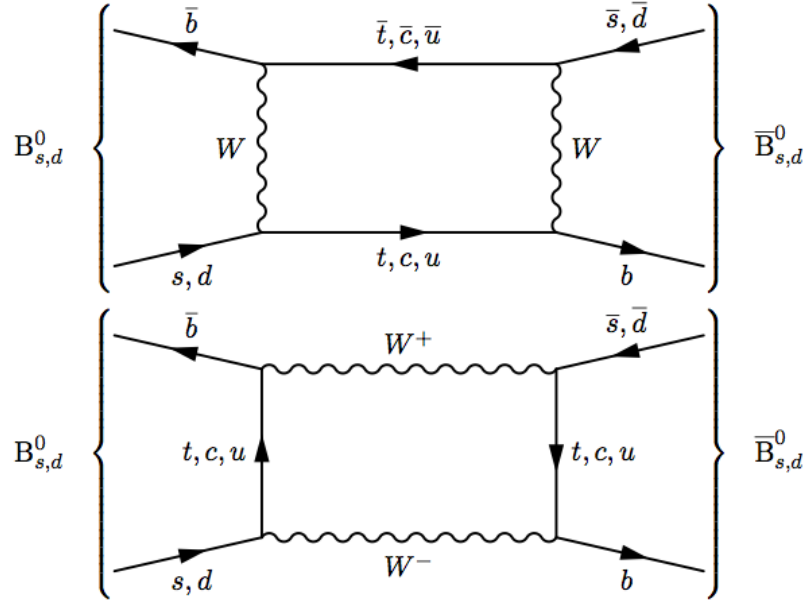


FIGURE 2.2: Leading order Feynman diagrams for  $B_s^0$  and  $B^0$  mixing.

This mixing means that the  $B_s^0$  meson mass eigenstates are different to the flavour eigenstates; the mass eigenstates are instead a linear combination of the  $B_s^0$  and  $\bar{B}_s^0$ , as shown in equation 2.26, where  $H$  and  $L$  refer to *heavy* and *light* mass eigenstates respectively:

$$|B_{L,H}\rangle = p|B_s^0\rangle \pm q|\bar{B}_s^0\rangle. \quad (2.26)$$

The time evolution of the  $B_s^0$  system described in equation 2.26 is given by the solution to the time dependent Schrödinger equation shown in 2.27:

$$i \frac{d}{dt} \begin{pmatrix} B_s^0(t) \\ \bar{B}_s^0(t) \end{pmatrix} = \mathbf{H} \begin{pmatrix} B_s^0(t) \\ \bar{B}_s^0(t) \end{pmatrix}. \quad (2.27)$$

Here,  $H$  is a  $2 \times 2$  effective Hamiltonian, which can be written in terms of the Hermitian mass and lifetime matrices,  $\mathbf{M}$  and  $\mathbf{\Gamma}$ , as shown in 2.28:

$$\mathbf{H} = \left( \mathbf{M} - \frac{i}{2} \mathbf{\Gamma} \right). \quad (2.28)$$

The diagonal elements of  $\mathbf{M}$  and  $\mathbf{\Gamma}$  represent flavour-conserving transitions, but the off-diagonal elements represent the oscillation of a  $B_s^0$  to  $\bar{B}_s^0$ , and the reverse. Solving the equation in 2.27 gives the ratio  $q/p$  in terms of the off-diagonal elements of  $\mathbf{M}$  and  $\mathbf{\Gamma}$ , as shown in equation 2.29:

$$\left( \frac{q}{p} \right)^2 = \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}}. \quad (2.29)$$

Any deviation from unity in the value of  $q/p$  would be a clear indication of  $CP$ -violation in  $B_s^0$  oscillations. The  $B_s^0$  system is commonly described in terms of the mass difference,  $\Delta m_s$ , the decay width difference,  $\Delta \Gamma_s$ , and the average decay width,  $\Gamma_s$ , of the two mass eigenstates, which are all defined in equation 2.30:

$$\begin{aligned} \Delta m_s &= m_H - m_L \\ \Delta \Gamma_s &= \Gamma_L - \Gamma_H \end{aligned} \quad (2.30)$$

$$\Gamma_s = \frac{\Gamma_L + \Gamma_H}{2}.$$

Measuring these values experimentally can help determine the level of  $CP$ -violation in the  $B_s^0$  environment; this is performed generally in one of two ways. The first method



involves selecting a final state that is a (near) pure  $CP$  eigenstate, and thus accessible through decays of only one of the  $B_H$  or the  $B_L$ , such as  $B_s^0 \rightarrow K^+ K^-$  [15] ( $CP$ -Even) or  $B_s^0 \rightarrow J/\psi f^0$  [16] ( $CP$ -Odd). A measurement of the effective lifetime of the  $B_s^0$  in one of these channels then translates as a measurement of  $\Gamma_H$  or  $\Gamma_L$ . The second method requires a time-dependent analysis of a channel that is accessible to both  $B_s^0$  and  $\bar{B}_s^0$  mesons. In this case, the interference between the direct decay amplitude and the amplitude for oscillation and then decay gives rise to a time dependent  $CP$ -violating asymmetry that can be expressed in terms of  $\Delta\Gamma_s$  and also a  $CP$ -violating phase,  $\phi_s$  [17]; this is the method employed in the analysis of decays such as  $B_s^0 \rightarrow J/\psi \phi$  [18].

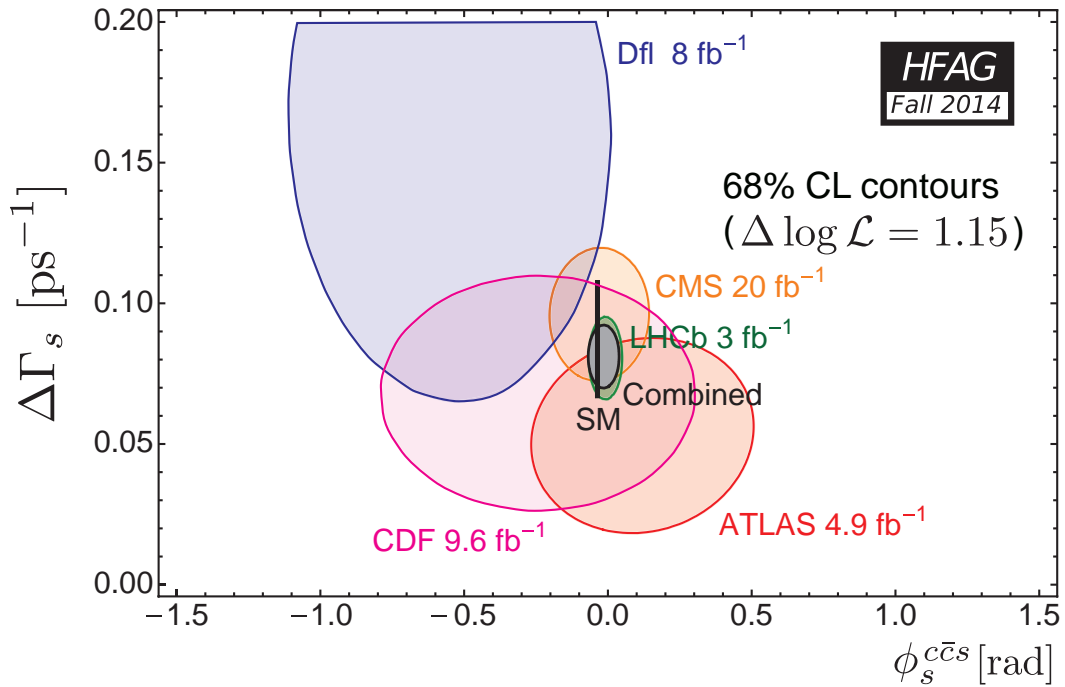


FIGURE 2.3: Current status of experimental measurements in the  $\Delta\Gamma_s$ - $\phi_s$  plane, where the contours represent a 68% confidence level. The grey circle shows the weighted average of all experimental results, and the thin black rectangle gives the Standard Model prediction. Here,  $\phi_s^{c\bar{c}s}$  means that  $\phi_s$  is determined from an average of measurements where the quark transition is  $b \rightarrow c\bar{c}s$  [19].

Figure 2.3 shows the experimental status of measurements of  $\Delta\Gamma_s$  and  $\phi_s$ , along with the world average as produced by the Heavy Flavour Averaging Group (HFAG) [19]. The Standard Model prediction for  $\Delta\Gamma_s$  is taken from [20], and the prediction for  $\phi_s$  is taken from [21]. The world average is consistent with Standard Model expectation, and

the small measured value of  $\phi_s = -0.015 \pm 0.035$  rad implies  $CP$ -violation in  $B_s^0$  mixing is minimal.

### 2.2.2 $CP$ -Violation

$CP$  symmetry, the combination of Charge conjugation and Parity reversal, was originally thought to be an exact symmetry of nature, meaning that particles and anti-particles would behave in exactly the same way. The apparent domination of matter over anti-matter in the Universe implied that this was not the case, and the possibility of  $CP$  symmetry being violated was postulated as a potential explanation for the matter and anti-matter imbalance.  $CP$ -violation was first observed experimentally in the neutral kaon system [22], and has subsequently been observed in other particle/anti-particle systems.

$CP$ -violation can occur in any of three distinct ways in the  $B_s^0$  system;  $CP$ -violation in decay,  $CP$ -violation in mixing, and  $CP$ -violation in interference between mixing and decay. These are described in brief below:

*$CP$ -violation in decay:* The first kind of  $CP$ -violation requires that the decay amplitudes of  $B_s^0 \rightarrow f$  and  $\bar{B}_s^0 \rightarrow \bar{f}$ , where  $f$  is any final state accessible to the  $B_s^0$  and  $\bar{f}$  is the equivalent final state for the  $\bar{B}_s^0$ , are different, as shown in equation 2.31:

$$\left( \frac{\bar{A}_{\bar{f}}}{A_f} \right) \neq 1. \quad (2.31)$$

Here,  $\bar{A}_{\bar{f}}$  is the decay amplitude of  $\bar{B}_s^0 \rightarrow \bar{f}$ , and  $A_f$  is the decay amplitude of  $B_s^0 \rightarrow f$ .

*$CP$ -violation in mixing:* This would require the oscillation of a  $B_s^0 \rightarrow \bar{B}_s^0$  to have a different probability to the reverse oscillation  $\bar{B}_s^0 \rightarrow B_s^0$ , as expressed in equation 2.32:

$$\left| \frac{q}{p} \right| \neq 1. \quad (2.32)$$

*$CP$ -violation in interference between mixing and decay:* This would occur if there was a complex phase between the decay amplitude for  $B_s^0 \rightarrow f$  and  $B_s^0 \rightarrow \bar{B}_s^0 \rightarrow f$ ,

where  $f$  is any final state accessible to both  $B_s^0$  and  $\bar{B}_s^0$  mesons. This can be expressed mathematically as shown in equation 2.33:

$$\mathcal{I}m\left(\frac{q}{p}\frac{\bar{A}_f}{A_f}\right) \neq 0. \quad (2.33)$$

## 2.3 Quark Mixing and the Cabibbo-Kobayashi-Maskawa (CKM) Matrix

The fact that quark flavour and generation are not conserved in weak decays implies that the weak flavour eigenstates and the strong mass eigenstates of the quarks are not the same. The relationship between the flavour eigenstates,  $d'$ ,  $s'$ , and  $b'$ , and the mass eigenstates,  $d$ ,  $s$ , and  $b$ , is shown in equation 2.34:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.34)$$

Here,  $V_{CKM}$  is a  $3 \times 3$  unitary matrix that describes the mixing between quark generations, known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix; this is shown in equation 2.35:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (2.35)$$

In the CKM matrix, each of the nine elements describes the coupling between two separate quark flavours, *e.g.*  $V_{cd}$  is the coupling between a  $c$  quark and a  $d$  quark. The CKM matrix can be written entirely in terms of three quark mixing angles,  $\theta_{1,2}$ ,  $\theta_{1,3}$ , and  $\theta_{2,3}$ , and a single  $CP$ -violating phase  $\delta$ , as shown in equation 2.36:

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \quad (2.36)$$

In this representation,  $c_{ij} = \cos \theta_{ij}$  and  $s_{ij} = \sin \theta_{ij}$ . If the three angles  $\theta_{ij}$  were all zero, this would produce the identity matrix, which would prevent quark mixing from occurring. It is also interesting to note that removing the third generation of quarks would reproduce the original Cabibbo matrix; the original two-generation quark mixing model required only one mixing angle, the Cabibbo angle  $\theta_C$ , but also allowed no possibility of  $CP$ -violation.

### 2.3.1 Wolfenstein Parameterisation

The experimentally determined relationship of the hierarchy of the quark transition strengths,  $s_{13} \ll s_{23} \ll s_{12} \ll 1$ , allows the CKM matrix to be parameterised as a power series expansion in  $\lambda$ , where  $\lambda = \sin \theta_{12}$  [23]. Expanding to  $\mathcal{O}(\lambda^3)$  yields the form of the matrix shown in equation 2.37:

$$\begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \quad (2.37)$$

As experiments have become more accurate, it has become important to expand the parameterisation to higher orders in  $\lambda$ . To this end, it is useful to switch from  $(\rho, \eta)$  space to  $(\bar{\rho}, \bar{\eta})$  space, where  $\bar{\rho}$  and  $\bar{\eta}$  are defined [24] in equation 2.38:

$$\begin{aligned} \bar{\rho} &= \rho \left( 1 - \frac{\lambda^2}{2} \right) \\ \bar{\eta} &= \eta \left( 1 - \frac{\lambda^2}{2} \right). \end{aligned} \quad (2.38)$$

### 2.3.2 The Unitarity Triangle

The unitarity of the CKM matrix requires that the relationships  $\sum_i V_{ij} V_{ik}^* = \delta_{jk}$  and  $\sum_j V_{ij} V_{kj}^* = \delta_{ik}$  are satisfied. These relationships yield 6 orthogonality relationships, which can be represented diagrammatically as triangles in the complex plane. Of the 6,



In the triangle in figure 2.4, the angles  $\alpha$ ,  $\beta$ , and  $\gamma$  are related to the CKM elements as show in equation 2.41:

$$\begin{aligned}\alpha &= \arg \left( -\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right) \\ \beta &= \arg \left( -\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right) \\ \gamma &= \arg \left( -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right).\end{aligned}\tag{2.41}$$

The angles in equation 2.41 can be constrained with measurements of  $CP$ -violating variables in  $B$  meson decays. The elements of the CKM matrix can also be determined experimentally with various measurements; the current experimental results for the three angles and the nine CKM elements are shown in 2.42 and 2.43:

$$\begin{pmatrix} 0.97427 \pm 0.00014 & 0.22536 \pm 0.00061 & 0.00355 \pm 0.00015 \\ 0.22522 \pm 0.00061 & 0.97343 \pm 0.00015 & 0.0414 \pm 0.0012 \\ 0.00886^{+0.00033}_{-0.00032} & 0.0405^{+0.0011}_{-0.0012} & 0.99914 \pm 0.00005 \end{pmatrix}\tag{2.42}$$

$$\begin{aligned}\alpha &= (85.4^{+3.9}_{-3.8})^\circ \\ \beta &= (21.6^{+0.75}_{-0.74})^\circ \\ \gamma &= (68.0^{+8.0}_{-8.5})^\circ.\end{aligned}\tag{2.43}$$

In the case of  $B_s^0$  mesons the angle  $\beta$  is instead referred to as  $\beta_s$ , and is related to the  $B_s^0$   $CP$ -violating phase through the relationship shown in 2.44:

$$\phi_s = -2\beta_s.\tag{2.44}$$

In the Standard Model,  $\phi_s$  is predicted to be small, meaning that  $CP$ -violation is predicted to be smaller in the  $B_s^0$  system than in the  $B^0$  system. The angle  $\beta_s$  can be measured through a time dependent analysis of a decay to a  $CP$  eigenstate, such as the  $CP$  even decay  $B_s^0 \rightarrow D_s^+ D_s^-$ . This will be discussed further in chapter 5.

For a detailed review of the experimental measurements used to extract the CKM matrix elements, and to constrain the three angles, the reader is directed to [26]. The precise determination of the CKM matrix elements and Unitarity Triangle angles is one of the key goals of the LHCb experiment, as this provides a stringent test of the Standard Model. Any indication that the CKM matrix is not unitary, which would result from the squares of the moduli of an individual row or column of the matrix not summing to 1, or alternatively the sum of the three Unitarity Triangle angles not equalling  $180^\circ$ , would be a clear indication of New Physics not accounted for by the Standard Model. Alternatively, discrepancies between various determinations of the same CKM parameter can also be an indication of New Physics [27].

## Chapter 3

# The LHC and the LHCb detector

Particle accelerators have long been used in experimental particle physics to reach the energies needed to produce higher mass particles in collisions. Two main types of particle accelerator are used: linear accelerators (LINACs), and synchrotrons.

In a LINAC, particles are accelerated in a straight line with the use of high energy alternating current electric fields. Most modern LINACs make use of radio frequency (RF) cavities to provide the field as, when the accelerating particle reaches a high speed, the polarity of the field must change many times a second to efficiently continue to accelerate the particle. The maximum achievable energy in a LINAC is governed by the accelerating length.

Synchrotrons overcome this limitation by allowing the particles to be accelerated in a continuous loop by using magnets to bend the path around in a circular motion. The majority of modern particle accelerators are synchrotrons, as these allow access to higher energies without the same space considerations that would come from a LINAC.

### 3.1 The LHC complex

The Large Hadron Collider (LHC) is currently the highest energy synchrotron in the world. At maximum design energy, the LHC will collide protons with a centre of mass collision energy of  $\sqrt{s} = 14$  TeV, although the analyses presented in chapters 4 and 5 use data taken during 2011 and 2012, at centre-of-mass energy  $\sqrt{s} = 7$  TeV and 8 TeV respectively. Situated beneath the Franco-Swiss border, the LHC occupies the same 27 km



tunnel that previously housed the Large Electron Positron (LEP) collider. Compared to electrons, colliding protons has a number of advantages. Firstly, colliding protons allows access to a much wider range of invariant masses at the same time; this allows for much broader physics studies than those possible with the collision of electrons, which are best suited to precision studies.

Another advantage of accelerating protons instead of electrons is that much higher collision energies can be achieved, as protons lose less energy through synchrotron radiation. Synchrotron radiation is electromagnetic radiation emitted by charged particles during acceleration, which causes the particle to lose an amount of energy per revolution of radius  $r$  governed by equation 3.1 [28]:

$$\Delta E = \frac{q^2 \beta^3 \gamma^4}{3\epsilon_0 r}. \quad (3.1)$$

Here,  $q$  is the charge of the particle,  $\epsilon_0$  is the permittivity of free space, and  $\beta$  and  $\gamma$  are the relativistic variables. As the particle approaches the speed of light,  $\gamma = E/m$ , and thus the energy loss is inversely proportional to the mass of the particle to the fourth power. This means that protons lose much less energy per revolution than electrons, and can therefore reach much higher energies during acceleration in a synchrotron.

The maximum achievable energy of a proton synchrotron is governed by the radius of the synchrotron, and also by the strength of the dipole magnets that are used to steer the protons around the ring. As the radius of the ring was already set, the key to maximising the energy of protons in the LHC was to use dipole magnets of extremely high field strength; to this end, the LHC makes use of 1232 superconducting dipole electromagnets. In addition to the dipole magnets, the LHC also contains 392 quadrupole magnets that prevent the proton beam diverging as it traverses the ring, and a number of higher order pole magnets to make more minor corrections to the beam.

To enable the LHC to effectively accelerate protons, they must be already travelling at a high energy when they enter. To achieve this, there are a number of smaller accelerators involved in accelerating the protons before they enter the LHC; the full chain of accelerators is shown in figure 3.1.

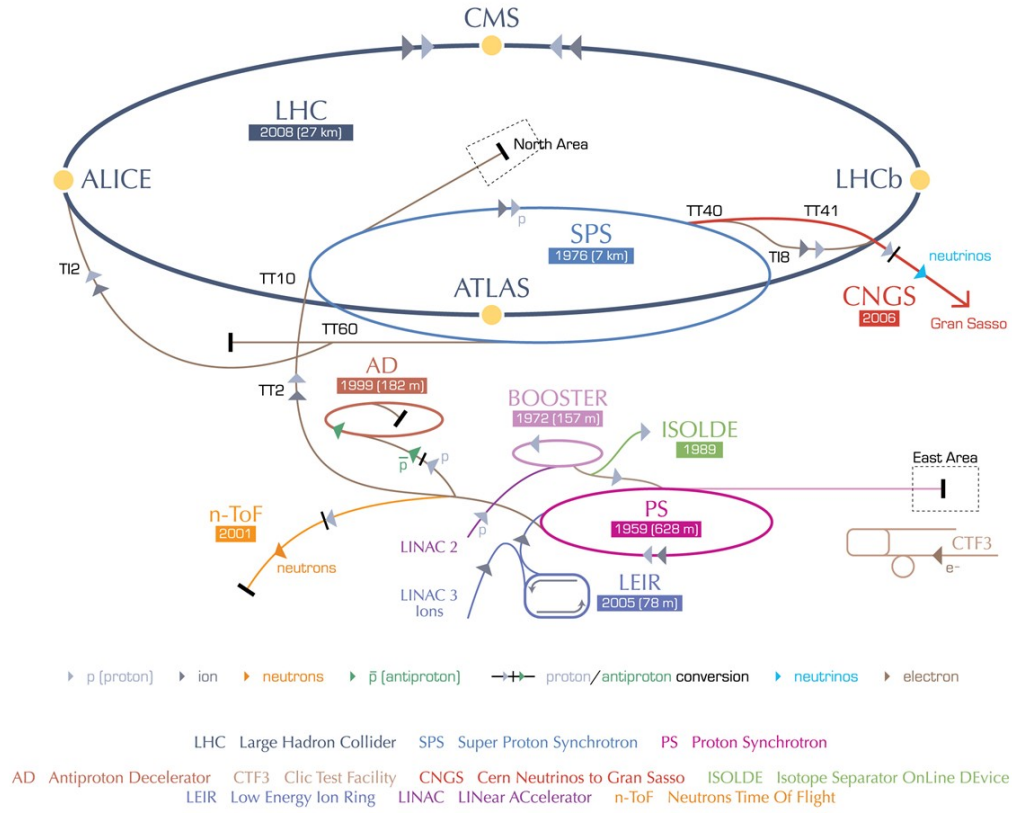


FIGURE 3.1: The complete chain of pre-accelerators for the LHC [29].

The protons for acceleration are obtained by applying a large electric field to hydrogen gas, in order to strip the electrons from the hydrogen atoms and leave only the protons. These protons are then passed in pulses into the first of the pre-accelerators for the LHC, LINAC-2. This is a linear accelerator 30 m in length, which uses radiofrequency (RF) cavities to accelerate the protons to an energy of 50 MeV.

From here, the protons are passed into the Proton Synchrotron Booster (PSB). The PSB is the first of the synchrotrons involved in accelerating the protons to the LHC injection energy and consists of four identical rings stacked on top of each other, each 157 m in circumference. Each of the four rings accelerates  $1.05 \times 10^{12}$  protons simultaneously, reaching an energy of 1.4 GeV.

Upon leaving the PSB, the protons pass into the Proton Synchrotron (PS). The PS, with a circumference of 628 m, was the first of CERN's synchrotrons and originally

the laboratory's flagship accelerator, but has since been modified to serve as a pre-accelerator of either protons or heavy ions for eventual transfer to the LHC. When accelerating protons, the PS takes two injection pulses from the PSB, with each pulse containing approximately  $0.83 \times 10^{13}$  protons. The PS makes use of 277 conventional electromagnets, including 100 dipole magnets, to guide the protons around the ring, and accelerates the protons to an energy of 25 GeV. During this acceleration stage the protons are separated into 72 bunches, with each bunch containing approximately  $10^{11}$  protons.

After reaching the maximum energy of the PS, the protons are transferred to the penultimate stage of acceleration, the Super Proton Synchrotron (SPS). Previously a particle collider in its own right, facilitating the discovery of the  $W$  and  $Z$  bosons in 1983, the SPS is now used to increase the energy of protons to a sufficient level to allow injection into the LHC. The SPS is approximately 7 km in circumference and, with the assistance of 1317 electromagnets (including 744 bending dipoles), accelerates the protons to an energy of 450 GeV. The bunches formed in the PS are maintained in the SPS, but with some deliberate gaps between the bunches. These gaps enable the powering of specially designed dump 'kicker' magnets, which can eject the proton beam during acceleration if required.

At this point the protons are travelling with enough energy to allow them to be efficiently accelerated by the LHC, and are injected in two beams travelling in opposite directions. To allow the acceleration of both beams, the LHC uses two beam pipes. Cost and space considerations inside the LHC tunnel meant that the two beam pipes must be inside the same ring; this is illustrated in the LHC cross section shown in figure 3.2. This means that the magnet structure inside the LHC is relatively complex, as the magnetic field provided must simultaneously control both the clockwise and anti-clockwise proton beams. This structure does ensure that only one cooling system is required; this is very cost effective, given that the superconducting magnets must be cooled to 1.7K using superfluid Helium.

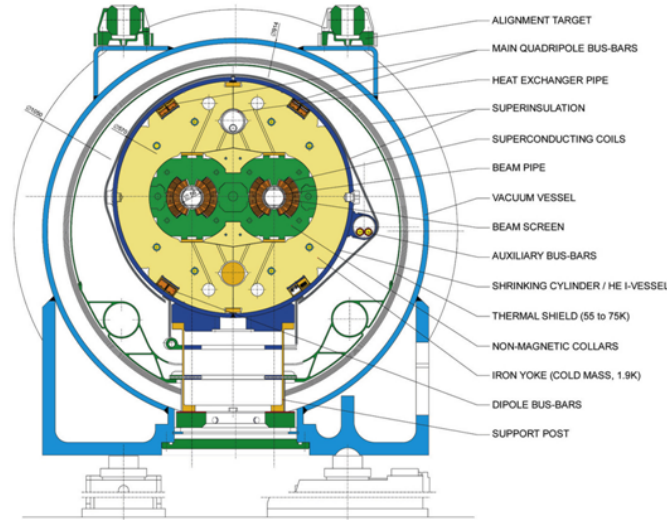


FIGURE 3.2: Illustration of a cross-section of one of the LHC dipoles [30].

To help ensure the Helium is not warmed during operation, the LHC must be thermally isolated from the surrounding tunnel. This is performed by using both a reflective thermal shield around the outside of the beam line, and also a vacuum layer that prevents heat being conducted into the beam line. Each beam pipe is also kept under ultra high vacuum conditions to prevent the circulating proton bunches colliding with gas molecules during acceleration. At full design capacity the LHC will contain 2808 bunches of protons separated by 25 ns, with each bunch containing approximately  $10^{12}$  protons. During the 2011 running, the LHC instead circulated 1380 proton bunches, with 50 ns spacing [31].

The LHC beam line consists of 8 straight sections and 8 bending sections. Each of the straight sections has the potential to be used as a crossing point for the proton beams, but four of these points are reserved to allow for beam cleaning and beam dumping. Beam cleaning ensures the superconducting magnets are not damaged by protons straying from the bunches, whereas a beam dump allows for extraction of the proton beam from the LHC at the end of a run, or if the beam becomes unstable.

The remaining four straight sections around the LHC are where the two proton beams are made to collide. Prior to collision, the beams are passed through a series of magnets to ensure the transverse spread of the proton bunches is as small as possible, in order to increase the chances of collisions occurring when two bunches cross. Each collision

point is surrounded by a specialised particle detector. Two of these detectors, CMS and ATLAS, are General Purpose Detectors (GPDs). These detectors are designed to study a broad range of particle decays, searching for evidence of previously undiscovered particles and New Physics. A third detector, ALICE, is designed to study quark-gluon plasma, a form of matter that was present very shortly after the Big Bang. ALICE primarily studies the collisions of lead ions in the LHC, and as such only takes data for a short period of time each year.

The last of the main experiments situated on the LHC is the LHCb experiment. This experiment, and the various components of the particle detector used in the experiment, will be the focus of the remainder of this chapter.

### 3.2 The LHCb detector

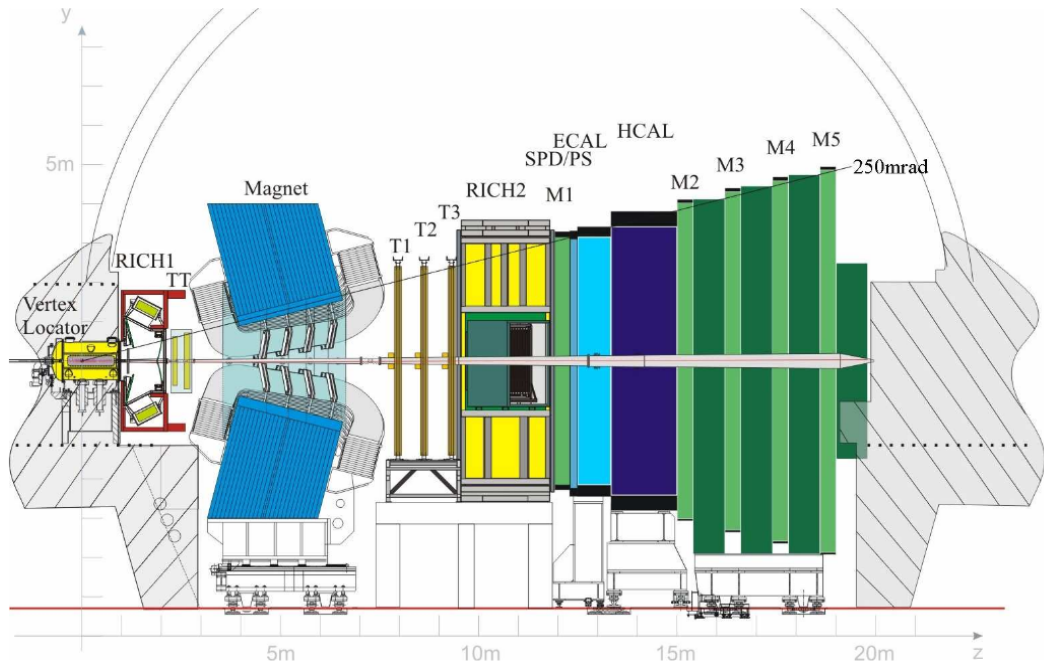


FIGURE 3.3: View of the LHCb detector in the  $y$ - $z$  plane; particles traverse the detector from left to right in the image [32].

The LHCb detector is a single-arm spectrometer, designed to make precision studies of decays of particles containing  $b$  quarks at the LHC. As shown in figure 3.4,  $b$  quarks are

produced predominately in the forward or backward directions in  $p$ - $p$  collisions. For this reason, LHCb covers only a forward angle between 10 mrad – 300 mrad in the bending plane of the magnet, and 10 mrad – 250 mrad in the non-bending plane. This corresponds to a pseudorapidity range  $1.8 < \eta < 4.9$ , where pseudorapidity is defined in equation 3.2, meaning LHCb is well positioned to provide complimentary measurements to those that will be made at ATLAS and CMS:

$$\begin{aligned}\eta &= \frac{1}{2} \ln \frac{|P| + P_z}{|P| - P_z} \\ &= -\ln \left( \tan \frac{\theta}{2} \right).\end{aligned}\tag{3.2}$$

Here,  $|P|$  is the modulus of the momentum,  $P_z$  is the component of the momentum in the  $z$  direction, and  $\theta$  is the angle that the particle track makes with the beamline. A schematic of the LHCb detector is shown in figure 3.3. The detector makes use of a number of highly specialised sub-detectors in order to reconstruct the decays of  $B$  hadrons. Using information from the VELO, tracking stations, calorimeters, and muon stations allows for the tracking of charged particles through the detector and enables precise vertex reconstruction. The RICH detectors, along with the calorimeters and muon stations, provide particle identification and allow the reconstruction of kaons and pions; this is especially important, as many interesting  $B$  hadron decays contain these particles in the final state. These decays are more difficult to identify in the GPDs, where methods to distinguish between kaons and pions must be more intricate, but LHCb is optimised for making these measurements. The rest of this chapter details the individual components of the LHCb detector, and explains why the detector is well suited to performing these studies.

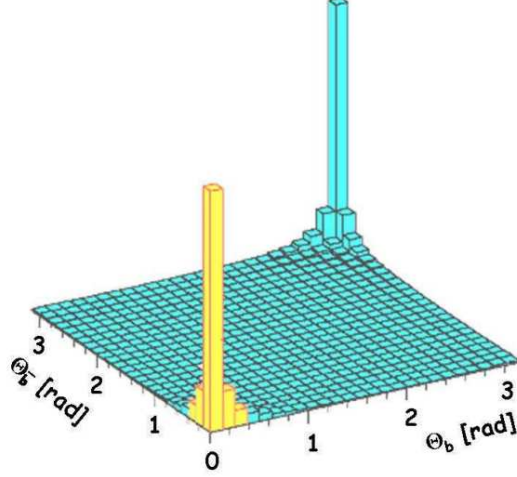


FIGURE 3.4: Pythia simulation of the relative  $b\bar{b}$  production cross section as a function of the  $b$  and  $\bar{b}$  polar angles from the primary vertex [33].

### 3.2.1 Luminosity levelling

The maximum instantaneous luminosity of a collider is governed by equation 3.3 [34]:

$$\mathcal{L} = f \frac{N_1 N_2}{4\pi\sigma_x\sigma_y}. \quad (3.3)$$

$N_1$  and  $N_2$  are the number of particles in each of the colliding bunches,  $f$  is the frequency of the beam crossing, and  $\sigma_x$  and  $\sigma_y$  govern the beam profile. At the LHC, a high luminosity is required by both the ATLAS and CMS experiments, where many of the physics analyses by these experiments are searching for processes with very small predicted branching fractions, and therefore rely heavily on having a large amount of data. Having a large luminosity has the effect of increasing the number of collisions per bunch crossing,  $\mu$ , and, whilst beneficial for the GPDs, a higher  $\mu$  is actually detrimental to the performance of LHCb. It is vitally important that LHCb is able to accurately reconstruct the secondary and tertiary decay vertices that are characteristic of the decays of  $B$  hadrons and  $D$  mesons, and correlate these with the correct primary vertices. A larger number of interactions in each bunch crossing has the effect of increasing the number of particles passing through the detector, which in turn makes it much harder to reconstruct the decay vertices. For this reason, LHCb is designed to run at a luminosity

a factor of 50 lower than the GPDs,  $2 \times 10^{32} \text{ cm}^2 \text{ s}^{-1}$ , and this is achieved by offsetting the beams as they collide in the detector; this process is known as luminosity levelling.

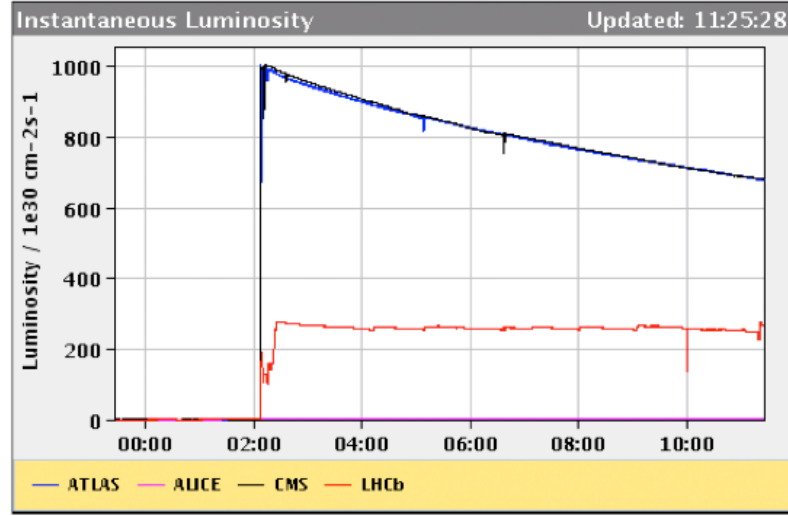


FIGURE 3.5: Instantaneous luminosity delivered to ATLAS, CMS, and LHCb over the course of a fill. The luminosity in the GPDs drops off with time as the bunches lose protons during collisions, whereas LHCb maintains a constant luminosity by adjusting the beams during the course of the fill [35].

The beams are constantly adjusted to account for the loss of protons from each bunch as they circulate and collide over the course of a fill. By gradually bringing the beams closer together, the instantaneous luminosity can be kept constant in LHCb; this is shown in figure 3.5. This levelling makes the determination of delivered luminosity much easier, which is important for analyses that rely on a precise measurement of luminosity in order to precisely determine branching fractions. The luminosity uncertainty can be avoided by measuring a branching ratio rather than a branching fraction, where the decay of interest is measured relative to a similar, previously well measured decay. This was the method used in the analysis presented in chapter 5.

The beam in the LHC is larger during injection and acceleration than during the stable beam period during which collisions can occur. As the beam is accelerated the positions of protons within the bunch oscillate around the centre of the beam, as the magnetic field is constantly increased. Collimators are used to remove stray ‘halo’ protons from the edges of each bunch, as these particles could potentially damage the surrounding



machine and the detectors. During acceleration, the proton beams have a transverse diameter of approximately 1 mm when passing through LHCb. When the beams are fully stabilised and ready for collisions, quadrupole magnets either side of LHCb focus the beam such that the transverse diameter at the collision point is 16  $\mu\text{m}$ . The design of the vertex locator was therefore driven by the requirements to be as close as possible to the collision point during collisions, whilst also not being damaged by stray protons during acceleration.

### 3.2.2 VELO

The Vertex Locator (VELO) is the first of the LHCb sub-detector systems. The main purpose of the VELO is to precisely reconstruct the origin vertex of each particle that traverses the detector; this allows for the identification of decay vertices that are displaced from the original interaction point, which is characteristic of the decay of particles containing  $b$  quarks. This requires a highly specialised detector that is capable of accurately measuring the track that each charged particle follows, in the short space close to the decay. The VELO consists of 42 modules split into two halves, one on each side of the beam pipe, covering approximately a metre in the  $z$  axis surrounding the interaction point, with modules both upstream and downstream of the collision region. Each module consists of two silicon strip detectors, glued back to back; one of these measures the  $r$  coordinate of a particle passing through the module and the other measures the  $\phi$  coordinate. Combining these two measurements, along with the  $z$  position of the module, gives a three dimensional position measurement of the particle. Tracking the particle through the various modules makes it possible to determine the origin vertex of that particle, by examining where the particle's path appears to coincide with the path from another particle (or particles for multi-body decays). The VELO will be discussed in detail in chapter 4.

### 3.2.3 Tracking system

To continue tracking the particles through the detector after the VELO, LHCb makes use of a number of tracking sub-detectors at various points through the detector. These stations consist of the Tracker Turicensis (TT), and the Tracking Stations, which are in turn divided into the Inner Tracker (IT) and the Outer Tracker (OT). The primary goal

of these tracking detectors is to measure the trajectory of a particle before and after the magnet, thereby obtaining a measurement of the particle's charge and momentum, by considering the magnitude of the deflection caused by the magnetic field.

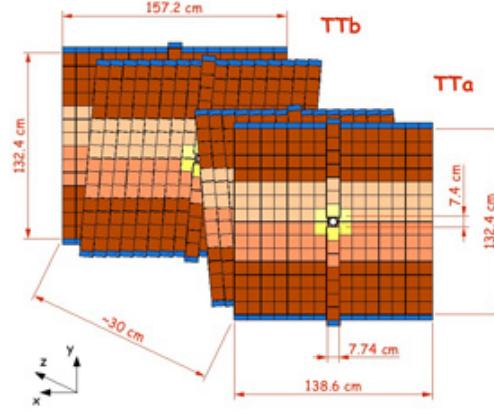


FIGURE 3.6: Illustration of the two pairs of silicon detectors used in the LHCb TT. Note the first and last of the detectors share the same alignment, whilst the second and third are rotated by  $+5^\circ$  and  $-5^\circ$  respectively [36].

The TT consists of two pairs of detectors,  $TT_a$  and  $TT_b$ , which are upstream of the magnet and are separated by approximately 27 cm in the  $z$  direction. The four stations are aligned in a  $x - u - v - x$  layout, where  $x$  denotes a vertical alignment, and  $u$  and  $v$  correspond to a rotation of  $x$  by  $+5^\circ$  and  $-5^\circ$  respectively; this layout is demonstrated in figure 3.6. The  $TT_a$  detectors consist of 15 vertically aligned modules above the beam pipe, and an identical set of modules below the beam pipe. Each module is made of 7 silicon microstrip sensors; these sensors are  $500\text{ }\mu\text{m}$  thick, single-sided  $p\text{-}in\text{-}n$  doped silicon, with 512 readout strips each and a strip pitch of  $183\text{ }\mu\text{m}$  [37]. Adjacent modules are offset by approximately 1 cm in the  $z$  axis, with a slight overlap in order to ensure full coverage. The two  $TT_b$  detectors are almost identical to the  $TT_a$  detectors but instead consist of 17 modules, which is necessary in order to achieve a similar angular coverage as the  $TT_b$  detectors are placed further downstream.

There are three tracking stations located downstream of the magnet,  $T_1$ ,  $T_2$  and  $T_3$ . Each station is split into two parts, where the IT forms the central portion of the stations surrounding the beam pipe, and the OT makes up the rest of the stations. This arrangement, along with the TT, is shown in figure 3.7.

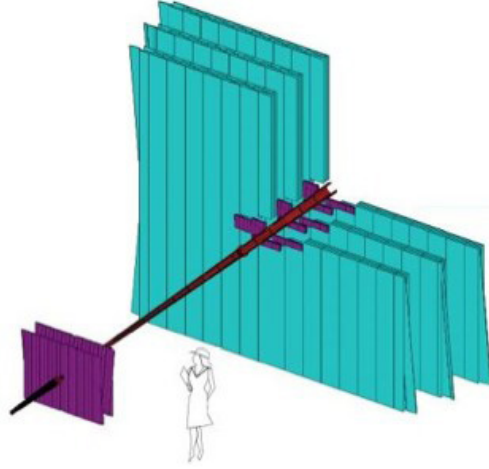


FIGURE 3.7: The layout of the tracking system in LHCb. This shows the two pairs of TT stations in purple on the left of the figure, and the larger tracking stations in light blue. The central portion of the tracking stations is the IT, in purple, while the rest is the OT. Note that the magnet is placed between the two sections in the detector.

Each IT station is made from four segments which are above, below, left and right of the beam pipe. Each module of the IT is made from 7 consecutive silicon microstrip sensors; each sensor is again single-sided *p-in-n* doped silicon, but with only 384 readout strips, and a slightly larger pitch than the TT, at  $198\text{ }\mu\text{m}$  [38]. The segments above and below the beam pipe consist of just one module, but the left and right segments are made from two modules each. Each of the IT stations is made from 4 layers, which follow the same  $x - u - v - x$  arrangement as in the TT. Therefore each station consists of 168 silicon sensors, meaning the IT utilises 504 sensors in total. Two layers of an IT station are illustrated in figure 3.8, where the layout and number of sensors is clearly seen.

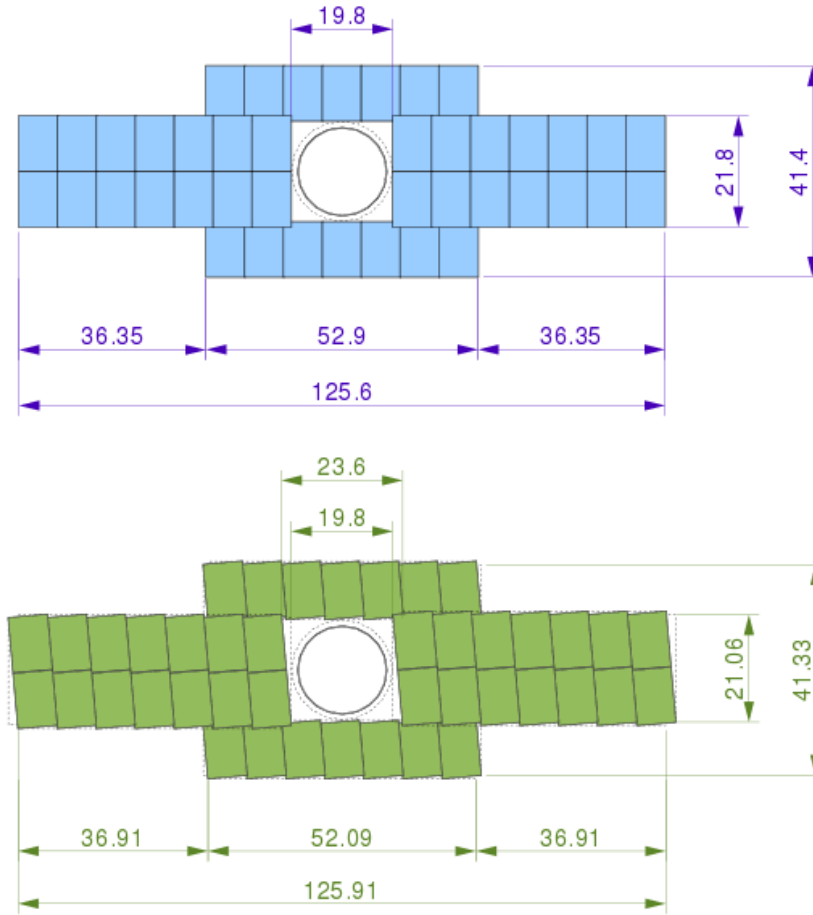


FIGURE 3.8: Illustration of two layers of the IT, an  $x$  layer (top) and a stereo rotated layer (bottom). Dimensions are given in cm, and refer to the sensitive silicon surface area covered [38].

The IT uses silicon sensors as it covers the area closest to the beam pipe, where the particle flux is higher. Using silicon in this area achieves a good hit resolution in spite of the higher flux. However, using silicon for the entire area of the tracking stations would prove far too expensive, as they have to be large in order to cover the entire LHCb acceptance so far downstream of the interaction point. For this reason, the OT surrounds the IT, from the edge of the IT stations out to the edge of the acceptance. The OT stations use drift time detector technology; this reduces costs without degrading the detector performance, as the resolution required in the OT is lower than the IT due to the lower flux experienced further away from the beam pipe.

The OT uses straw tubes with a 4.9 mm inner diameter, consisting of a gas mixture surrounding an anode wire, with the whole straw kept under a 1.5 kV bias voltage [39]. When a charged particle passes through the gas mixture, it ionises the gas. The liberated electrons drift towards the anode wire, creating an avalanche of electrons as they travel through the electric field. These electrons are then collected by the anode wire, which is read out as an electric current. The time taken to read out the straw is determined by the time taken for the electrons to drift towards the anode wire; thus the drift time can be converted to a position measurement of the traversing charged particle. The gas mixture of Argon (70%) and CO<sub>2</sub> (30%) is chosen to provide a drift time of  $< 50$  ns, which is short enough to allow rapid readout of the detector, but long enough that the position resolution is not adversely affected. Within a module, adjacent straws are staggered in order to prevent acceptance gaps. Each of the OT stations consists of 2 modules, with each module again consisting of 4 layers arranged in the  $x - u - v - x$  format. This means there are 8 detection layers per station, and therefore 24 layers in total across the entire OT.

### 3.2.4 RICH detectors

The ability to separate between kaons and pions over a wide momentum range is vitally important to LHCb, as these particles often appear in the final states of many of the  $B$  hadron decays that LHCb will analyse. For example, the analysis presented in chapter 5 studied the branching fraction of the decay  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  normalised to the decay  $B^0 \rightarrow D_s^\pm D^\mp$ . In the analysis,  $D_s^\pm$  decays are reconstructed through the  $D_s^\pm \rightarrow K^+ K^- \pi^\pm$  final state, and  $D^\pm$  decays are reconstructed through the  $D^\pm \rightarrow K^\mp \pi^\pm \pi^\pm$  final state. The ability to separate kaons and pions was clearly vitally important for the analysis, especially as the final states of the signal and normalisation channels differed only through the interchange of a  $K^\pm$  for a  $\pi^\pm$ . This analysis would have been much tougher in the GPDs, where it would be necessary to use more difficult ways of identifying hadron flavour, such as requiring the  $K^+$  and  $K^-$  to come from an intermediate  $\phi$  resonance.

In order to separate the particles, LHCb makes use of two Ring Imaging Cherenkov detectors, RICH1 and RICH2. These detectors utilise the principle of Cherenkov radiation in order to measure the velocity of the particles that traverse them. When a particle

travels through a medium at a speed greater than the speed of light in that medium it emits a cone of photons, where the opening angle of the cone is related to the particle's velocity by the equation given in 3.4, where  $\theta_c$  is the opening angle,  $c$  is the speed of light in a vacuum,  $n$  is the refractive index of the medium, and  $v$  is the particle's velocity:

$$\cos \theta_c = \frac{c}{nv}. \quad (3.4)$$

This velocity measurement can then be combined with the momentum measurements from the tracking systems to provide an estimate of the mass of the particle, which can be interpreted as an identification of the type of particle. LHCb is able to identify and separate kaons and pions over a momentum range from about 2 GeV/ $c$  to 100 GeV/ $c$ . The lower limit comes from the need to properly separate hadrons from multi-body B hadron decays, and also the fact that hadrons with momenta below 2 GeV/ $c$  will generally be bent out of the LHCb acceptance when traversing the magnetic field. The upper limit is governed by the momentum spectrum of two body B hadron decays, such as  $B^0 \rightarrow K\pi$  and  $B^0 \rightarrow \pi\pi$ , where there is also a great need to be able to properly identify the daughter hadrons [15]. In order to achieve this, the RICH detectors make use of 3 different media, known as radiators, to provide a range of refractive indices to give good  $K$ - $\pi$  separation over the whole of the required range. The use of Silica Aerogel,  $C_4F_{10}$ , and  $CF_4$  provides enough information to allow for the efficient separation over the whole momentum range; figure 3.9 shows the Cherenkov angle versus momentum plot produced with the  $C_4F_{10}$  radiator, clearly showing the separation of kaons, pions, and protons. Figure 3.10 is a simulation of the Cherenkov angle over the full momentum range in the three different materials, illustrating the benefit of using multiple media in order to increase the momentum range to which the detectors are sensitive. In figure 3.10,  $n=1.03$  is the refractive index for the Aerogel,  $n=1.0014$  is the refractive index for  $CF_4$ , and  $n=1.0005$  is the refractive index for  $C_4F_{10}$ . This is also why the scale in figure 3.9 is much smaller, as only the  $C_4F_{10}$  radiator is shown.

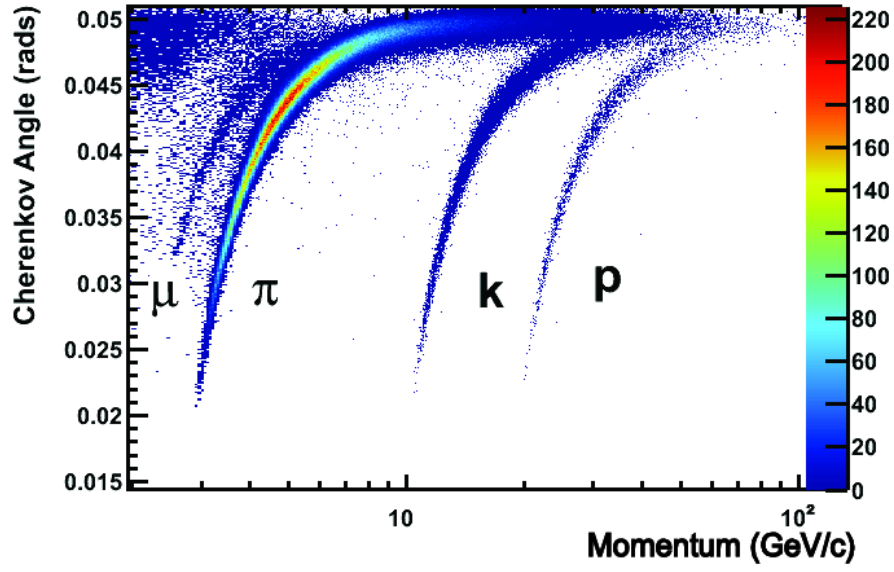


FIGURE 3.9: Cherenkov angle as a function of track momentum in the  $C_4F_{10}$  radiator in the RICH1 detector; the ability to separate kaons, pions, and protons (and to a lesser extent, muons) is clearly evident [40].

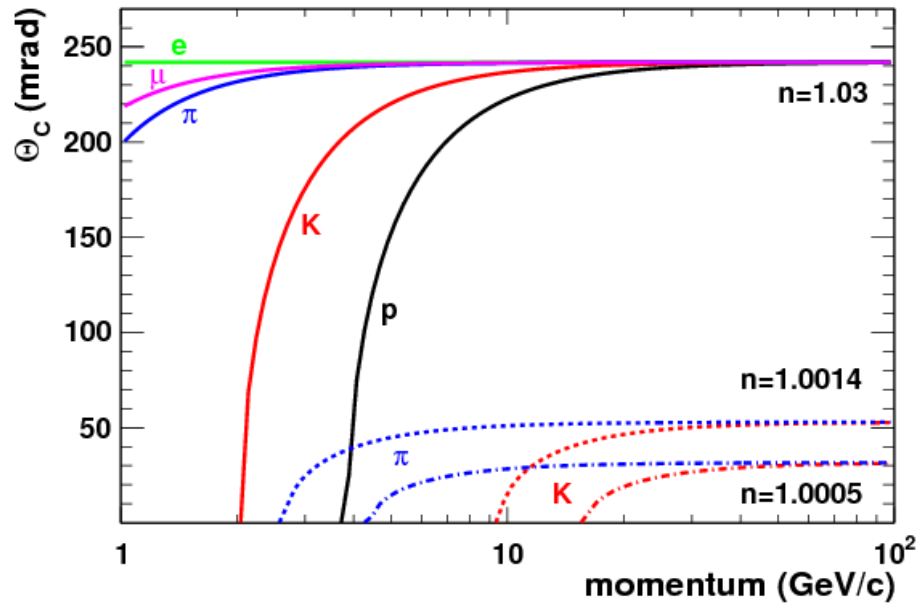


FIGURE 3.10: Simulated Cherenkov angle as a function of track momentum in the three different media used in the RICH1 and RICH2 detectors. Using three materials allows for separation across a wider momentum range [41]. The refractive index values,  $n$ , are: Aerogel = 1.03,  $CF_4$  = 1.0014, and  $C_4F_{10}$  = 1.0005.

The geometry of the RICH detectors is shown in figure 3.11. RICH1 is situated upstream of the magnet, and gives  $K/\pi$  separation up to momenta of approximately 40 GeV/ $c$ . RICH2 measures the high-momentum particles up to approximately 100 GeV/ $c$ , and is situated downstream of the magnet, as the higher momentum particles are affected less by the magnetic field, and are less likely to be bent out of acceptance. RICH1 covers the full angular acceptance of LHCb, whereas RICH2 covers a more central portion of the acceptance as, in the forward region, high momentum particles are much more likely to have low angles compared to lower momentum particles. In order to reduce the amount of detector material that particles must pass through, the RICH photodetectors used to measure the light emitted by hadrons are located outside of the LHCb acceptance in both RICH1 and RICH2. The Cherenkov light is focused and reflected out by spherical mirrors, onto the series of photodetectors. It is also necessary to use a series of flat mirrors to reflect the light further out onto the photodetectors in RICH2, in order to move the photodetectors out of the detector acceptance; this is also illustrated in figure 3.11.

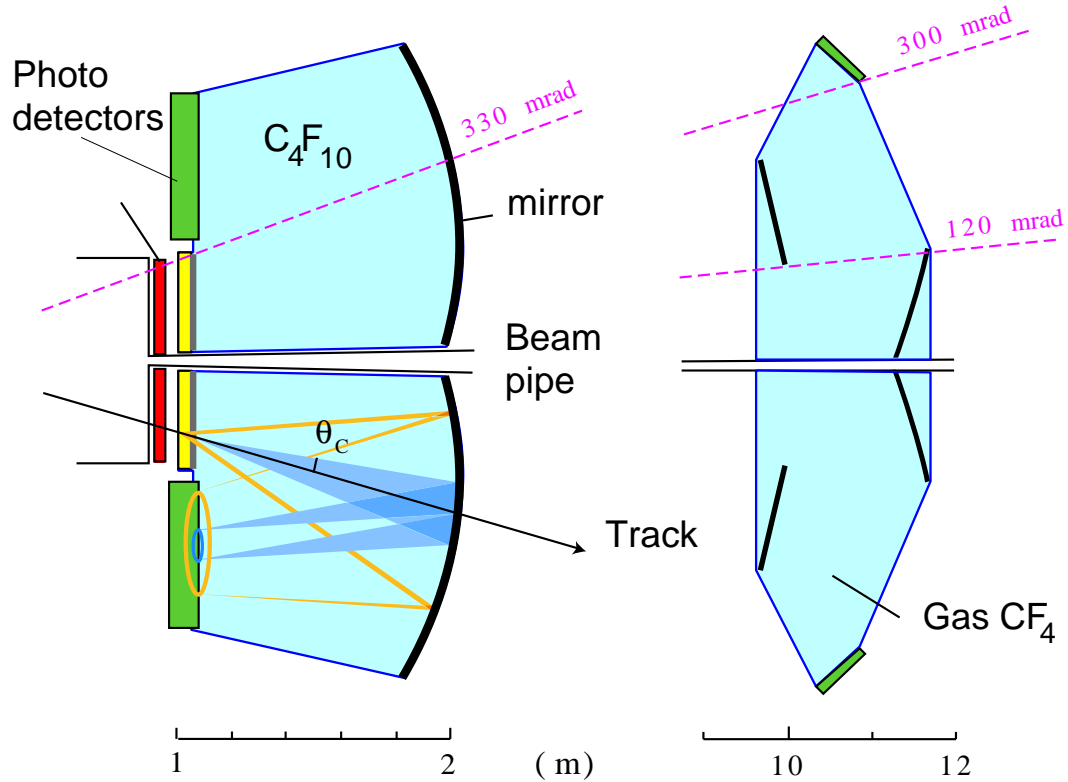


FIGURE 3.11: Schematic of RICH1 (left) and RICH2 (right), showing the mirror structure and the acceptance of the two detectors [42].



The  $K$ - $\pi$  separation that the RICH detectors provide is not perfect, and this can give rise to numerous backgrounds in LHCb analyses when a hadron is misidentified by the RICH as a different hadron. It is often important to properly account for these backgrounds in the analyses; for example, section 5.4 shows how the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  analysis dealt with backgrounds resulting from the misidentification of protons and pions as kaons.

### 3.2.5 Dipole magnet

LHCb makes use of a magnet in order to bend the path of charged particles. This allows for a measurement of a particle's charge and momentum by the tracking stations, by considering the trajectory of the particle both before and after traversing the magnetic field. The LHCb magnet is a warm dipole magnet which produces a integrated magnetic field of 4Tm in the  $y$  axis, which is large enough to have a measurable effect on high momentum particles. The polarity of the magnet can be reversed, and this is performed regularly in order to allow the assessment of charge asymmetric responses or slight alignment offsets of the detector [43].

### 3.2.6 Calorimeters

The LHCb calorimetry system consists of a number of detectors that are designed to make measurements of the energy of the particles as they pass through. There are four sub-detectors making up the calorimetry, each with its own distinct role: the Scintillating Pad Detector (SPD) and the Pre-Shower detector (PS) are mainly used for electron identification, the Electromagnetic Calorimeter (ECAL) measures the energy of particles that interact primarily via the electromagnetic force, and the Hadronic Calorimeter (HCAL) measures the energy of particles that interact primarily through the strong force. The calorimeters are also used in the event trigger of LHCb; this will be discussed further in section 3.2.8.

All of the calorimeter sub-detectors use the same basic principles to measure the energy of the particles passing through the detector [44]. When a high energy particle passes through a certain amount of material, it gives up a portion of its energy in the form of a particle shower. The energy of the particles produced in the shower can then be directly measured using scintillating tiles, and this can be interpreted as a measurement

of the energy of the original particle. For particles interacting electromagnetically, their energy loss is governed by the radiation length of the material,  $X_0$ . After traversing one radiation length, the particle will lose energy such that it only has  $1/e$  of its original energy. For particles interacting mainly through the strong force, the energy loss is instead related to the nuclear absorption length of the material,  $\lambda_{int}$ .

The SPD/PS system consists of two planes of scintillator pads, separated by a 15 mm block of lead (corresponding to  $2.5 X_0$ ). Both are segmented into inner, middle, and outer sections, each with progressively larger dimensions. The SPD detects charged particles before any showers are initiated, and can therefore be useful for differentiating between electrons and photons. The PS detector provides another measurement of showering effects before the particles reach the main calorimeters.

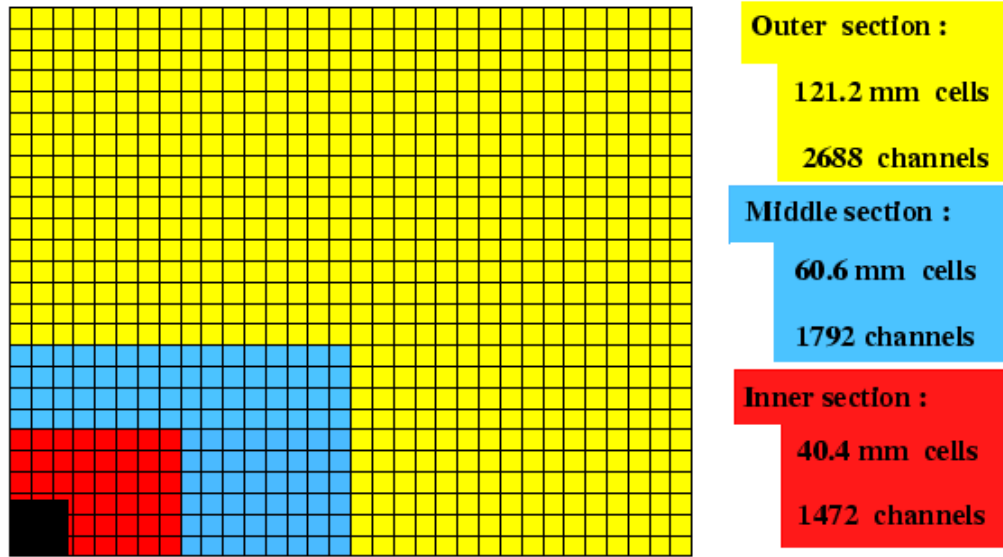


FIGURE 3.12: Illustration of the segmentation of the inner, middle, and outer sections of the LHCb ECAL, including the length of the square cells in each section [44].

The ECAL consists of 60 modules placed along the beam line, with each module made up of alternating layers of 2 mm thick lead plates and 4 mm thick scintillating tiles. This set up initiates electromagnetic showers from mainly photons and electrons, where the showers are caused by a combination of Bremsstrahlung and pair production. The entire length of the ECAL corresponds to a thickness of  $25X_0$ . As with the SPD/PS, the ECAL is split into inner, middle, and outer sections, which become gradually larger to account for the greater hit density closer to the beam pipe. These sections, and their

individual dimensions, are illustrated in figure 3.12. The energy resolution of the ECAL is  $\frac{8\%}{\sqrt{E}} \oplus 0.8\%$  [45].

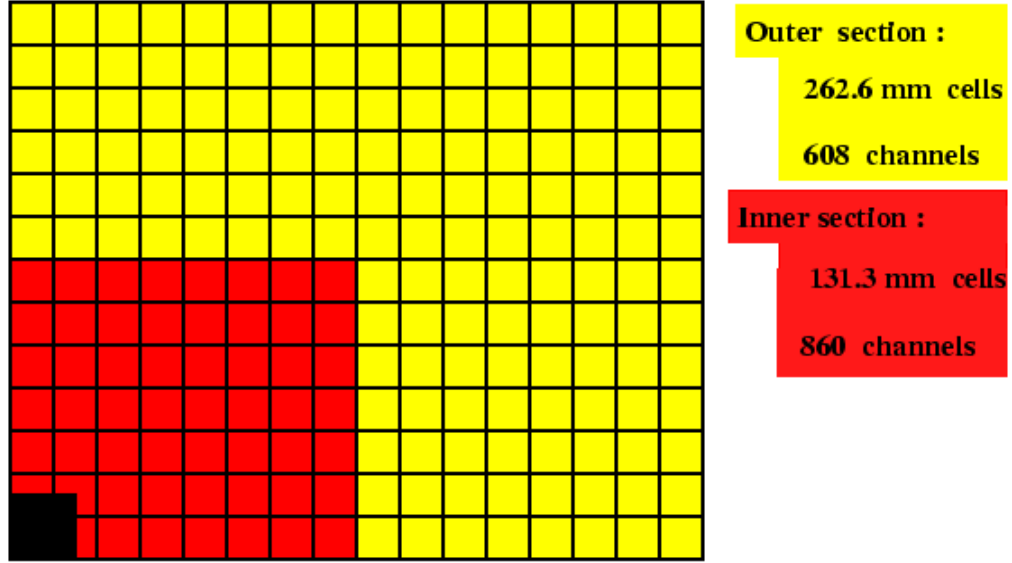


FIGURE 3.13: Illustration of the segmentation of the inner and outer sections of the LHCb HCAL, including the length of the square cells in both sections [44].

The last of the calorimeters is the HCAL. The HCAL uses iron as the absorber material, which causes hadrons to produce showers by interacting through the strong force. As  $\lambda_{int}$  is in general greater than  $X_0$ , the HCAL must be larger than the ECAL to measure the energy of the hadrons with a similar precision. The entire size of the HCAL corresponds to approximately  $5.6\lambda_{int}$ . Unlike the prior two calorimeters, the HCAL is only segmented into inner and outer sections, as illustrated in figure 3.13. The energy resolution of the ECAL is  $\frac{69\%}{\sqrt{E}} \oplus 9\%$  [45].

### 3.2.7 Muon chambers

The final sub-detectors of LHCb are the muon detectors. Many of the decays that LHCb will measure have final states that include muons, such as the  $B_s^0 \rightarrow \mu^+\mu^-$  measurement [46], so it is vitally important that the detector is able to identify and reconstruct muons. Muons are the most highly penetrating detectable particle, so the muon detectors are placed at the point of the detector furthest from the interaction point. There are five muon detectors in total, M1-M5, with M1 situated before the calorimeters, and

the other four coming after [47]. Thick iron absorbers are placed between the individual detectors in order to remove any hadronic background and ensure that the only particles detected at this stage are muons. Each of the muon detectors is split into four regions which surround the beam pipe with progressively larger radii. The regions are designed to have roughly equal occupancies, which means that the spatial resolution of the muon detectors decreases with distance from the beam line [48].

### 3.2.8 Trigger and Stripping

Due to cost considerations for data storage, and also the speed at which data can be written to disk, it is impossible for LHCb to record all of the collisions producing detectable particles, estimated to occur at a rate of 10 MHz. The LHCb data acquisition system was designed to read out data events at a rate of approximately 3 kHz, meaning only a small fraction of the collisions that occur in the detector are stored as data. This reduction is performed by the LHCb trigger system, which uses information from both hardware and software systems to narrow down the huge number of initial collisions and store only the events deemed to be most interesting for offline analysis by the various LHCb analysis groups. The trigger system consists of three separate stages: the Level-0 (L0) hardware trigger, and two High Level Trigger (HLT) software triggers. Only events that pass all three stages of the trigger are stored, as shown in figure 3.14.

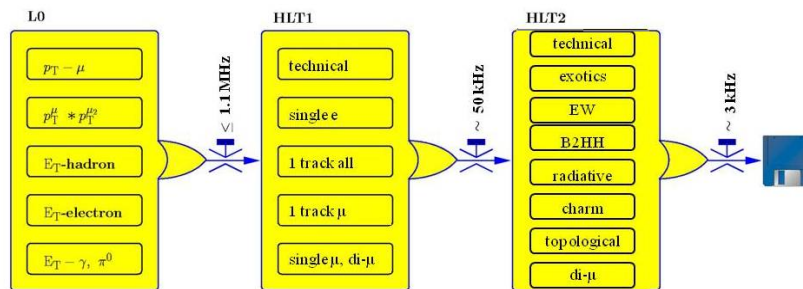


FIGURE 3.14: Flow chart showing how the LHCb trigger system works to reduce the initial event rate down to approximately 3 kHz in order to be saved for offline analysis [49].

## L0

The first stage of the trigger system is a hardware level system that makes use of the response from various parts of the detector in order to rapidly decrease the number of events to a rate of approximately 1 MHz. There are three separate triggers acting at the L0 stage, with the output of each used by the Decision Unit to decide whether to accept or reject an event.

The first of the L0 triggers is the Pile-Up Trigger. This trigger was designed to use two VELO R sensors, placed downstream of the VELO, to look for evidence of multiple primary interaction vertices (PVs) in an event. Rejecting these events at trigger stage is beneficial as the analysis of events with multiple PVs is much more difficult, meaning these events are of little use to the analysis groups.

The L0 also makes use of a Calorimeter Trigger, which uses the various calorimeter systems described in section 3.2.6 to look for evidence of high transverse energy ( $E_T$ ) hadrons, electrons, photons, and neutral pions. High  $E_T$  decay products can often indicate an initial  $B$  hadron decay, making these events potentially interesting for offline analysis. The Calorimeter Trigger also records the total HCAL  $E_T$  deposit and the SPD multiplicity, as these can help identify events with many final state particles that would command a long processing time.

The final part of the L0 trigger is the Muon Trigger. This trigger uses the information from the muon chambers to select events containing one or two high transverse momentum ( $p_T$ ) muons. This can be indicative of either the presence of a  $J/\psi$  particle at some point in the detector, which can be often produced in  $B$  hadron decays, or alternatively of various other  $B$  hadron decays of interest to LHCb analyses.

The L0 Decision Unit then combines the information from each of the three separate triggers, and compares the outputs to five sets of pre-defined criteria that select events that contain any of the following: one high  $p_T$  muon, two high  $p_T$  muons, a high  $E_T$  hadron, a high  $E_T$  electron, or a high  $E_T$  photon or  $\pi^0$ .

Events in LHCb that are selected by the trigger can either be Triggered On Signal (TOS), or Triggered Independent of Signal (TIS). If an event was recorded as TOS, then the event was selected by the trigger due to one or more of the particles involved in the signal event passing the required criteria for a particular trigger. If an event was

recorded as TIS, then the event was selected as a result of a particle that was not part of the signal decay meeting the criteria for the event to pass the trigger.

### **HLT1**

The HLT1 is a software trigger that is designed to further reduce the event rate to approximately 50 kHz. This trigger stage reconstructs tracks in the VELO, and then attempts to match these tracks to the particles selected by the L0 trigger. This allows HLT1 to confirm or reject the decision from the L0 Decision Unit.

### **HLT2**

HLT2 is the final stage of the LHCb trigger, and uses the rest of the detector information to fully reconstruct tracks in the events that have passed HLT1. HLT2 compares each event to a number of different selection criteria, where each selection is designed to identify events that appear to match a particular physical process or decay. The output rate of HLT2 is approximately 3 kHz, and all remaining events are saved for offline analysis. The HLT2 stage of the trigger works in much the same way as the actual final selections used in the analyses at LHCb, often making use of the same information.

### **Stripping**

The first stage of each analysis on LHCb is to pass the triggered data events through stripping. Each analysis makes use of a specific stripping line, which is a tighter selection designed to pick only events that may be of interest to that particular analysis. For example, the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  analysis presented in chapter 5 used a stripping line designed to look for events containing decays of  $B^0$  or  $B_s^0$  mesons into two oppositely charged  $D$  mesons. The selection criteria applied to select such events are discussed in section 5.3.

#### **3.2.9 Software chain**

LHCb makes use of a number of software packages for Monte Carlo production and real data analysis. These are described in brief below.

## **Gauss**

Monte Carlo (MC) production in LHCb is carried out using the GAUSS software package [50, 51]. The first stage of GAUSS simulates the proton collisions using PYTHIA [52, 53], and decays  $B$  hadrons using the EVTGEN package [54]. Final state radiation is simulated using PHOTOS [55]. All of the MC particles are then passed through a full simulation of the LHCb detector, which is produced using the GEANT4 framework [56, 57].

## **Boole**

The second stage of MC production is to digitise the output of the full simulation from GAUSS, and this is performed using the BOOLE [58] package. This stage makes the simulated events look like data, by giving the detector response that would be produced if the event had been an actual data event.

## **Moore**

The MOORE [59] package is used to run the trigger on digitised detector output.

## **Brunel**

BRUNEL [60] is responsible for fully reconstructing the events based on the digitised output, either of simulated events from BOOLE, or from the detector output from actual collisions. Following this stage, events are prepared for full analysis in .dst (data summary tape) files.

## **DaVinci**

Physics analyses are performed on the data and MC files using DAVINCI [61]. Event selections and specific decays can be programmed in the options in order to produce analysis specific outputs for further study. It is also possible to select the variables to include in the output that will be of interest in further analysis. DAVINCI outputs .root files that can be fully analysed using the ROOT analysis software [62].

## Chapter 4

# Radiation Damage in the Vertex Locator

The use of semiconductor technology has become a popular choice for modern particle detectors, as it offers better resolution and faster read out than has been previously achievable; the first part of this chapter examines how this technology is used to track charged particles. LHCb makes use of semiconducting silicon strip sensors in both the VELO and the tracking stations, as described in chapter 3, and section 4.3 details the set up and characteristics that allow the VELO to effectively track high energy particles close to the interaction point in LHCb. One drawback of using silicon sensors so close to the interaction region is that they are damaged by the high energy particles that traverse the sensor, and this can ultimately degrade the performance of the sensor. The final sections of this chapter examine the causes and effects of this radiation damage, focusing specifically on the increase in leakage current in the sensors, and shows how this damage is monitored and predicted at LHCb.

### 4.1 Semiconductors

A semiconductor material is characterised by the presence of an energy gap between the valence and conduction electron bands that is large enough that the material behaves as an insulator at very low temperatures, with no movement of electrons between bands, but which allows some transitions of electrons at room temperature. An illustration



of the difference between valence and conduction bands in conductors, insulators, and semiconductors is shown in figure 4.1.

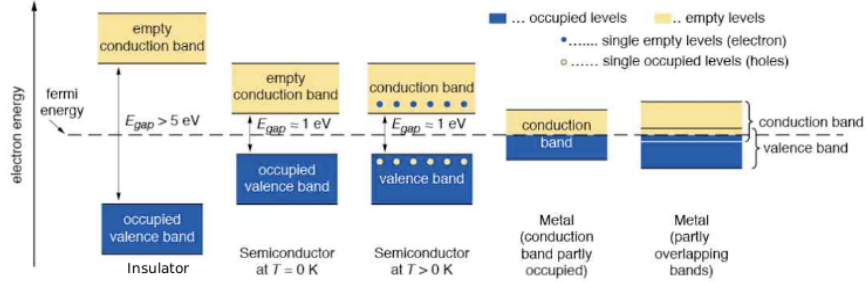


FIGURE 4.1: Illustration of the band gap structure in insulators, semiconductors, and conductors (metals).

When an electron moves from a state in the valence band into a state in the conduction band, it becomes a free negative charge carrier. The hole in the valence band left behind by the electron may also be thought of as a positive charge carrier which will also contribute to the conduction process [63]. In an intrinsic semiconductor, electrons and holes are formed in pairs, and as such the number of free electrons is always equal to the number of holes. These electron-hole pairs continuously form and then recombine.

It is possible to alter the characteristics of a semiconductor through a process known as doping. Doping involves introducing small levels of impurities to the semiconductor lattice in order to increase the number of free charge carriers, and can be performed in two ways. The first type of doping is positive, or *p*-type doping. In this case, a number of atoms of an element with one fewer outer shell electrons than the semiconductor material are implanted in the lattice. This results in the presence of extra holes in the valence band, and therefore a net increase of free charge carriers. The second type of doping is negative, or *n*-type doping. Here, the lattice is doped with an element with an extra outer shell electron, resulting in the presence of extra electrons in the material. In silicon, phosphorus is commonly used as the *n*-type donor element, whereas boron is used as the *p*-type material.

## 4.2 *p-n* junctions

A *p-n* junction is created when *n*-type and *p*-type doped silicon are joined in the same semiconductor crystal. Electrons from the *n*-type region diffuse into the *p*-type region, and similarly holes from the *p*-type region diffuse into the *n*-type region. This creates a third region between the *p*-type and *n*-type regions which is devoid of free charge carriers, and is therefore known as the depletion region. This also creates an electric field across the junction; the system reaches an equilibrium state where the net migration of electrons and holes is balanced by electrons and holes moving in the opposite way due to the electric field.

It is possible to increase the size of the depletion region by applying a reverse bias voltage across the silicon. Connecting the *n*-type region to the positive terminal causes more of the free electrons to drift away from the junction, with holes from the *p*-type region being simultaneously pulled towards the negative terminal; this results in the depletion region widening by an amount governed by equation 4.1 [64]:

$$V_D = \frac{q}{2\epsilon} |N_{\text{eff}}| d^2. \quad (4.1)$$

In this equation,  $V_D$  is the voltage required to deplete a region of thickness  $d$ , where  $q$  is the electron charge,  $\epsilon$  is the permittivity of the sensor material, and  $N_{\text{eff}}$  is the number of effective charge carriers present. When  $d$  is equal to the thickness of the sensor full depletion is said to have occurred, and increasing the applied voltage further will no longer increase the depletion region.

### 4.2.1 *p-n* junctions as particle detectors

When a charged particle passes through a material it gives up a portion of its energy in the form of ionising radiation. This charge deposition is governed by the Bethe-Bloch formula shown in equation 4.2:

$$-\frac{dE}{dx} = \frac{4\pi}{m_e c^2} \frac{nz^2}{\beta^2} \left( \frac{e^2}{4\pi\epsilon_0} \right)^2 \left[ \ln \left( \frac{2m_e c^2 \beta^2}{I(1 - \beta^2)} \right) - \beta^2 \right]. \quad (4.2)$$

In this equation,  $-\frac{dE}{dx}$  represents the energy loss of the traversing particle over the distance travelled through the material,  $z$  is the charge of the particle,  $\beta$  is the particle's velocity divided by  $c$ ,  $n$  is the number of charged nucleons in the material,  $I$  is the mean excitation potential of the electrons in the material,  $m_e$  is the electron mass, and  $\epsilon_0$  is the permittivity of free space. An ionising particle passing through silicon will deposit energy that liberates a number of electron-hole pairs, providing the particle has a high enough energy.

By applying a high reverse bias voltage across the silicon  $p$ - $n$  junction, these electron-hole pairs can be separated before they can recombine in the lattice. The electrons drift towards the  $n$ -type silicon, and the holes drift towards the  $p$ -type; this induces a charge that can be read out and interpreted by external software in order to reconstruct the hit position of the traversing particle.

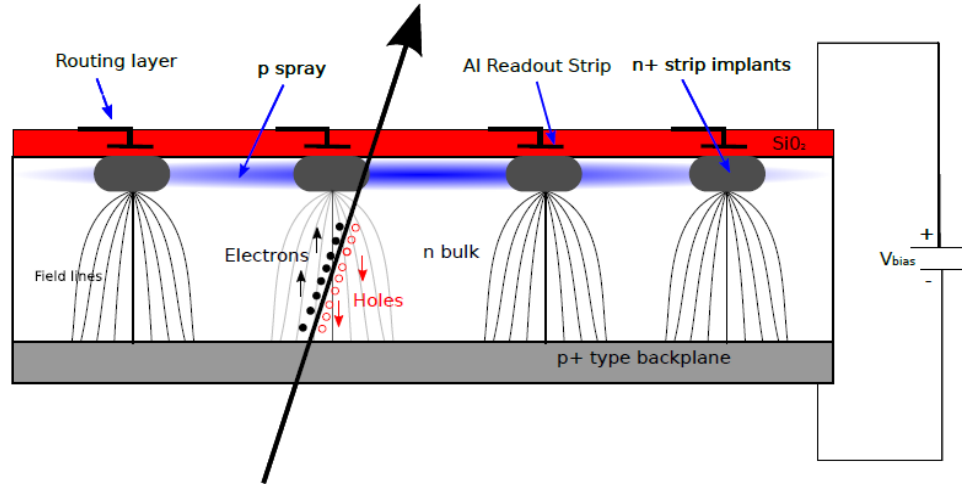


FIGURE 4.2: Diagram of an  $n^+$ -in- $n$  silicon detector, as used in the LHCb VELO. Electrons are carried by the electric field to the  $n^+$  strips, where they induce a mirror charge in the aluminium readout strips [65].

Figure 4.2 shows the sensor set up used in the LHCb vertex locator detector, known as an  $n^+$ -in- $n$  type detector. In this arrangement, the sensor consists of an  $n$ -type bulk, with a more heavily doped  $n$ -type ( $n^+$ -type) silicon as the strip implants, and a  $p^+$ -type backplane. The main sensor is separated from the readout strips by a layer of silicon oxide and an additional layer of  $p$ -type silicon, in order to prevent an excessive

build up of positive charge on the surface layer; the p-spray isolates the readout lines. As an ionising particle passes through the silicon it liberates a number of electron-hole pairs. The electrons then drift towards the  $n^+$  strip implants, where they induce a mirror charge in the aluminium readout strips. This is then readout and converted to Analogue-to-Digital-Converter (ADC) counts, which are analysed by the reconstruction software. A full description of the VELO follows in the next section.

### 4.3 The LHCb vertex locator

One of the main goals of the LHCb experiment is the precise measurement of the properties of hadrons containing  $b$  quarks.  $B$  hadrons are notable for their relatively long lifetime, which means that they travel a short distance after creation before decaying into lighter hadrons. This is also the case, albeit to a lesser extent, for  $D$  mesons; as the  $b \rightarrow c$  quark transition dominates over other  $b$  transitions, this is also of importance when measuring the decays of  $B$  hadrons. The ability to separate these secondary vertices is vital for precision measurements of  $B$  hadron decays as it allows the decays to be isolated from various background processes; in LHCb, this vertex identification is performed using the Vertex Locator (VELO).

The VELO is a silicon strip detector that surrounds the collision point within LHCb. It is split into two halves, with each half consisting of 21 modules spaced out along the  $z$  axis. The modules immediately around the interaction region are evenly spaced, whereas the modules further downstream are staggered at various intervals; the positioning of the modules in the VELO is illustrated in figure 4.3. Each module is made from two sensors, glued back to back, with one sensor measuring the  $r$  coordinate of a traversing particle, and the other measuring the  $\phi$  coordinate. Combining these measurements with the known  $z$  position of the module gives a 3-dimensional position measurement of a particle passing through the sensor.

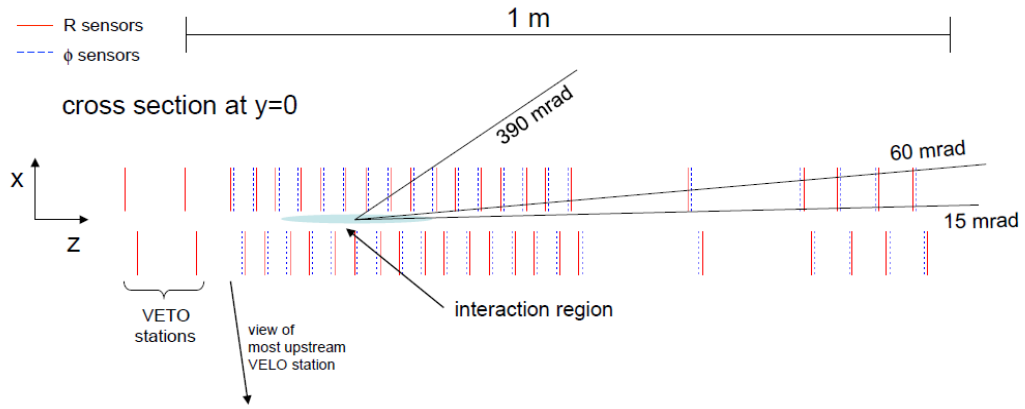


FIGURE 4.3: Illustration of the location of the modules within the VELO, along with the positioning of the interaction region where the protons are brought to collide within LHCb [66].

To increase the precision achievable with the VELO, it is essential to locate the sensors as close to the collision point as possible. For this reason, when collisions are occurring within LHCb, the sensitive area of each VELO sensor is just 8 mm from the proton beams. As discussed in section 3.2.1, the LHC requires a larger beam aperture when accelerating and stabilising the beams, and therefore the VELO sensors must retract to 30 mm from the beam in order to minimise the potential for damage from the beams. During collisions the sensors are then moved back into position surrounding the beam, with a slight overlap between the two halves in order to ensure complete coverage of the active silicon area [66].

The sensors are all made using 300  $\mu\text{m}$  thick  $n^+$ -*in*- $n$  silicon, except for two  $n$ -*in*- $p$  sensors (the replacement VELO sensors are all produced using  $n$ -*in*- $p$  technology). They are semicircular in shape, with an active region that covers from 8 mm to 42 mm in the transverse plane when in the closed position. The  $r$  measuring sensors are split into four  $45^\circ$  areas in order to reduce occupancy, with each area containing 512 circular silicon strips. The strip pitch varies from 38  $\mu\text{m}$  at the inner radius to 102  $\mu\text{m}$  at the outer radius, which accounts for the higher particle flux closer to the beam line. The  $\phi$  measuring sensors are split into two sections, inner and outer, to again account for relative occupancies. The strip pitch in both sections increases with distance from the beam line, from 38  $\mu\text{m}$  to 78  $\mu\text{m}$  in the inner section and 39  $\mu\text{m}$  to 97  $\mu\text{m}$  in the outer

section [67]. There is also a small skew between alternating  $\phi$  sensors in order to improve the tracking performance of the detector. Alternating the direction of the  $\phi$  sensors introduces a stereo angle between strips on adjacent sensors, which resolves the stereo ambiguities that result from multiple hits in the detector and reduces combinatoric background in the pattern recognition [68]. The sensor geometries are illustrated in figure 4.4.

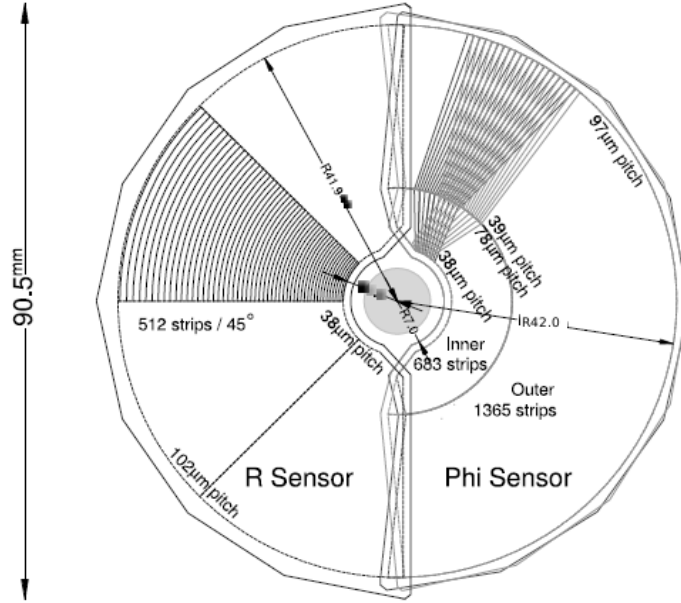


FIGURE 4.4: Illustration of the geometry of the two sensor types used in the VELO. The illustration of the  $\phi$  sensor shows two adjacent sensor geometries overlaid, in order to illustrate the alternating skew in these sensors [68].

The  $r$  and  $\phi$  sensors are glued to a printed circuit board structure known as a hybrid. The Beetle chips are also mounted on the hybrid, and are in turn connected to the sensors; the Beetle chip is the front end readout chip in LHCb with each sensor utilising one chip for every 128 silicon strips, totalling 16 chips for every sensor. The hybrid also contains the negative temperature coefficient thermistors (NTCs) that are used to monitor the temperature of the sensor; this is especially important when assessing annealing effects after radiation damage, as will be discussed further in section 4.6.2. The hybrid is connected to a rigid carbon fibre paddle, which contains the cooling mechanics; this setup constitutes a single module. Figure 4.5 shows a photo of a single

VELO module before installation, and figure 4.6 shows numerous modules fixed in place ahead of installation in the detector.

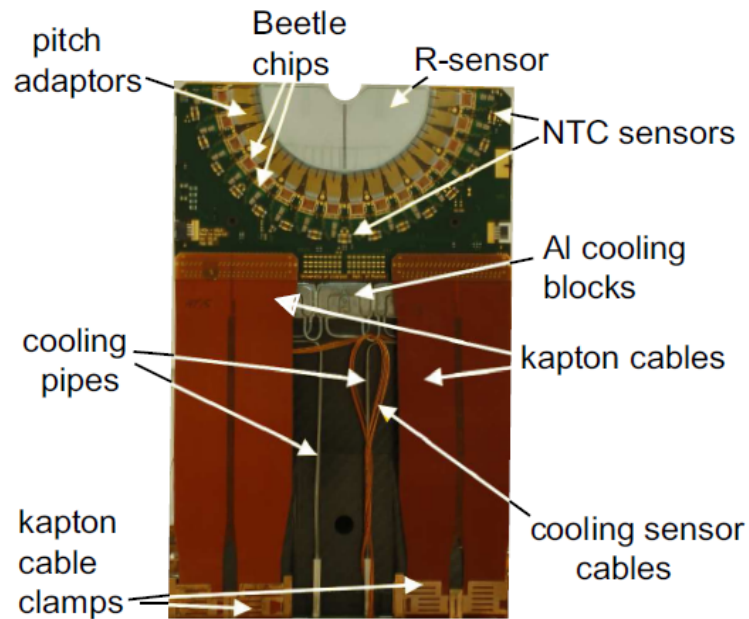


FIGURE 4.5: Photograph of a single VELO module before installation in the detector.

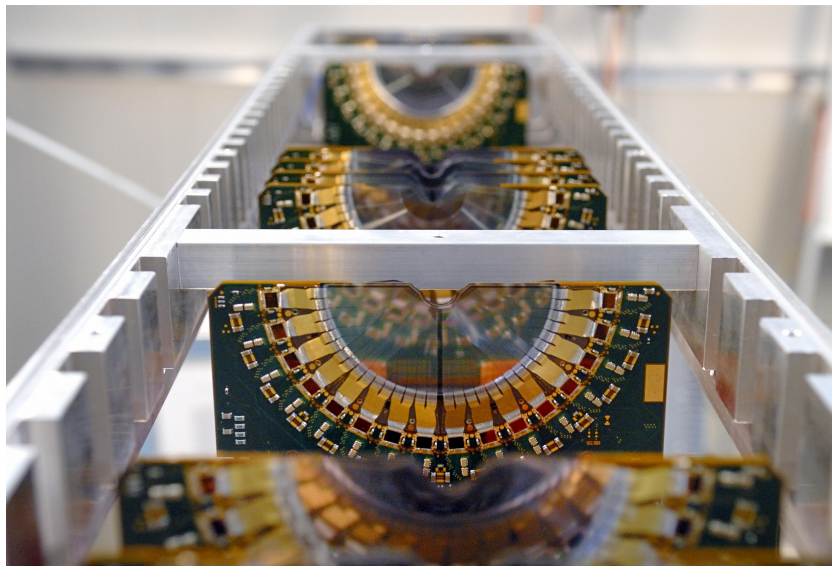


FIGURE 4.6: Photograph of multiple modules in the VELO, after being secured and awaiting installation into the full detector setup [69].

Cooling of the VELO modules is vitally important to the successful operation of the silicon sensors. Each of the modules requires 20W of power during normal operation, and the heat created by the chips would raise the temperature of the silicon to such an extent that the electronics would be damaged and the silicon would be destroyed by thermal runaway. Thermal runaway results from the exponential temperature dependence of the sensor leakage current. Raising the temperature of the sensor raises the leakage current, and the increase in current causes an increase in sensor leakage power which further heats the sensor. This feedback loop causes the temperature to run away, destroying the sensor [70]. The increase in temperature would also cause uncontrollable annealing after the silicon had been irradiated. To prevent this, each module is cooled by a bi-phase CO<sub>2</sub> system built into the carbon fibre support, as shown in figure 4.5. Heat is conducted away from the silicon by a 400  $\mu\text{m}$  thick Thermal Pyrolytic Graphite (TPG) core, which is in turn connected to aluminium cooling blocks known as cooling cookies. This aluminium is melted around the capillaries which carry the CO<sub>2</sub> coolant to and from the supply. The cooling system operates at a temperature of  $-30^\circ\text{C}$  and provides temperatures of approximately  $-7^\circ\text{C}$  to the active silicon area during operation, although the determination of the actual silicon temperature during operation is non-trivial, as will be further discussed in section 4.6.3.

The proximity of the VELO modules to the beam also introduces further mechanical issues which require the inclusion of further structure for the detector. The potential for damaging RF pickup from the proton beam passing between the VELO modules necessitates the use of a shield around each VELO half. The shield, known as the RF-foil, is made from 300  $\mu\text{m}$  thick aluminium, and is corrugated in shape in order to allow the two halves to close as near as possible to each other. A diagram of the RF-foil showing this corrugation is shown in figure 4.7. In order to prevent the RF-foil being deformed due to the very low pressure of the LHC beam vacuum, the VELO is kept in a secondary vacuum system. The majority of the full VELO detector, including the two sets of modules, the readout electronics, and the cooling system, is situated inside this secondary vacuum vessel; a pressure of  $10^{-6}$  mbar is maintained across the detector, as opposed to  $10^{-9}$  mbar for the main LHC vacuum. A diagram of the VELO vacuum vessel is shown in figure 4.8.



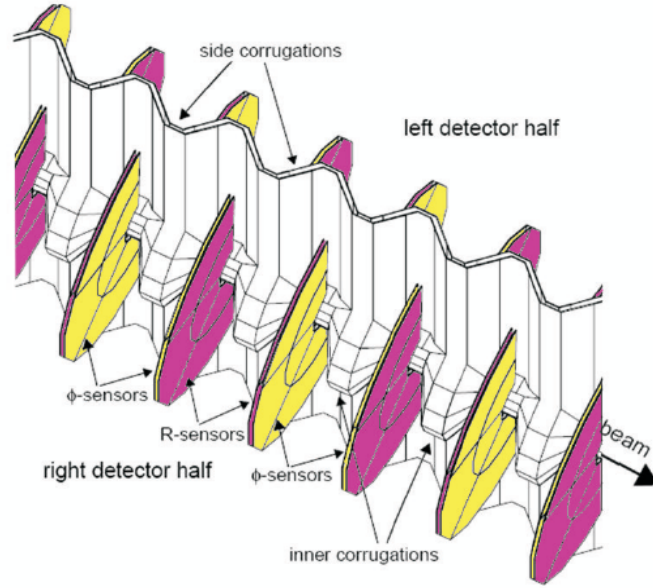


FIGURE 4.7: Diagram of the corrugated structure of the RF-Foil used in the VELO to prevent RF pickup from the proton beam passing between the modules [66].

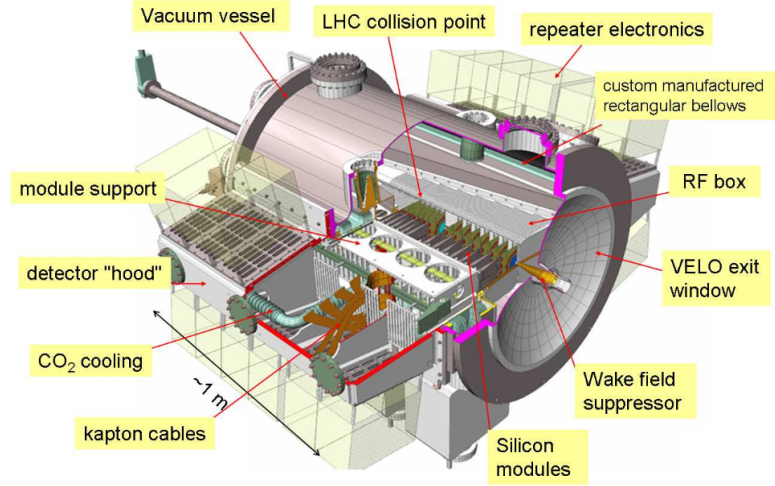


FIGURE 4.8: Illustration of the secondary vacuum vessel in which the majority of the VELO detector is situated within LHCb [71].

When a charged particle traverses one of the VELO modules, it deposits energy through ionisation which is then collected at the readout strips, as described in section 4.2.1. The charge on the strips is read out by the aforementioned Beetle chips in analogue form,

and then sent to a repeater board outside the VELO vacuum tank for amplification. From here, the signal is sent to the TELL1 boards, which are common read out boards designed for LHCb. These boards make use of four 10-bit Analogue to Digital Converters (ADCs) in order to convert the original analogue signal from the Beetle chips to ADC counts.

At this point, the collected charge is ready to be assessed as either signal or noise. As a traversing particle can deposit charge on more than one strip in a sensor, there are two separate threshold criteria that, if fulfilled, can define a signal. The first of these is the seed threshold; for a signal to be detected, a strip must have a charge that is six times the mean noise for that strip, which is typically equal to between 12 and 18 ADC counts. Once a seed strip is found, the adjacent strips are then searched to see if either of them pass the inclusion threshold criteria, which is set at 40% of the seed threshold. If the strip does indeed pass the inclusion criteria, it is said to form a cluster with the original seed strip. A cluster can contain a maximum of four strips, and the charge on each strip in the cluster is then used in an algorithm that assesses exactly where the incident particle traversed the detector. Figure 4.9 shows the distribution of ADC counts in a two strip cluster in a VELO  $r$ -sensor; this represents the charge deposition across a series of events. The distribution is modelled by a Landau function, which is well known to describe the energy lost by a particle traversing material, convoluted with a Gaussian function, which models the noise in the sensor. It should be noted that the noise contribution is small, and as such the Gaussian portion of the fit does not greatly transform the overall distribution from the underlying Landau distribution.

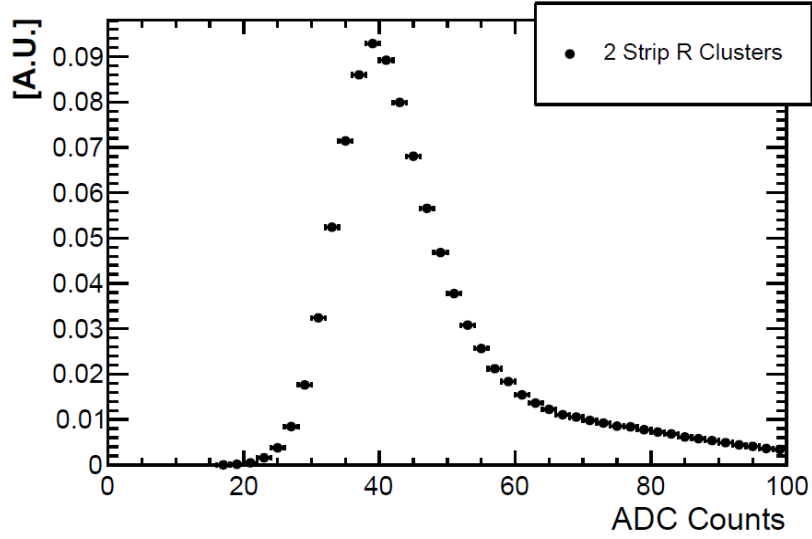


FIGURE 4.9: Example of the distribution of ADC counts in a two strip cluster in a VELO  $r$ -sensor [65].

The best possible single hit resolution of a VELO sensor is approximately  $4\text{ }\mu\text{m}$ , although this is dependent on a number of factors. By assessing the signal presence in all of the modules in the VELO, charged particles can be tracked through the detector, allowing the location of the various decay vertices present in an event in LHCb to be determined. Whilst the precision with which a vertex can be measured depends strongly on the number of tracks emanating from that vertex, the VELO can determine the position of a vertex with typically approximately  $20\text{ }\mu\text{m}$  resolution in the transverse plane, and  $100\text{ }\mu\text{m}$  resolution in the  $z$ -direction. Given that a  $b$ -hadron travels a distance of the order of  $1\text{ cm}$  before decaying, the VELO is easily able to separate  $b$ -hadron decay vertices from the original interaction region.

## 4.4 Radiation damage in silicon detectors

Silicon sensors can be damaged by the high energy particles which they are measuring, which can result in a degradation in their performance over time. This damage is classed as either ionising energy loss (IEL) or non-ionising energy loss (NIEL), depending on the mechanism causing the damage. The bulk damage in silicon detectors is primarily due to the displacement of a silicon atom, known as a primary knock on atom (PKA),

from its lattice site [72]. The damage resulting from a PKA depends on the amount of recoil energy transferred from the traversing particle; at low recoil energies damage is restricted to fixed point defects, but at higher recoil energies the PKA has enough energy to create a further defect cluster at the end of its track [73].

These defects alter the performance of the detector as they can ‘trap’ charge. For example, a cluster of holes in the silicon lattice can trap electrons produced by an incident ionising particle. This will delay the electrons reaching the implant strips, and therefore prevent the signal being read out by the electronics; this has the effect of decreasing the signal-to-noise ratio, and hence the tracking performance, of the silicon detector. Further problems result from ionisation damage; this occurs when the trapped electrons cause a concentration of charge to build up in an area of the silicon lattice which alters the electric field at that point. Local variations in the electric fields within the sensor cause changes in the mobility of electrons and holes in the lattice, which also affect the charge collection efficiency of the detector.

As the impact of this radiation damage on the performance of the silicon detector can be ultimately quite severe, it is important to understand the propensity for damage depending on the environment in which the detector will be used. The energy transferred to the silicon lattice by an incident particle varies greatly depending on both the type of particle and its initial energy, and for this reason it is necessary to scale all damage predictions and calculations to a reference point. The NIEL scaling hypothesis corrects all particles and energies to the equivalent number of 1 MeV neutrons, and has been shown to be a good approximation over a range of energies [74]. Incident particle fluences are therefore described in terms of 1 MeV neutron-equivalent particles, as opposed to the actual composition of incident particles; figure 4.10 shows the normalised damage factors of a range of particles and energies.

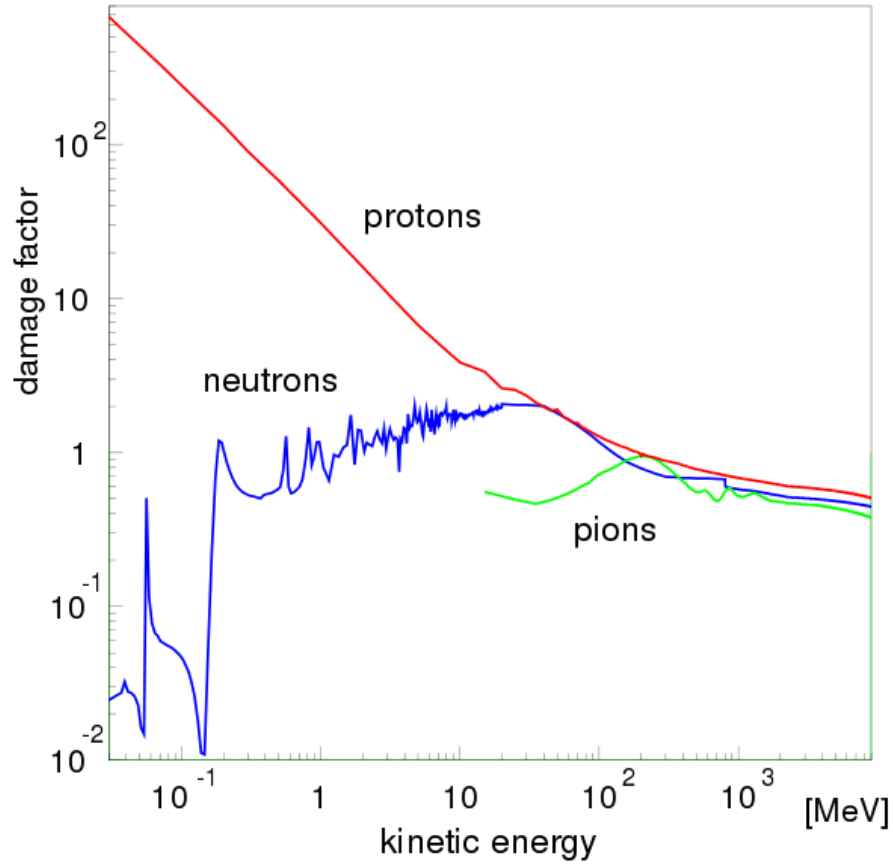


FIGURE 4.10: A plot showing the damage factors of various particles at different energies. The damage factor is normalised such that a 1 MeV neutron has a damage factor of 1. Plot produced using tables from [75].

This normalisation makes the prediction of total fluence received by the VELO silicon sensors possible, as shown in section 4.6.1, which then allows for a prediction of the likely damage that the silicon will sustain during use.

## 4.5 Monitoring leakage currents in the VELO

One of the main issues affecting the performance of a silicon detector after irradiation is an increase in the leakage current present in the detector. When a new silicon detector is operated under a reverse bias voltage, diffusion across the depletion region is heavily suppressed. However, there is still a very small current that flows across the detector,

which is known as the leakage current. This current increases with the reverse bias voltage applied to the detector, and saturates after the detector is fully depleted.

In a detector that has not been subjected to radiation this leakage current is due mainly to the thermal generation of electron-hole pairs in the depletion zone, with the applied voltage preventing recombination, and is thus small. As the silicon lattice becomes damaged by radiation, more defects are created in the depletion zone, meaning the leakage current increases. An increased leakage current is detrimental to the performance of the detector, as the noise present within the detector also increases. This alters the signal-to-noise ratio achievable by the detector, and as such can affect the ability to properly measure hit clusters. As the number of free charge carriers in the silicon is altered by radiation damage, it can also become necessary to operate the detector at a higher voltage to ensure full depletion. In order to maintain an efficient performance from the detector, it is vital to monitor the leakage currents in the silicon at regular points during its lifetime.

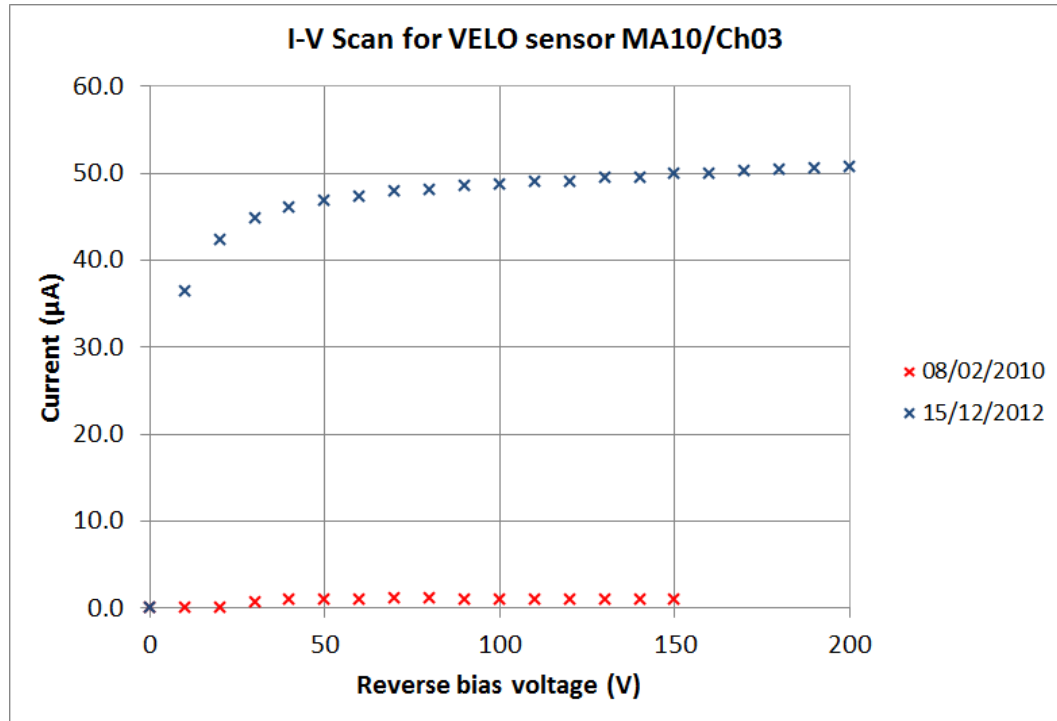


FIGURE 4.11: Plot of the I-V characteristic curve for one of the VELO sensors, as recorded at the start and the end of the first full data taking period at LHCb.

In LHCb, this monitoring is made possible by performing regular I-V (current-voltage) scans on the silicon sensors. Whilst the detector is not in use, the leakage current in the sensors is assessed by ramping the voltage from 0V to the operational voltage in 10V steps, and recording the current at each point. Figure 4.11 shows the results for one sensor from I-V scans performed at the start of 2010, and also at the end of 2012 after the detector had been in operation for almost 3 years, with approximately  $3 \text{ fb}^{-1}$  of delivered luminosity. The large increase in leakage current in the sensor is clearly seen.

Figure 4.12 shows the variation in leakage current for every sensor at full depletion over the period of operation prior to the first long shutdown of the LHC. The trend of increasing leakage current with delivered luminosity is clearly visible in all of the VELO sensors. Also apparent from this figure is the high leakage current present in some of the sensors at the start of operation; these increased currents result from surface defects present in the sensor at installation. The leakage current due to surface defects decreases with radiation damage, which causes the leakage current in these high production current sensors to initially decrease with luminosity, before eventually reverting to the standard behaviour displayed in the other sensors.

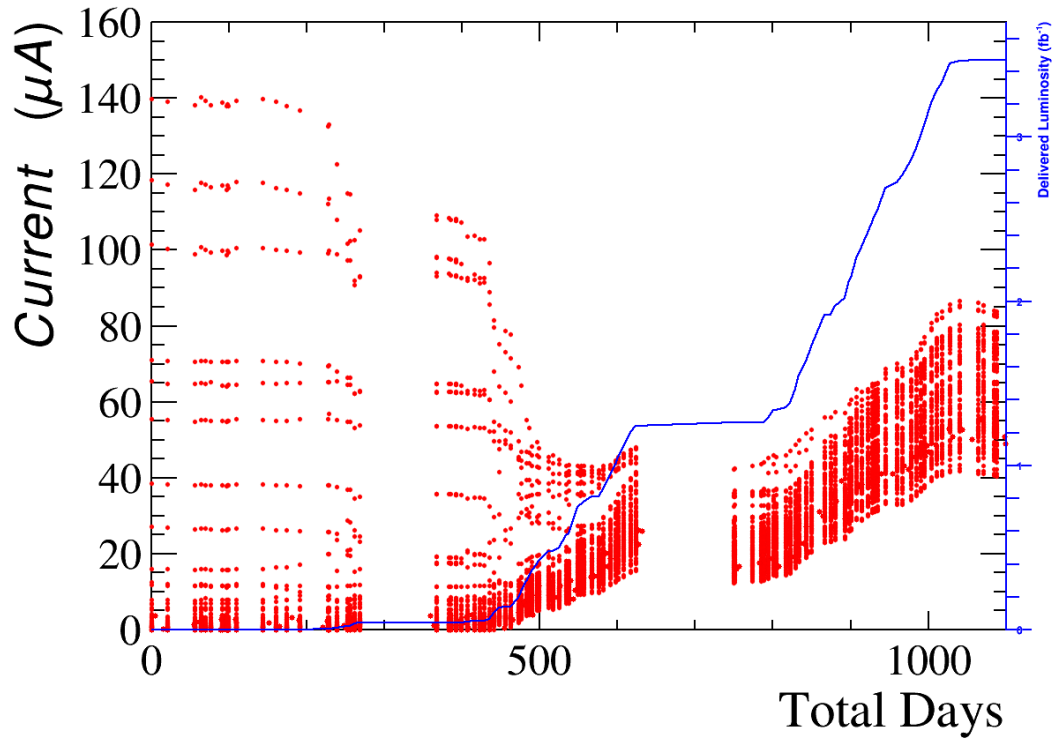


FIGURE 4.12: Showing the increase in leakage current in the various silicon sensors over time, as more luminosity is delivered to LHCb. The red dots represent the measurement for each individual silicon sensor at each IV scan at full depletion, and the blue line shows the total delivered luminosity in LHCb at that point.

The bulk damage to the silicon lattice resulting from irradiation is the main cause of increased leakage current [76], and as such it is important to understand this effect in order to maximise the long term performance of the VELO.

## 4.6 Predicting leakage currents

The leakage current in the VELO has two separate components: bulk leakage current, and surface leakage current. Bulk leakage currents, as the name suggests, arise from the current inside the detector volume. These currents have a known temperature dependence, and can therefore be predicted by considering the characteristics of the detector. Surface leakage currents result from imperfections or irregularities on the sensor surface, often from the manufacturing process or scratches suffered during testing and installation. In contrast to the bulk currents, surface currents do not display the same



temperature dependence and are much more difficult to model. For this reason, the predictions detailed in the remainder of this chapter are valid for describing only the bulk leakage currents in the VELO sensors.

It is possible to calculate the bulk current in a sensor as a function of the fluence received by the sensor using equation 4.3 [77]:

$$I = \alpha V \phi. \quad (4.3)$$

$I$  is the bulk current when the sensor is fully depleted (A),  $V$  is the active volume of the sensor ( $\text{cm}^3$ ),  $\alpha$  is a constant that relates the increase in the current to the fluence received by a sensor (A/cm) and will be discussed further in section 4.6.2, and  $\phi$  is the fluence received by the sensor, in units of 1 MeV neutron-equivalent per  $\text{cm}^2$ . The calculation of  $\phi$  is discussed further in the following section.

#### 4.6.1 Sensor fluence profiles

The sensors in the VELO are subjected to a very high flux of highly energetic particles. The flux received by the sensors is simulated in order to allow the estimation of the fluence profile across each sensor, for use in the bulk current predictions. The fluence profile of each sensor has a high radial dependence, as the parts of the sensor closest to the beam line receive a much greater flux than the outermost parts of the sensor. The fluence profile across a sensor is described by the equation in 4.4:

$$\phi = A r^{-\beta}. \quad (4.4)$$

In this equation,  $A$  is the 1 MeV neutron-equivalent fluence per  $\text{fb}^{-1}$  at a radius of  $r = 1 \text{ cm}$ , and  $\beta$  describes the exponential radial dependence of the fluence across each sensor. The fluence received depends not only on the region of the sensor, but also on the  $z$  position of the sensor relative to the interaction point. The sensors closer to the interaction point experience a higher flux of particles, meaning the  $A$  and  $\beta$  parameters must be calculated for each individual sensor. Figure 4.13 shows how the two parameters vary with the  $z$  position of the sensor, as well as two examples of the radial distribution of the fluence across an individual sensor.

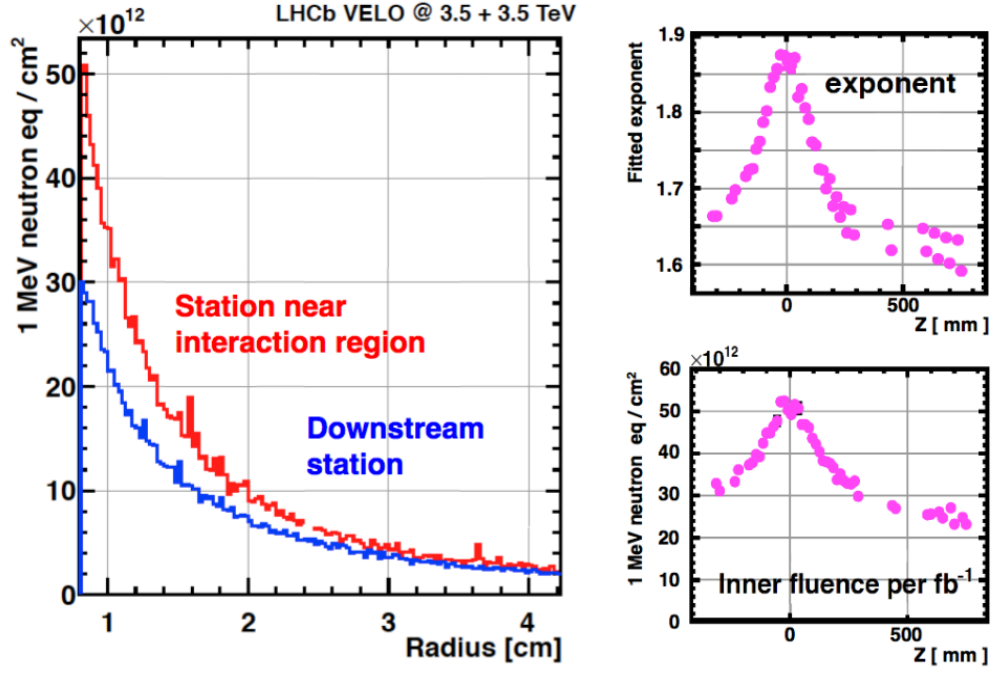


FIGURE 4.13: (Left) The simulated fluence profile across VELO sensors both near to (red) and downstream of (blue) the interaction region, and (right) the variation of the  $\beta$  (top) and  $A$  (bottom) parameters as a function of  $z$  position of the sensor [78].

In order to assess the increase in current resulting from the fluence received by the sensor, the sensor can first be modelled as a series of infinitesimal strips of radius  $\delta r$  at radius  $r$ . Substituting the volume of one of the strips into equation 4.3, and combining with equation 4.4, gives the current increase in the strip using the relationship shown in equation 4.5:

$$\delta I = \left( \alpha t A r^{-\beta} \right) \pi r \delta r. \quad (4.5)$$

Where  $t$  is the thickness of the silicon. Integrating this relationship, with respect to radius, across the active area of the silicon then gives the increase in current across the whole sensor, as shown in equation 4.6:

$$\Delta I = \frac{A \alpha \pi t}{2 - \beta} \left[ r^{2-\beta} \right]_{r_i=0.8 \text{ cm}}^{r_o=4.2 \text{ cm}}. \quad (4.6)$$

The final ingredient needed to allow for a prediction of the bulk current after a certain luminosity is a calculation of the factor  $\alpha$ ; this is discussed in the following section.

#### 4.6.2 Annealing

In order to obtain an accurate value for the  $\alpha$  parameter in equation 4.6, it is necessary to account for annealing in the silicon after radiation damage. Annealing is a process whereby thermal energy modifies the damage in the silicon lattice, further altering the performance of the sensor.

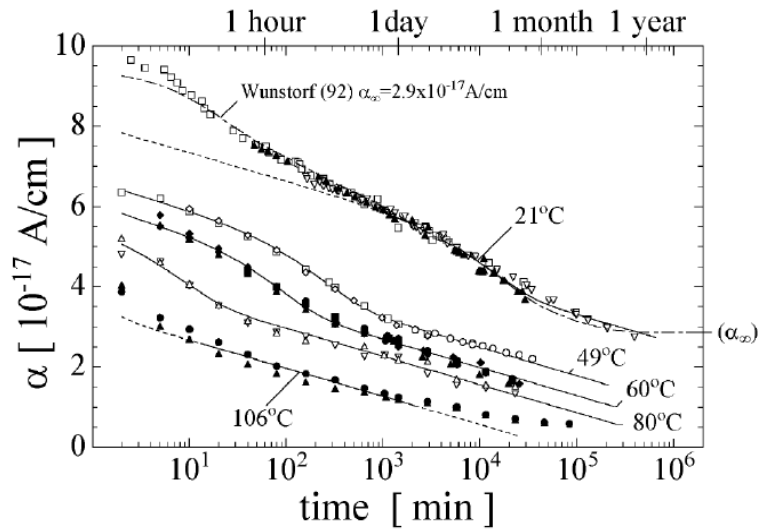


FIGURE 4.14: Results of annealing studies from [79]. The parameters in equation 4.7 are determined empirically by fitting to the 21°C data in this figure.

Annealing must be calculated empirically, and is modelled using equation 4.7, which is derived from fitting to data as shown in the plot in figure 4.14:

$$\alpha(t) = \alpha_I \times e^{-\frac{t}{T_I}} + \alpha_0 - \beta \ln \left( \frac{t}{t_0} \right). \quad (4.7)$$

The parameters in equation 4.7, as determined using the 21°C data, are shown in table 4.1.

Parameter	Value at 21°C
$\alpha_I$	$1.23 \times 10^{-17} \text{ A/cm}$
$T_I$	$1.4 \times 10^4 \text{ minutes}$
$\alpha_0$	$7.07 \times 10^{-17} \text{ A/cm}$
$\beta$	$3.29 \times 10^{-18} \text{ A/cm}$
$t_0$	1 minute

TABLE 4.1: The parameters for equation 4.7, determined empirically at 21°C by fitting to the data in figure 4.14, as taken from [79].

When the LHCb detector is not in operation, either between LHC fills or during shut-down periods, the VELO is kept at approximately  $-30^\circ\text{C}$  by the cooling mechanics described in section 4.3. At this temperature, the effect of annealing is negligible. The effects of annealing are more pronounced during data taking, when the VELO modules are much closer to  $0^\circ\text{C}$ . The most problematic periods occur when the cooling for the VELO is switched off, either temporarily due to cooling failures or intervention, or over a longer period such as the Christmas period of 2010/2011. All of these various periods are accounted for in the full prediction calculations.

### 4.6.3 Temperature corrections

As the annealing relationship used in equation 4.7 is only valid at 21°C, it is necessary to correct all annealing periods to this standardised temperature in order to accurately predict the effect on the leakage current present in the sensor. The temperature of the VELO is constantly monitored by an online Supervisory Control And Data Acquisition (SCADA) system called PVSS, with temperature information from the NTCs read out and stored at 10 second intervals<sup>1</sup>. This information is then analysed offline; each time the temperature varies by more than  $1^\circ\text{C}$ , the date and time is stored. This makes it possible to assess what temperature the sensor has been at, and for how long. By correcting this to an equivalent time spent at 21°C using equation 4.8, and then adding up the total time for the period between each IV scan, a prediction of the amount of annealing for that period can be made:

$$t_{T_2} = \left( e^{-\frac{E_a}{kT_1}} / e^{-\frac{E_a}{kT_2}} \right) \times t_{T_1}. \quad (4.8)$$

<sup>1</sup>The author wishes to thank Karol Hennessy for downloading and making available the PVSS data for this study.

In this equation,  $E_a$  is the silicon activation energy (1.31 eV),  $k$  is Boltzmann's constant,  $T_1$  is the initial temperature,  $T_2$  the target temperature, and  $t_{T_1}$  and  $t_{T_2}$  are the times spent at these temperatures.  $t_{T_2}$  is then used in equation 4.7 to obtain the  $\alpha$  factor for the corresponding period, which is then used to predict the total current increase.

As the predictions are made at the same point in time as the IV scans are taken, it is important to accurately account for the luminosity delivered in the time between scans. Due to the luminosity levelling in LHCb, described in section 3.2.1, the luminosity delivered can be accurately approximated as one single injection, equal to the total delivered luminosity in the period between the two scans, midway between the scans. This method makes the final predictions much simpler and quicker to calculate. The uncertainty on the prediction resulting from this luminosity approximation was examined in [80], and was found to be negligible.

As this prediction is made at  $21^\circ\text{C}$ , and the bulk current drawn in silicon is heavily dependent on the temperature of the silicon, the final step to obtain a prediction comparable to the measurements shown in section 4.5 is to correct the current increase back down to the operational temperature at which the scans are performed. This current correction is performed using equation 4.9:

$$\frac{I(T_1)}{I(T_2)} = \left(\frac{T_1}{T_2}\right)^2 \times e^{-\frac{E_{eff}}{2k} \times \left(\frac{1}{T_1} - \frac{1}{T_2}\right)}. \quad (4.9)$$

Here,  $E_{eff}$  is the effective silicon band gap (1.21 eV [81]),  $k$  is Boltzmann's constant,  $T_1$  and  $T_2$  are the initial and final temperatures, and  $I(T_1)$  and  $I(T_2)$  are the currents drawn at these temperatures.

The operational temperature of the silicon is not easy to determine. The silicon temperature relative to the temperature of the surrounding components was originally assessed using thermal photography in a test environment prior to the installation of the modules, as shown in figure 4.15, and was thought to be approximately  $-8.3^\circ\text{C}$ . However, due to differences between the test environment set up and the conditions in the installed VELO, this is now thought to give a temperature estimate that is a few degrees too low. The extraction of precise silicon temperatures will require lengthy work that is beyond the scope of the analysis presented here, but will be included in future iterations. The NTC chip calibrations are known to be incorrect, with a temperature dependent

adjustment needed in order to properly correct the temperature information provided by the chips. This will require a comparison of the NTC temperatures with the temperatures recorded at the cooling cookies when the power is switched off, and therefore the temperature gradient across the VELO modules is minimal. It will then be possible to properly analyse the calibration error of the NTC chips, and therefore increase the accuracy of the temperature information used in the analysis.

After discussion with the VELO group, it was decided to set the final correction temperature to  $-5^{\circ}\text{C}$ , with a large uncertainty of  $\pm 2^{\circ}\text{C}$ . This uncertainty accounts for the imperfect knowledge of the NTC calibration error, and also the temperature dependence of the error. A better knowledge of the calibration error will allow a more precise estimate of the operational silicon temperature in future, and therefore decrease the size of the uncertainty due to the final temperature correction. This uncertainty is by far the dominant source of systematic uncertainty in the analysis, and is thus the only source considered in the final plots in the following section. When the NTC calibrations are better understood and the current temperature correction uncertainty of 40% is decreased, it will be necessary to consider systematic uncertainties resulting from various other sources, including: the LHCb absolute delivered luminosity determination ( $\approx 3.5\%$ ), the uncertainty resulting from the simulated fluence profiles ( $\approx 8\%$ ), the separation of the bulk and surface portions of the leakage current (estimated to be  $\approx 2\%$ ), and as a result of the  $1^{\circ}\text{C}$  step size used when analysing the temperature changes in the silicon (estimated to be less than 1%).

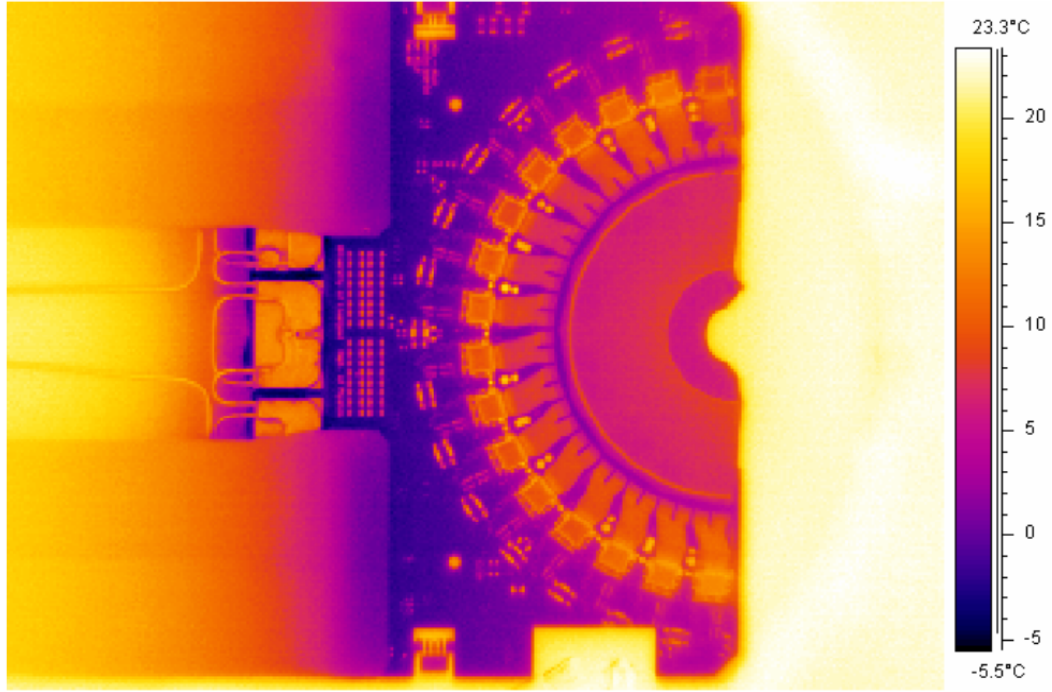


FIGURE 4.15: Example of a thermal photograph of one of the VELO modules from the vacuum tank test environment before installation [82].

#### 4.6.4 Results

Figure 4.16 shows the initial plot from figure 4.12 with the mean sensor leakage current prediction overlaid. The solid black line represents the mean prediction calculated from all sensors at a final temperature of  $-5^{\circ}\text{C}$ , and the green dashed lines show the mean value calculated at  $\pm 2^{\circ}\text{C}$  of the central value. The predicted trend clearly matches the observed trend well, except for the high production current sensors, where the high initial currents are due to surface leakage currents that are not modelled in the predictions.

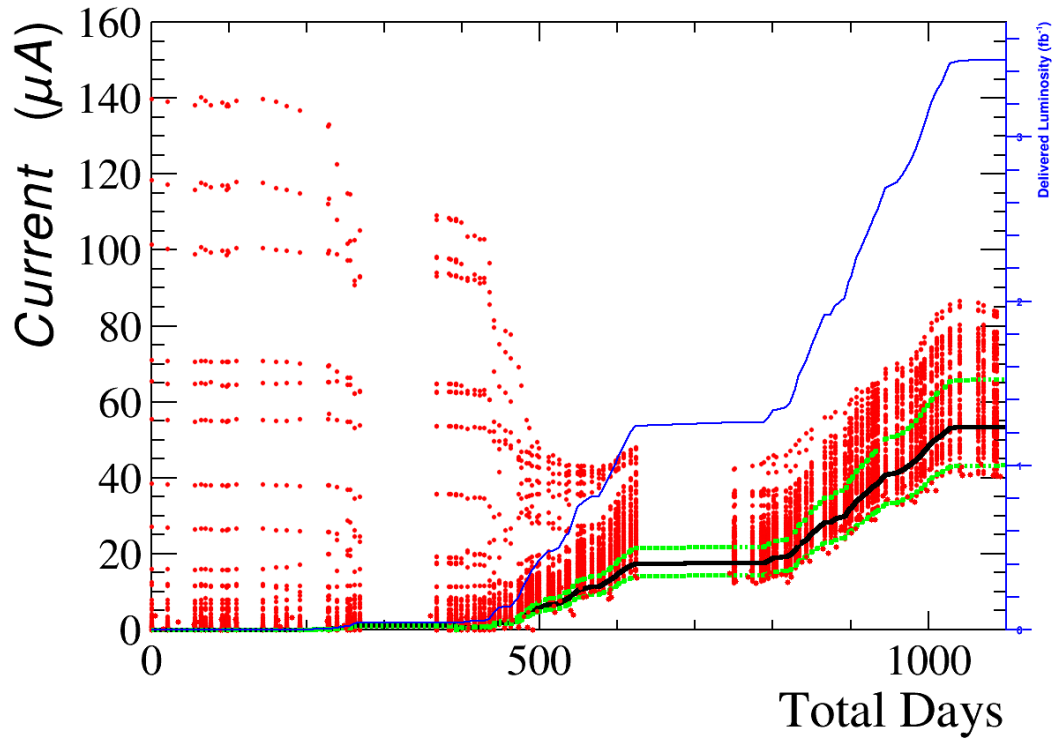


FIGURE 4.16: Showing the calculated mean leakage current prediction overlaid on the sensor by sensor leakage currents as measured in the I-V scans. The prediction mean is represented by the black line, corresponding to a temperature of  $-5^{\circ}\text{C}$ , while the green dotted lines represent a  $\pm 2^{\circ}\text{C}$  area around the mean.

Figures 4.17 and 4.18 show the leakage current increase in two individual VELO sensors. The sensor in figure 4.18 is closer to the interaction point than the sensor in figure 4.17, and as such has a higher leakage current at the end of the analysis period after the same luminosity. This behaviour is also modelled well by the prediction. In both cases, the evolution of the leakage current with delivered luminosity is seen to be well modelled by the prediction formalism.



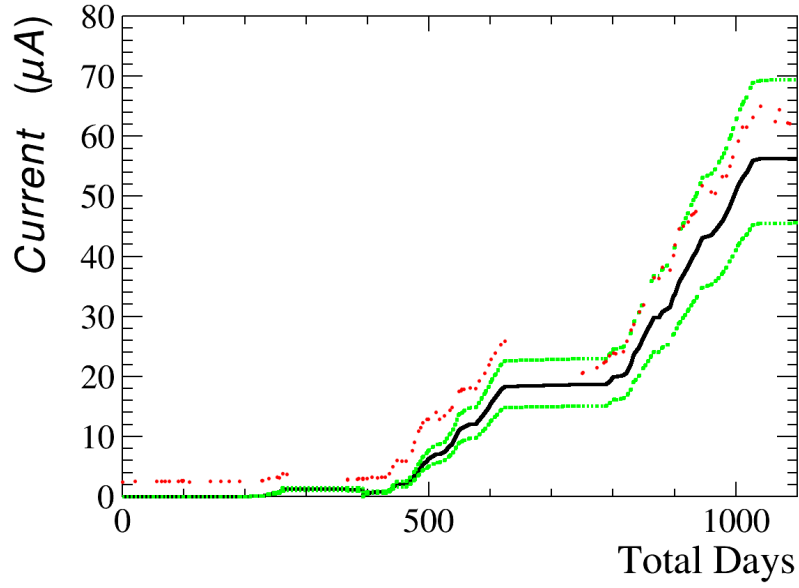


FIGURE 4.17: Example behaviour of the leakage current over time of one of the VELO sensors, both as measured in regular I-V scans (red dots), and as predicted (black line, with dashed green lines representing upper and lower uncertainty bounds).

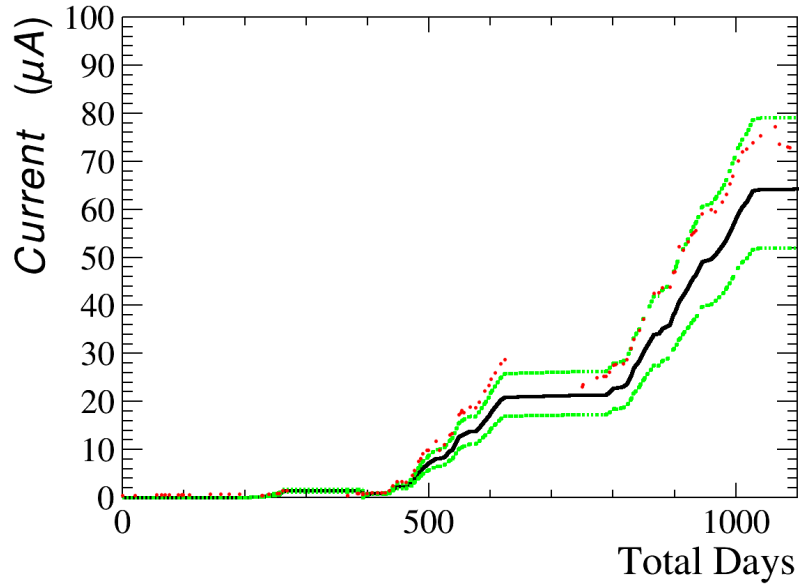


FIGURE 4.18: Example behaviour of the leakage current over time of another of the VELO sensors, both as measured in regular I-V scans (red dots), and as predicted (black line, with dashed green lines representing upper and lower uncertainty bounds).

#### 4.6.5 Summary and outlook

The effect of ionising radiation on the silicon detectors used in the LHCb VELO detector must be well controlled and understood in order to maximise the performance of the detector. Of particular importance is the monitoring of the leakage currents within the VELO sensors; regular I-V scans allow for the study of the evolution of these currents over time and with increasing levels of delivered luminosity.

The increase in leakage current due to the bulk damage inflicted on the silicon is well understood and modelled using the method detailed throughout this chapter. Understanding this current increase allows for a confident assessment of the capabilities of the VELO detector and its life expectancy with regards to delivered luminosity in LHCb. Providing there are no significant annealing periods, and excluding the likelihood of unforeseen circumstances destroying the detector, the VELO will be capable of efficient operation until being replaced by an upgraded VELO during the second long shutdown period of the LHC during 2018 [83].

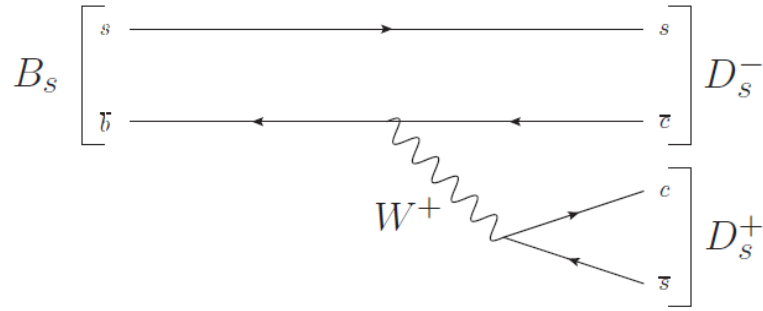
There are numerous improvements that will be made to the predictions during the second data taking period in LHCb that will allow more precise assessments of the expected leakage currents within the VELO sensors, and therefore make it easier to ensure that the current evolution is well understood. A detailed analysis of the NTC chips on a sensor by sensor basis should allow for a better estimation of the silicon temperature during running, which will decrease the uncertainty on the final predicted currents. It may also be possible to separate the bulk and surface leakage currents with the use of current-temperature (I-T) scans in the VELO [84]. As the predictions documented throughout this chapter are valid for the bulk leakage currents only, this could then allow for a more accurate comparison of the prediction with the observed current change in the VELO sensors.

## Chapter 5

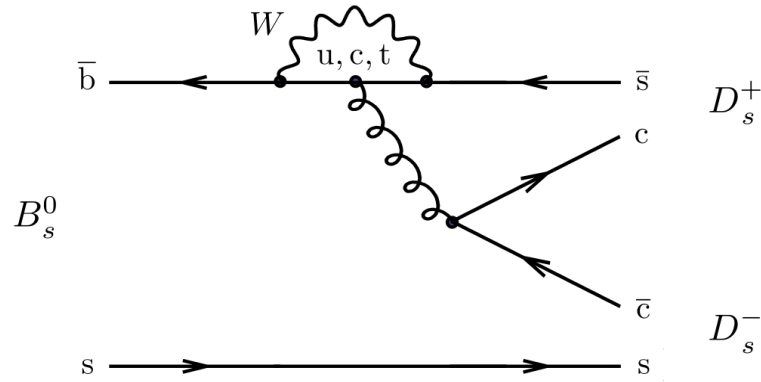
# $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ branching fraction

The decay of a  $B_s^0$  meson into a  $D_s^+$  and a  $D_s^-$  is of interest for a number of different analyses. The final state is  $CP$ -Even, and measuring decays to  $CP$  eigenstates provides a convenient way to measure numerous parameters arising from  $B_s^0$  mixing, as discussed in detail in section 2.2.1. Measuring decays to  $CP$  eigenstates avoids the necessity for a full angular analysis, which is needed to untangle  $CP$ -Even and  $CP$ -Odd states in mixed final states. A time dependent analysis of the  $B_s^0 \rightarrow D_s^+ D_s^-$  decay can determine the CKM angle  $\beta_s$ , which in turn determines the  $CP$ -violating phase  $\phi_s$ . For more information regarding the CKM angles and  $\phi_s$ , please see section 2.3.2.

At quark level, the decay is a  $b \rightarrow c\bar{c}s$  transition, and proceeds at tree level via the Feynman diagram illustrated in figure 5.1.

FIGURE 5.1: Leading order Feynman diagram for the decay  $B_s^0 \rightarrow D_s^+ D_s^-$ .

It is also possible for the decay to occur via a penguin topology, as shown in the Feynman diagram in figure 5.2.

FIGURE 5.2: Feynman diagram showing the decay  $B_s^0 \rightarrow D_s^+ D_s^-$  proceeding through a penguin diagram.

As  $b \rightarrow c$  is the Cabibbo favoured quark transition, the majority of  $B_s^0 \rightarrow D_s^+ D_s^-$  decays occur through the tree-level process, and penguin pollution is small. It is of interest, if possible, to calculate observables such as  $\Gamma_s$ ,  $\Delta\Gamma_s$ , and  $\phi_s$  both using tree-level dominated processes and also in decays that have a greater contribution from penguin diagrams in the final decay amplitude. Differences between the two measurements could be an indication of New Physics entering the penguin loops.

## 5.1 Previous measurements from $B_s^0 \rightarrow D_s^+ D_s^-$

The following section briefly describes two previous measurements made at LHCb using the  $B_s^0 \rightarrow D_s^+ D_s^-$  decay. For full details of each analysis, please refer to the relevant paper indicated in each subsection.

### 5.1.1 $B_s^0 \rightarrow D_s^+ D_s^-$ effective lifetime

The following is a brief summary of the LHCb measurement of the effective lifetime of the  $B_s^0$  when decaying to  $D_s^+ D_s^-$ ; a full description of the analysis can be found in [85].

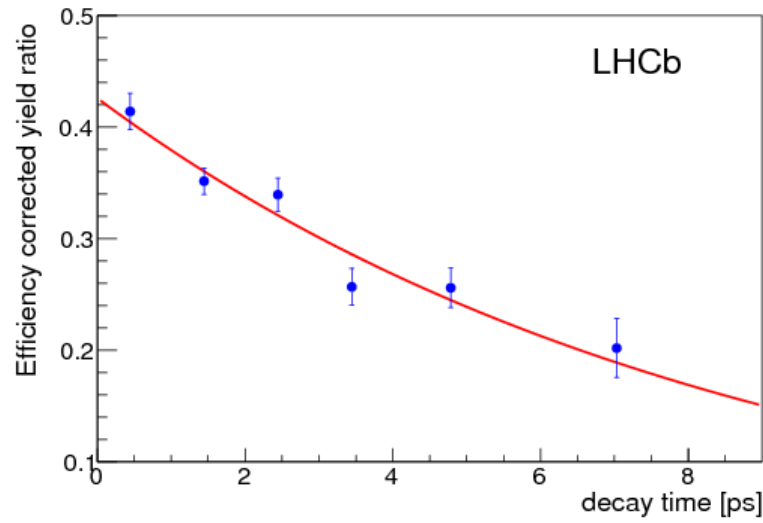


FIGURE 5.3: Efficiency corrected yield ratio for  $B_s^0 \rightarrow D_s^+ D_s^-$  relative to  $B^- \rightarrow D^0 D_s^-$  as a function of decay time. Also shown is the exponential fit used to extract the effective lifetime of the  $B_s^0$  in the  $D_s^+ D_s^-$  decay channel.

As the  $B_s^0 \rightarrow D_s^+ D_s^-$  has a final state that is almost entirely  $CP$ -Even, measuring the lifetime of this decay translates as a measurement of  $\Gamma_L$ , the decay width of the light  $B_s^0$  mass eigenstate, providing  $\phi_s$  is approximately 0. The fitted value of  $(-0.015 \pm 0.035)$  rad for  $\phi_s$ , as described in section 2.2.1, justifies this assumption. LHCb measured the yield of  $B_s^0 \rightarrow D_s^+ D_s^-$  relative to  $B^- \rightarrow D^0 D_s^-$  in bins of decay time. Fitting an exponential to the resulting plot, as shown in figure 5.3, made it possible to extract the effective lifetime using the already measured lifetime of the  $B^- \rightarrow D^0 D_s^-$  decay. The results for

the effective lifetime,  $\tau_{B_s^0 \rightarrow D_s^+ D_s^-}^{eff}$ , and  $\Gamma_L$  are shown below, where, in all cases, the first uncertainty is statistical and the second is systematic:

$$\tau_{B_s^0 \rightarrow D_s^+ D_s^-}^{eff} = (1.379 \pm 0.026 \pm 0.017) \text{ ps}$$

$$\Gamma_L = (0.725 \pm 0.014 \pm 0.009) \text{ ps}^{-1}.$$

These results are interesting to compare with measurements of the same observables from the channel  $B_s^0 \rightarrow K^+ K^-$ . The decay to  $K^+ K^-$  is also a  $CP$ -Even eigenstate, but the decay amplitude is expected to have larger contributions from loop processes [86]. The contribution of additional decay amplitudes from non-Standard Model particles is expected to be larger in  $B_s^0 \rightarrow K^+ K^-$ , if they exist [87]. The most recent measurement of the effective lifetime in the  $B_s^0 \rightarrow K^+ K^-$  decay is shown below [15]:

$$\tau_{B_s^0 \rightarrow K^+ K^-}^{eff} = (1.407 \pm 0.016 \pm 0.007) \text{ ps}.$$

No significant differences between the two effective lifetimes are observed.

### 5.1.2 $\phi_s$ from $B_s^0 \rightarrow D_s^+ D_s^-$

LHCb has also used the  $B_s^0 \rightarrow D_s^+ D_s^-$  decay channel to measure the  $CP$ -violating phase,  $\phi_s$ . For full details about the analysis, please refer to [88].

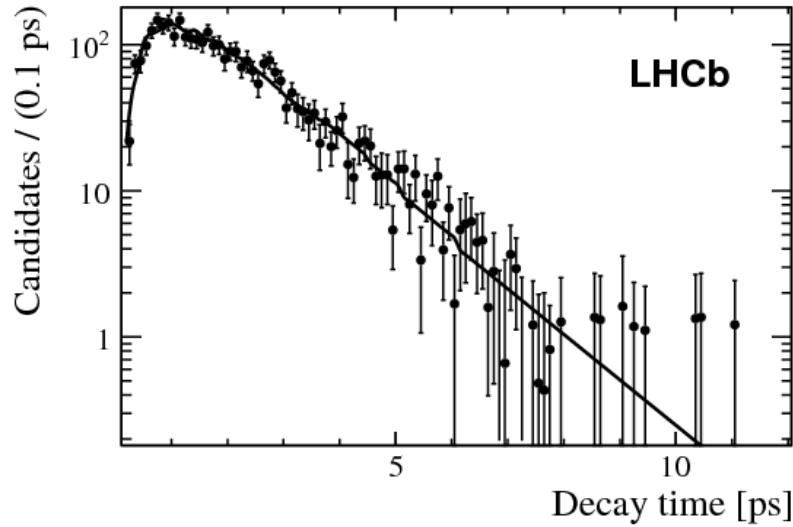


FIGURE 5.4: Distribution of the decay time for  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  signal decays used to extract a value for  $\phi_s$ . Discontinuities in the shape of the fit line are a result of the binned acceptance [88].

Measuring the decay time distribution of  $B_s^0 \rightarrow D_s^+ D_s^-$  candidates allows the extraction of  $\phi_s$ ; figure 5.4 shows the final fit in bins of decay time. The final data sample had the background component subtracted by performing a fit to the invariant mass of the  $B_s^0$ , and using this fit to weight events by their probability of being signal. The shape of the remaining distribution is dependent on a number of factors which are combined to give the full final fit to extract  $\phi_s$ . The values of  $\Gamma_s$ ,  $\Delta\Gamma_s$ , and  $\Delta m_s$  were all set to their world average values in the fit, with Gaussian constraints. The fit also includes efficiency contributions from the  $b$ -tagging algorithms used, along with the bin by bin decay time efficiencies. The initial event selection used was biased against  $B_s^0$  candidates with a low decay time, due to the selection requiring the  $B_s^0$  to have travelled in order to effectively identify it. This is accounted for in the full fit, and is the reason that there is a large drop in yield in the lowest decay time region of figure 5.4. The  $B_s^0$  candidates must be initially flavour-tagged so that the evolution of the system can be properly determined, and  $\phi_s$  can therefore be determined. An experimental measurement of  $\phi_s$  from  $B_s^0 \rightarrow D_s^+ D_s^-$  decays is of interest theoretically, as an open charm<sup>1</sup> final state may have a different penguin pollution compared to the  $B_s^0 \rightarrow J/\psi \phi$  decay which is most

<sup>1</sup>Meaning a final state where the  $c$  and  $\bar{c}$  are not in a bound state, *i.e.* a final state without a  $J/\psi$  or other excited charmonium resonance ( $\psi(2S)$  *etc.*).

commonly used to measure  $\phi_s$ . The LHCb result for  $\phi_s$  determined from  $B_s^0 \rightarrow D_s^+ D_s^-$  decays is:

$$\phi_s = (0.02 \pm 0.17 \pm 0.02) \text{ rad.}$$

The Standard Model prediction is [21]:

$$\phi_s = (0.0363_{-0.0015}^{+0.0016}) \text{ rad.}$$

The measurement is consistent with both the Standard Model prediction, and all other measurements of  $\phi_s$  using decays to different final states.

## 5.2 $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$

The remainder of this chapter focuses on the measurement of the branching fraction of the inclusive  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decay. The  $D_s^{*+}$  is an excited state of the  $D_s^+$ , and decays to a  $D_s^+$  and a photon 94.2% of the time, and to a  $D_s^+$  and a  $\pi^0$  the remaining 5.8% of the time. The branching fraction of the inclusive decay, where one, both, or neither  $D_s^\pm$  can pass through the excited state, was originally of great interest from a theoretical viewpoint.

The branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  was thought to provide a good approximation to the value of  $\Delta\Gamma_s/\Gamma_s$  (for a detailed discussion of the meaning and relevance of this quantity, please refer to section 2.2.1), the difference in decay width between the two  $B_s^0$  mass eigenstates divided by the average decay width, using equation 5.1 [89]:

$$\frac{\Delta\Gamma_s^{CP}}{\Gamma_s} = \frac{2 \times \mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-})}{1 - \mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-})}. \quad (5.1)$$



Here,  $\Delta\Gamma_s^{CP}$  is the difference in decay width between the  $CP$  eigenstates, and is equal to  $\Delta\Gamma_s$  in the limit that  $\phi_s = 0$ . This approximation was based on some important assumptions and limits:

1. that the decay of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  is predominantly  $CP$ -Even
2. that the decays saturate the value of  $\Delta\Gamma_s$
3. the Heavy Quark limit applies in the calculation.

More recent theoretical updates have cast doubt on the validity of these assumptions, suggesting that the decay modes containing at least one  $D_s^{*\pm}$  will have a non-negligible  $CP$ -Odd component, and also that other 3-body  $B_s^0$  decays will contribute to the value of  $\Delta\Gamma_s$  [90]. This update suggested that the contribution from the other decays could be a similar size to that from  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ , and hence that the original approximation to  $\Delta\Gamma_s/\Gamma_s$  is a poor one.

The decay  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  represents a pseudo-scalar decaying to two vectors. The spin structure of the decay means that the  $D_s^{*\pm}$  mesons can be either longitudinally or transversely polarised with respect to the decay axis. The longitudinally polarised states are  $CP$ -Even, whereas the transversely polarised decays are a superposition of  $CP$ -Even, when the polarisation vectors of the  $D_s^{*+}$  and  $D_s^{*-}$  are parallel, and  $CP$ -Odd, where the polarisation vectors of the  $D_s^{*+}$  and  $D_s^{*-}$  are perpendicular. This means that the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  decay can contain a  $CP$ -Odd component. The amplitudes of the three polarisation states are termed  $A_L$ ,  $A_{\parallel}$ , and  $A_{\perp}$ , and are related to the helicity eigenstates of the decay,  $A_0$ ,  $A_{+1}$ , and  $A_{-1}$ , as shown in 5.2 (where the helicity of a particle corresponds to the projection of its spin along its momentum vector) [91]:

$$\begin{aligned} A_L &= A_0 \\ A_{\parallel} &= \frac{A_{+1} + A_{-1}}{\sqrt{2}} \\ A_{\perp} &= \frac{A_{+1} - A_{-1}}{\sqrt{2}}. \end{aligned} \tag{5.2}$$

The three helicity amplitudes are complex and have a magnitude and phase, and can therefore interfere depending on the phase difference between the various eigenstates. In

order to separate the various components, a full angular analysis of the  $D_s^{*+}$  and  $D_s^{*-}$  decays would be required, requiring the reconstruction of the  $\gamma$  or  $\pi^0$  resulting from the  $D_s^{*\pm}$  decay. The neutral particle reconstruction was not performed in this analysis, and therefore a full angular analysis to extract the polarisation amplitudes was not possible. However, Belle were able to reconstruct the  $\gamma$  and  $\pi^0$  particles as the events are much cleaner in a  $e^- - e^+$  collider, and therefore were able to measure the longitudinally polarised fraction as  $f_L = 0.06_{-0.17}^{+0.18} \pm 0.03$  [92]. In the analagous decay  $B^0 \rightarrow D^{*+} D^{*-}$ , the contribution of the  $CP$  odd component,  $A_\perp$ , is estimated to be only 6% [93], and as such it can be expected that the  $CP$  odd component of the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  decay will be small.

This does not end the interest in a measurement of the branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ . The value of this branching fraction will still enable a calculation of  $\Delta\Gamma_s/\Gamma_s$ , although it would require further information from future analyses. Firstly, the  $CP$  odd component of the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  decay would have to be measured, which could be done using the method described in [94]. Furthermore, it would be necessary to measure the  $CP$  even component in all other  $b \rightarrow c\bar{c}s$  decays in order to assess the contribution to  $\Delta\Gamma_s$  [95]. A detailed discussion of theoretical predictions of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  branching fractions, and the predicted contribution of other modes to the value of  $\Delta\Gamma_s/\Gamma_s$ , is given in [96].

Furthermore, a precise measurement of the branching ratio  $B(b \rightarrow c\bar{c}s)$  is interesting, as it is an important ingredient in model independent searches for New Physics in  $B$  decays [97]. As one of the dominant contributions to this value, it is important to have an accurate measurement of the branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ .

This branching fraction has most recently been measured by CDF [98] and Belle [92], who found respectively:

$$B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.38 \pm 0.25 \pm 0.30 \pm 0.56)\%$$

$$B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (4.32_{-0.39-1.03}^{+0.42+1.04})\%$$

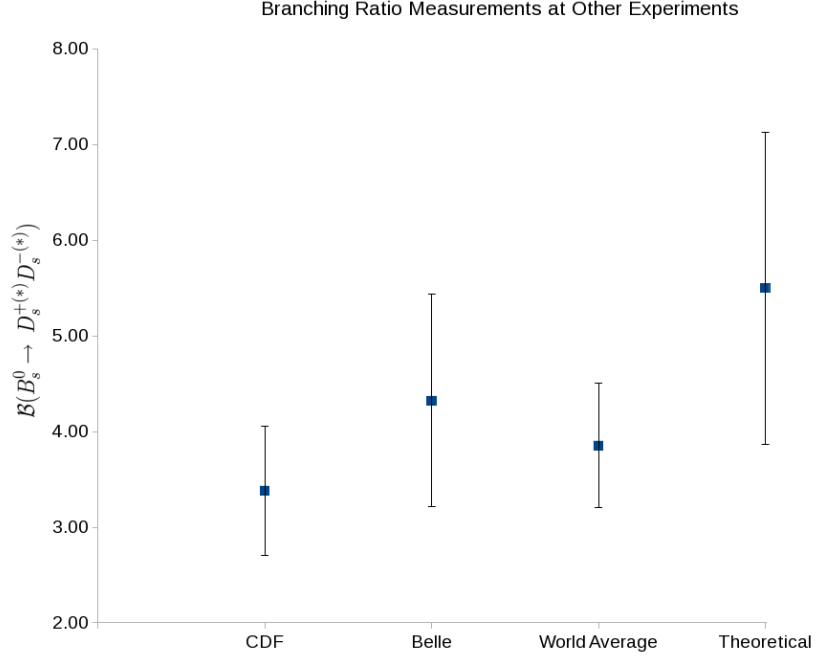


FIGURE 5.5: The previous measurements of the branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ , and the theoretical prediction from [96].

where the errors are statistical, systematic, and due to the normalisation for CDF, and statistical and systematic for the Belle result. These previous results, along with average of these results, and the theoretical prediction as taken from [96], are displayed in figure 5.5.

### 5.3 Selection

The analysis is performed using the entire 2011 LHCb data set, comprising approximately  $1.1 \text{ fb}^{-1}$  of proton-proton collisions with a centre of mass collision energy of  $\sqrt{s} = 7 \text{ TeV}$ .

In order to reduce the effect of systematic uncertainties on the final result, the branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  is measured relative to the decay  $B^0 \rightarrow D_s^\pm D^\mp$ , which has a very similar final state and kinematics.  $D_s^\pm$  candidates are reconstructed through the decay to  $K^+ K^- \pi^\pm$ , whereas  $D^\pm$  candidates are reconstructed through decays to a  $K^\mp \pi^\pm \pi^\pm$  final state. Therefore the final states of the signal and normalisation channel differ only through the interchange of a  $K^\pm$  for a  $\pi^\pm$ , meaning that many systematic

uncertainties will cancel, and only the differences resulting from this interchange will still contribute. The final branching fraction can then be calculated using equation 5.3. In this equation,  $\frac{f_d}{f_s}$  is the ratio of the  $B$  meson fragmentation fractions governing  $B_s^0$  and  $B^0$  production, and  $\epsilon_{\text{rel}}^{B^0/B_s^0}$  is the relative selection efficiency between the two channels.  $N_{B_s^0}$  and  $N_{B^0}$  are the total yields after full selection for the  $B_s^0$  and  $B^0$  respectively.

$$\mathcal{B}(B_s^0 \rightarrow D_s^{+(* )} D_s^{-(* )}) = \frac{f_d}{f_s} \cdot \epsilon_{\text{rel}}^{B^0/B_s^0} \cdot \frac{\mathcal{B}(D^+ \rightarrow K^- \pi^+ \pi^+)}{\mathcal{B}(D_s^+ \rightarrow K^- K^+ \pi^+)} \cdot \frac{N_{B_s^0}}{N_{B^0}} \cdot \mathcal{B}(B^0 \rightarrow D_s^\pm D^\mp). \quad (5.3)$$

The  $D_s^{*\pm}$  always decays to a  $D_s^\pm$  and either a  $\pi^0$  or a  $\gamma$ . In this analysis only the  $D_s^\pm$  is reconstructed in the decay of a  $D_s^{*\pm}$ , meaning the channels containing one or two  $D_s^{*\pm}$  are partially reconstructed and are identified through the missing invariant mass that results from not reconstructing the  $\pi^0$  or  $\gamma$ .

The first stage of event selection is to use a stripping line to select events with a candidate  $B$  meson decaying into two oppositely charged  $D$  mesons (see section 3.2.8 for more details about stripping lines in LHCb). In order to make this initial loose selection, the stripping line includes a number of different criteria which events must pass in order to be selected. The different variables used in the selection are described in brief in the following; for a more detailed description of each variable used, please refer to appendix F.

The stripping line includes loose cuts on a number of variables for candidate  $B_{(s)}^0$  mesons, including requiring a mass in the range 4750 MeV – 7000 MeV, a lifetime longer than 0.2 ps, a  $\text{DIRA}^2$  greater than 0.999, a  $\chi^2/\text{ndf}$  less than 10 for the end vertex (the  $B_s^0$  decay vertex), and an impact parameter  $\chi^2$  less than 25 with the primary vertex. There are also loose cuts applied to the  $D_s^\pm$  and  $D^\pm$  candidates; the maximal Distance Of Closest Approach ( $\text{DOCA}$ ) between the two charm daughters must be less than 0.5 mm, their daughters must have a transverse momentum ( $p_T$ ) sum of greater than 1.8 GeV, each must have a vertex  $\chi^2/\text{ndf}$  of less than 10, a  $\chi^2$  of greater than 36 to any primary vertex, and also a mass in the range 1769.62 MeV – 2068.49 MeV. Loose cuts are also

<sup>2</sup>Where  $\text{DIRA}$  represents the cosine of the angle between the momentum vector of the particle and the flight direction from primary to secondary vertex; this cut ensures the  $B_s^0$  is travelling forwards through the detector.

applied to all daughter candidates; all must have a total momentum ( $p$ ) greater than 100 MeV and a  $p_T$  greater than 1 GeV. They are also required to have a track  $\chi^2/\text{ndf}$  less than 4, and an impact parameter  $\chi^2$  greater than 4 to the primary vertex. Loose particle identification (PID) requirements are placed on all kaons and pions, using information from the RICH detectors. Finally, all events are required to have fewer than 500 long tracks (a long track is a track with a trajectory that traverses the VELO and also the tracking stations after the magnet), and must also be selected by specific triggers at the HLT2 stage of trigger which look for decays of  $B$  hadrons containing multiple particles in the final state [99](see section 3.2.8 for further information about the LHCb trigger).

As this line selects events matching the decay of any  $B$  meson to any two  $D$  mesons, it is necessary to make a number of further cuts in order to isolate both the signal and normalisation decays.  $D_s^+$  and  $D_s^-$  candidates are required to have a mass in the range 1940 MeV – 2000 MeV, and  $D^+$  and  $D^-$  candidates are required to have a mass between 1840 MeV – 1900 MeV. As the full  $B_s^0$  distribution is of interest, only high mass cuts are applied to  $B_s^0$  candidates; each candidate must have a mass lower than 5800 MeV. Only the main  $B^0$  peak is used for normalisation, and as such the  $B^0$  candidates are required to have a mass in the range 5050 MeV – 5500 MeV. There are also particle identification requirements placed on all final state kaons; the probability of particles being kaons is determined with the aid of the RICH detectors, and a minimum probability threshold is set which all kaons must fall above to be used in the analysis. All events are required to be triggered either on signal by the L0Hadron trigger (L0HadronTOS), or independent of signal by the L0Global trigger (L0GlobalTIS). The L0Hadron trigger looks for hadrons leaving high transverse energy deposits in the calorimeters, and the L0Global trigger requires that any of the various L0 triggers selected the event. Section 3.2.8 contains more details about the LHCb trigger system, and the requirements for TOS and TIS events.

The final cut applied makes use of a Boosted Decision Tree (BDT) [100, 101, 102] in order to remove combinatorial background. The BDT is trained using simulated signal events from the three signal channels (one, two or zero  $D_s^{*\pm}$ ) as the signal sample, and events from the data distribution sidebands as the background sample, where a sideband event is required to have a  $B_s^0$  mass greater than 5600 MeV, and also both  $D_s^+$  and  $D_s^-$  masses less than 1930 MeV or greater than 2010 MeV. The discriminating variables used in the training of the BDT are listed in table 5.1. This set of variables was selected from

an initial wide range of additional variables, by considering the effect that each variable had on the overall effectiveness of the BDT performance; the ranking of the variables as indicated by the BDT training is shown in table 5.2. The 14 variables (2 for the  $B_s^0$ , and 6 each for the  $D_s^+$  and  $D_s^-$ ) selected were seen to be the most useful in the BDT, giving a similar performance to the BDT with more variables, but preventing any issues coming from the finite sample sizes used for signal and background in the BDT training and testing. As the training and testing signal samples are taken from simulated events, the agreement between the simulated and actual distributions for all of the BDT variables is demonstrated in appendix A. The signal and background samples are split randomly and evenly into two separate samples, one for training, and the other for testing the BDT.

The BDT was trained using the AdaBoost boosting algorithm to reweight events that are wrongly classified at the end of previous trees. Events were reweighted using a parameter  $\alpha = \log((1 - \epsilon)/\epsilon)$ , where  $\epsilon$  is the error rate for that event. The forest consisted of 400 trees, and each tree had a maximum depth of 3 levels. No pruning was used in the BDT training. Also tested were a Fisher Discriminant and a Neural Net Multi-Layer Perceptron (MLP). The MLP was trained using the Cross-Entropy (CE) estimator type, used  $N + 5$  hidden layers (where  $N$  is the number of input variables, leading to a total of 19 hidden layers), and was trained over 600 cycles. As shown in figure 5.7, the BDT method was observed to be the best performing, and as such this was the method used in the analysis.

$B_s^0$ Variables	$D_s^+$ and $D_s^-$ Variables
Impact Parameter $\chi^2$ to own PV	Impact Parameter $\chi^2$ to own PV
Transverse momentum	Transverse momentum
	DIRA to own PV
	Lifetime
	DIRA to origin vertex
	Product of daughter transverse momentum

TABLE 5.1: Variables used in the BDT.

Figure 5.6 shows the overtraining check that is performed, indicating that the BDT does not suffer from overtraining. Figure 5.7 shows the background rejection vs signal efficiency curves for the various methods tested in TMVA, justifying the use of the BDT due to its superior performance compared to the Neural Net and Fisher methods. The

BDT Variable	Variable Importance $\times 10^{-2}$
Ds2 transverse momentum	9.705
Ds1 transverse momentum	9.508
$B_s^0$ Impact Parameter $\chi^2$ to own PV	8.582
$B_s^0$ transverse momentum	8.334
Ds2 DIRA to origin vertex	8.185
Ds1 lifetime	7.118
Product of Ds1 daughter transverse momentum	6.658
Ds1 DIRA to own PV	6.568
Ds1 DIRA to origin vertex	6.504
Product of Ds2 daughter transverse momentum	6.392
Ds2 DIRA to own PV	6.325
Ds2 Impact Parameter $\chi^2$ to own PV	6.309
Ds2 lifetime	5.104
Ds1 Impact Parameter $\chi^2$ to own PV	4.707

TABLE 5.2: The ranking of variable importance for the 14 variables used to train the BDT, listed in order of importance. Note that Ds1 and Ds2 are arbitrary labels for the two  $D_s^\pm$  candidates in the events, and represent a random mixture of  $D_s^+$  and  $D_s^-$  candidates.

correlation matrices for the input variables for both signal and background events can be seen in appendix A.

The BDT cut is chosen to maximise the value of the total  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data yield divided by the square root of the total signal and background events, with background events taken from the full range from 4750 MeV – 5800 MeV. This is performed by fitting the data events remaining after various BDT cuts, and taking the signal and background yields from these fits. The final mass distributions with various BDT cuts are shown in appendix C. The top plot in figure 5.8 shows how the significance changes as the selection progress, and how it varies with the BDT cut applied; the plot shows that the significance is stable and maximum for any BDT cut from -0.10 to 0.00, so the cut value was chosen to allow the most signal events in the final plot, at -0.10.

## 5.4 Backgrounds

There are a number of further backgrounds present in the full mass range, mainly due to either the misidentification of another hadron as a kaon, or partial reconstruction of other decays. These backgrounds must be either eliminated from the data, or be understood well enough that an extra distribution can be added to the final mass fit.

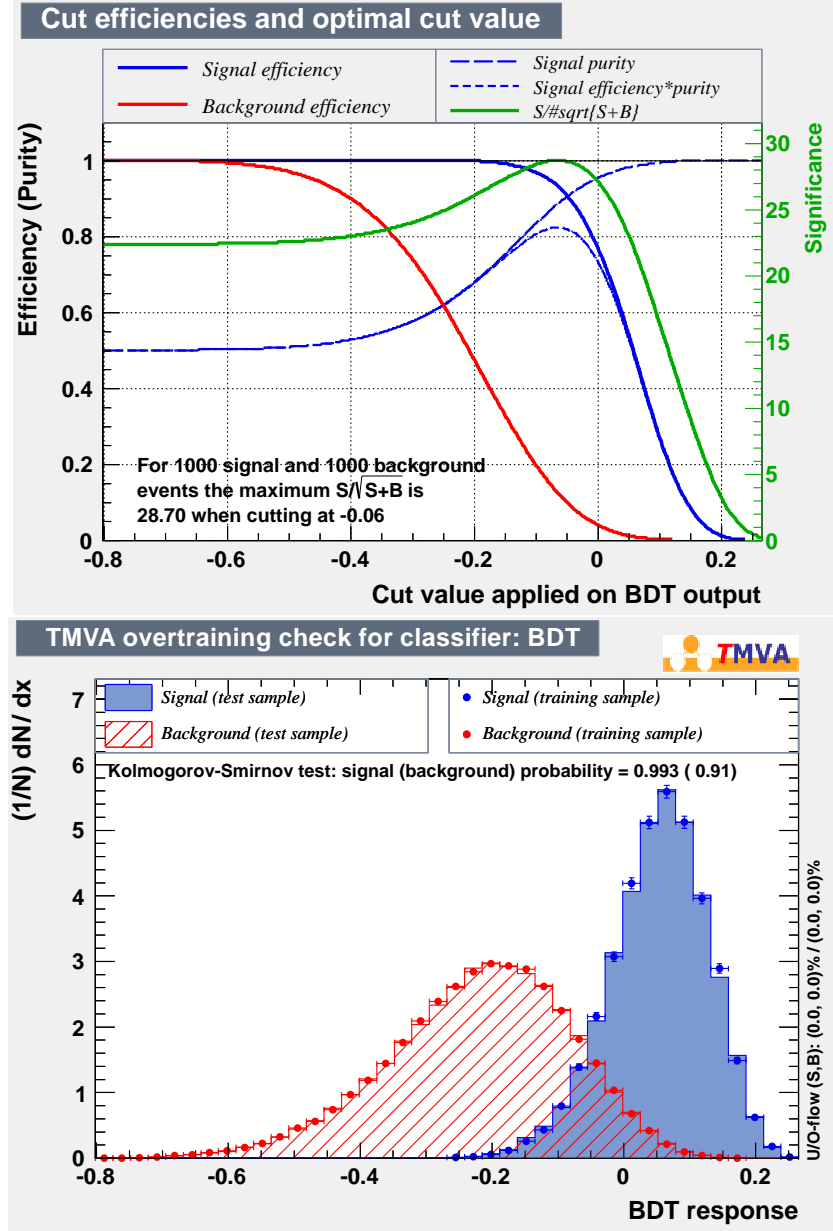


FIGURE 5.6: (Top) Plot showing the various figures of merit for the BDT using 1000 signal and 1000 background events. The efficiency of the selection is the fraction of events remaining at a particular BDT cut, and the purity is the fraction of the total final sample that was signal, (Bottom) Plot showing the signal and background training and testing samples for the BDT overlaid.

#### 5.4.1 $B^0 \rightarrow D_s^\pm D^\mp$

The normalisation channel itself provides a background to the signal channel. If a pion in the decay of the  $D^+$  is misidentified as a kaon, and the resulting reconstructed mass falls in the  $D_s^+$  mass window, this decay can fake a signal event. In order to remove these events, a veto is applied to remove any  $D_s^+$  or  $D_s^-$  consistent with being a  $D^+$



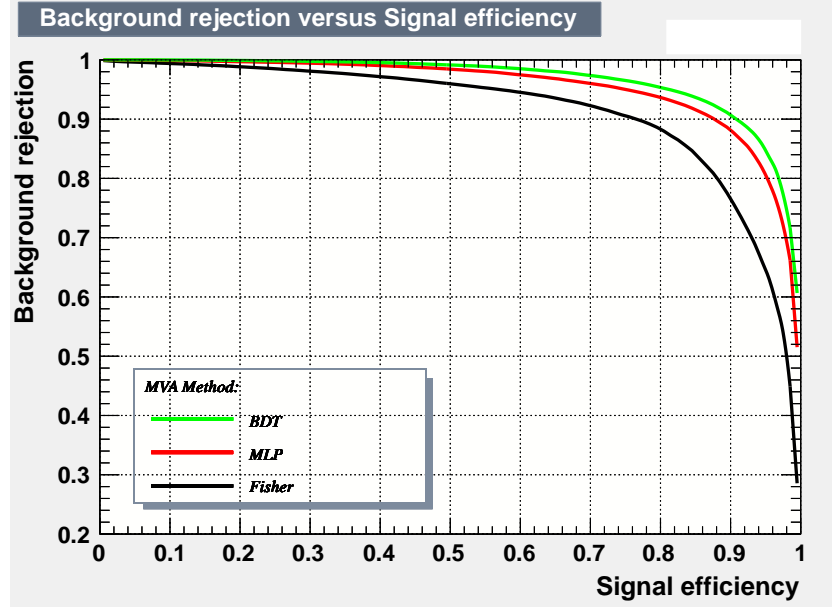


FIGURE 5.7: Background rejection vs signal efficiency for the TMVA methods tested, showing the superior performance of the BDT.

or  $D^-$  under the exchange of the same sign kaon in the decay for a pion. In order to virtually remove all of this background, a fairly wide cut is applied; any  $D_s^\pm$  candidate with a mass in the range 1835 – 1905 MeV after the switch is rejected by the veto. The effect also occurs in the normalisation channel, where the signal channel can provide a background to the decay. Therefore any  $D^+$  with a mass in the range 1950 – 1990 MeV after exchanging either of the pions in the decay for kaons is rejected. The efficiencies of the signal channel vetoes are given in section 5.5.1, whilst the normalisation channel veto efficiencies are given in section 5.5.2.

This veto technique was found to be the most efficient at removing this background whilst keeping as much signal as possible. The effect of tightening the PID cut in order to remove this background was examined, and the results are shown in figure 5.9. The plots in this figure show the mass distribution of the  $D_s^+$  candidate when the  $K^+$  is switched to a  $\pi^+$ . The peak from the misidentified  $D^+$  is clearly seen, and is only removed by using a cut of greater than 0.5 (the value is an estimation of the probability of a particle identified as a kaon actually being a kaon).

A cut of greater than 0.5 leaves a final data sample of 1595 for analysis, and a cut of greater than 0.6 leaves only 950 data events; this is compared to 4004 events when using the veto as described previously. For this reason the veto was used in the analysis to

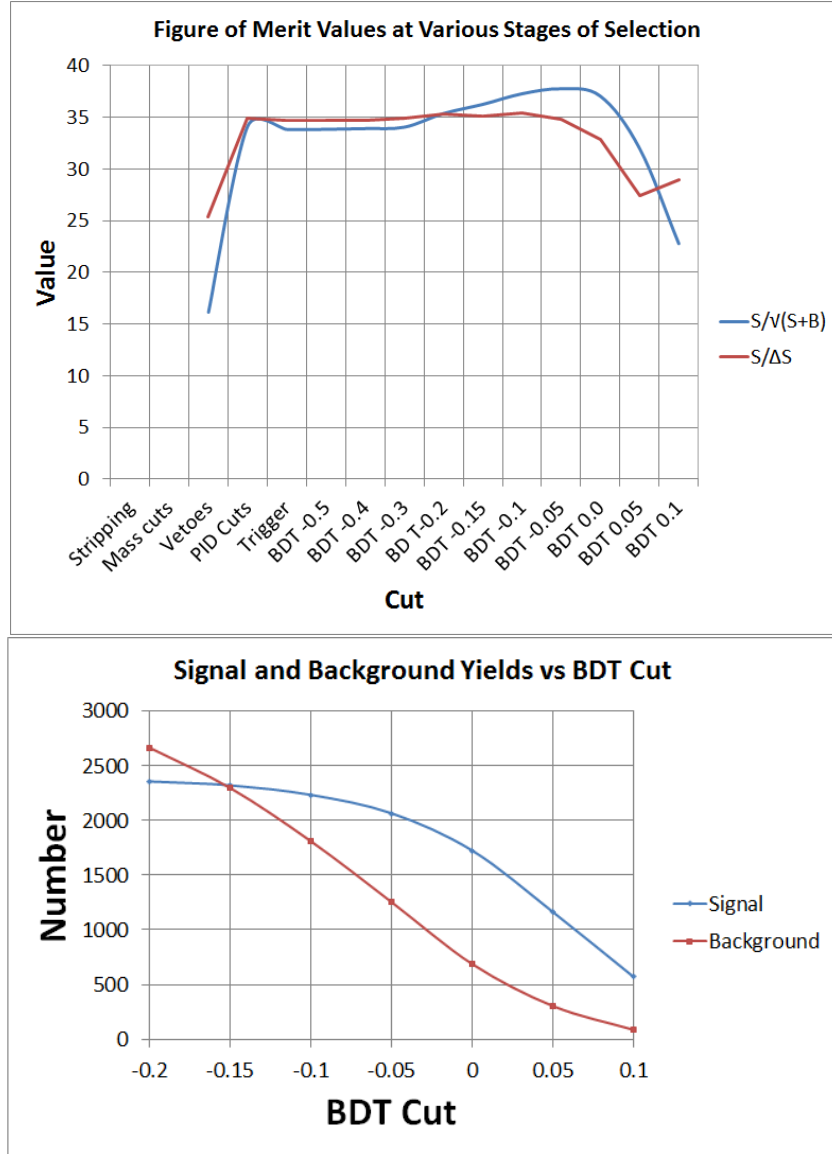


FIGURE 5.8: (Top) Figure of merit value vs BDT cut. Each consecutive stage includes all of the previous selection criteria applied up to that point.  $S/\Delta S$  represents the total signal yield divided by the statistical uncertainty on the yield, as indicated by the full data fit. For the stripping line selection, and also when using only mass cuts, the full fit to the  $B_s^0$  and  $D_s^\pm$  mass distributions failed, hence the lack of a value in both of these columns, (Bottom) Signal and background yields vs BDT cut.

remove the misidentified  $B^0 \rightarrow D_s^\pm D^\mp$  background, with the PID cut left at  $> 0.1$  as before.

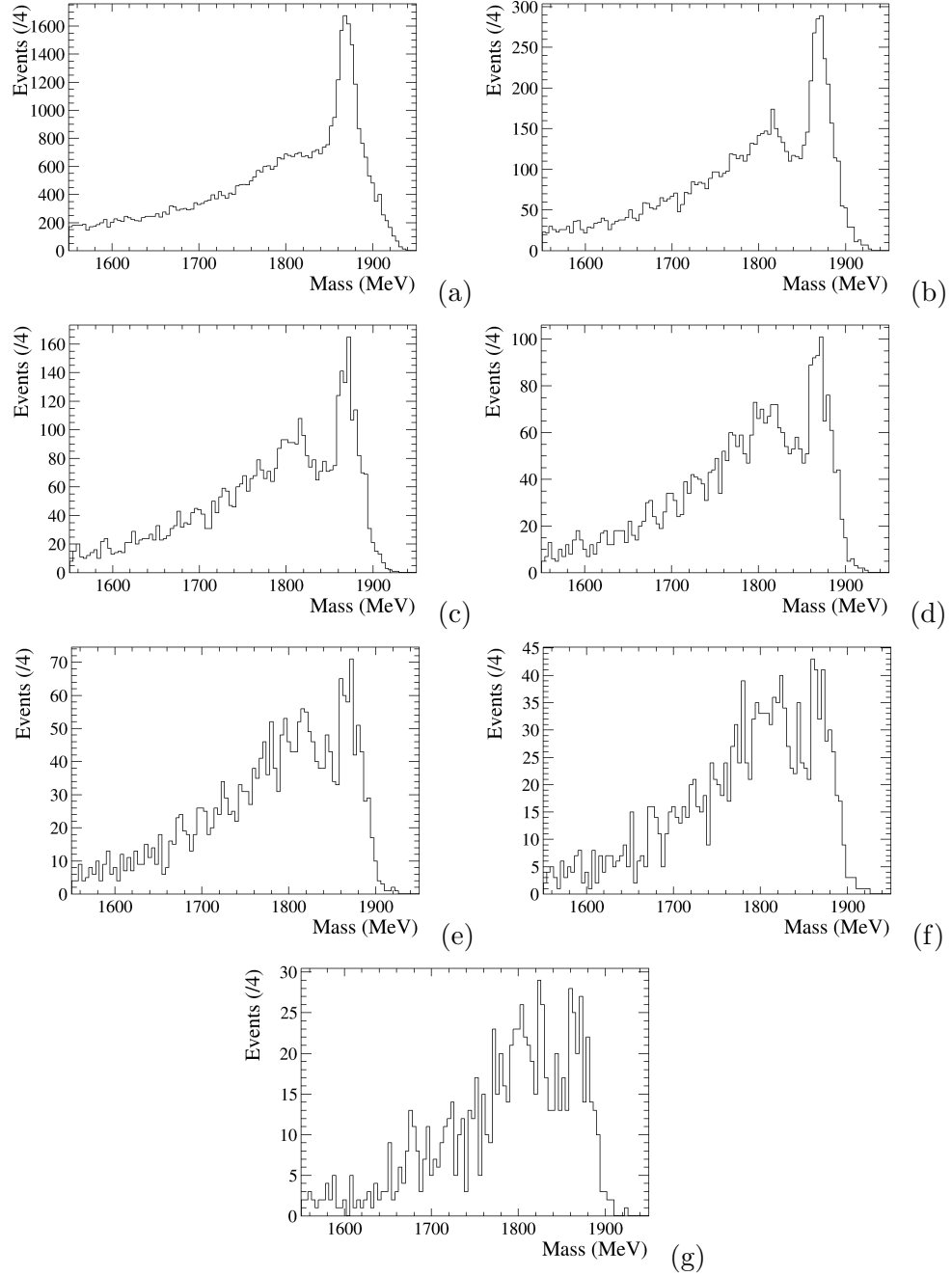


FIGURE 5.9: Plot showing the invariant mass distribution of the  $D_s^+$  data candidate when the  $K^+$  mass hypothesis is switched to a  $\pi^+$ , to illustrate the presence of a background from misidentified  $D^\pm \rightarrow K^- \pi^+ \pi^+$  decays. The effect of using a tighter PID cut to remove this background is displayed, using a cut of (a) 0.0, (b) 0.1, (c) 0.2, (d) 0.3, (e) 0.4, (f) 0.5, and (g) 0.6.

### 5.4.2 $\Lambda_b^0 \rightarrow \Lambda_c^+ D_s^-$

This channel provides a background to the signal decay in a similar way to the previous channel. The  $\Lambda_c^+$  decays to a proton, a kaon, and a pion; if the proton is misidentified as a kaon, and the resulting reconstructed mass falls in the  $D_s^+$  mass window, then this decay can fake a signal event. This background is removed by vetoing any  $D_s^+$  or  $D_s^-$  candidates consistent with being a  $\Lambda_c^+$  under the exchange of the same sign kaon for a proton. A veto on any  $D_s^+$  candidate with a mass in the range 2271 – 2301 MeV after the switch was found to be sufficient to remove this background, as tested by examining the number of simulated  $\Lambda_b^0 \rightarrow \Lambda_c^+ D_s^-$  found to pass the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  selection. The efficiencies in the signal and normalisation channels are again available in sections 5.5.1 and 5.5.2 respectively. As can be seen in these sections, this invariant mass veto gives a very high efficiency as it is, and thus it was not deemed necessary to make use of any proton PID information.

### 5.4.3 $B_s^0 \rightarrow D_s(2460)^+ D_s^-$

It is important to consider the possibility of decays of the form  $B_s^0 \rightarrow D_s^{+**} D_s^-$  contributing to the final mass distribution in the data. Of the various  $D_s^{+**}$  mesons the best understood is the  $D_s(2460)^+$ , and therefore the  $B_s^0 \rightarrow D_s(2460)^+ D_s^-$  decay is chosen to model this background. The  $D_s(2460)^+$  mostly decays to a  $D_s^+$  and some combination of photons and  $\pi^0$ 's. For simulation studies, the  $D_s(2460)^+$  is generated to decay to the following states in the given fractions, where the fractions are taken from the PDG and normalised to 100%:  $D_s^{*+} \pi^0 = 58.82\%$ ,  $D_s^+ \gamma = 23.53\%$ ,  $D_s^+ \pi^+ \pi^- = 5.89\%$ ,  $D_s^{*+} \gamma = 5.88\%$ , and  $D_s^+ \pi^0 \gamma = 5.88\%$ . As with the signal channel decays involving the excited  $D_s^+$  states, the reconstructed  $D_s^+$  is real, but there is missing mass in the  $B_s^0$  reconstruction owing to the neutral decay products not being reconstructed. It is therefore impossible to remove these background decays in any similar way to the previous two channels. Dealing with this background properly is also complicated by the fact that the branching fraction is as yet unmeasured. However, by using the branching ratios of the corresponding  $B^0 \rightarrow D_s^{+**} D^-$  decays, which have been measured, relative to the  $B^0 \rightarrow D_s^\pm D^\mp$  decays as guidance, it can be expected that the branching fraction of these  $B_s^0 \rightarrow D_s(2460)^+ D_s^-$  decays could be of the same order as the signal channel.

In order to account for this background in the data fit, simulated events were generated and the efficiency of the selection on the events in this channel was measured. It was found that the efficiency in this channel relative to the main signal channel was 0.548; it was therefore necessary to add a term for this decay to the final data fit. This is covered in more detail in section 5.6.6.

#### 5.4.4 Further potential backgrounds

Checks were also made to see if there would be any expected contribution from decays of the form  $B^\pm \rightarrow D_{(s)}^+ D_{(s)}^- h^\pm$ , where the hadron corresponds to either a pion or a kaon. The decay  $B_s^0 \rightarrow D_s^+ D_s^-$  was also considered where one of the  $D_s^\pm$  mesons decays to  $K^\mp \pi^\pm \pi^\pm$  instead. Also considered was the decay  $B_s^0 \rightarrow D_s^+ K^+ K^- \pi^-$ , where the  $K^+ K^- \pi^-$  combination does not come from a  $D_s^\pm$ , but could have an invariant mass in the selection range. In all cases, simulated events were put through the full selection, and the efficiency of the selection for these events was calculated. It is shown in table 5.3 that the efficiency was negligible in all cases except for the decay  $B^\pm \rightarrow D_s^+ D_s^- \pi^\pm$ , where presumably the decay is partially reconstructed without the bachelor pion and the resulting decay falls within the  $B_s^0$  mass window. However, this decay is heavily suppressed with its branching fraction expected to be less than  $10^{-5}$ , much smaller than the signal decay. For this reason, this decay is neglected as a potential background contribution. It is deemed unnecessary to include contributions from any of these decays in the final data fit.

Decay Mode	$\epsilon_{rel}$ to $B_s^0 \rightarrow D_s^+ D_s^-$	Branching Ratio to $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$
$B_s^0 \rightarrow D_s^+ D_s^- (K^+ \pi^- \pi^-)$	$0.0000 \pm 0.0000$	$0.121 \pm 0.009$
$B_s^0 \rightarrow D_s^+ K^+ \pi^- \pi^-$	$0.0082 \pm 0.0026$	$0.0073 \pm 0.0027$
$B_s^0 \rightarrow D_s^+ K^+ K^- \pi^-$	$0.0132 \pm 0.0021$	$< 0.0073$
$B^\pm \rightarrow D_s^+ D_s^- \pi^\pm$	$0.1320 \pm 0.0103$	Unobserved
$B^\pm \rightarrow D_s^+ D_s^- K^\pm$	$0.0118 \pm 0.0030$	Unobserved
$B^\pm \rightarrow D_s^+ D^- \pi^\pm$	$0.0009 \pm 0.0009$	Unobserved
$B^\pm \rightarrow D_s^+ D^- K^\pm$	$0.0063 \pm 0.0022$	Unobserved
$B^\pm \rightarrow D^+ D^- \pi^\pm$	$0.0000 \pm 0.0000$	Unobserved
$B^\pm \rightarrow D^+ D^- K^\pm$	$0.0000 \pm 0.0000$	$0.0049 \pm 0.0022$

TABLE 5.3: Potential backgrounds and their selection efficiencies relative to the  $B_s^0 \rightarrow D_s^+ D_s^-$  decay channel. Note that the efficiency uncertainties arise from the finite size of the MC samples used to test the various potential background channels.

## 5.5 Efficiencies

### 5.5.1 Signal channel efficiencies

Table 5.5 shows the efficiencies of the various cuts on the truth matched simulated signal events for the various signal decay channels. The efficiencies in the table are presented as relative to the previous cuts so, for instance, the BDT efficiency is measured on the events that remain after both the stripping line and mass cuts have been applied. The mass and BDT cut efficiencies are both measured by counting the number of truth matched MC events both before and after the cut is applied.

The mass cut efficiency is defined as the total number of events that simultaneously pass the mass selection criteria relating to the  $B_s^0$ , and both of the  $D_s^\pm$  daughters. As the efficiency here is taken from truth matched MC, the shape of the  $D_s^\pm$  was studied in the three different channels to ensure there was no difference in data that could change the efficiency. To separate the channels, the final mass distribution was split into three sections relating to areas that were predominately  $D_s^+ D_s^+$  (5300 MeV – 5410 MeV),  $D_s^{*\pm} D_s^\mp$  (5140 MeV – 5300 MeV), and  $D_s^{*+} D_s^{*-}$  (4900 MeV – 5050 MeV). The two floating parameters in the  $D_s^\pm$  fit, the mean value of the distribution and a parameter governing

the width of the peak, were then recorded after the full selection had been applied, and also with a tighter BDT cut in case the combinatorial background shape changed the shape of the  $D_s^\pm$  distributions. The results are summarised in table 5.4 and the 6 distributions are also shown in figure 5.10; this confirmed that there was no distinct difference in the distributions, as all parameters agreed within uncertainties.

BDT	Parameter	$B_s^0 \rightarrow D_s^+ D_s^-$	$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$
Normal	Mean	$1969.34 \pm 0.29$	$1969.74 \pm 0.21$	$1969.20 \pm 0.24$
	Width	$10.28 \pm 1.47$	$11.98 \pm 0.99$	$11.01 \pm 1.98$
Tight	Mean	$1969.30 \pm 0.54$	$1969.87 \pm 0.39$	$1969.13 \pm 0.39$
	Width	$12.62 \pm 0.95$	$12.28 \pm 0.61$	$10.60 \pm 0.72$

TABLE 5.4: Mean and width parameters for the  $D_s^\pm$  distributions in the various channels under consideration. ‘Normal’ BDT cut refers to the cut used in the final data selection, whereas ‘Tight’ is a BDT cut at 0.1 which removes virtually all combinatorial background.

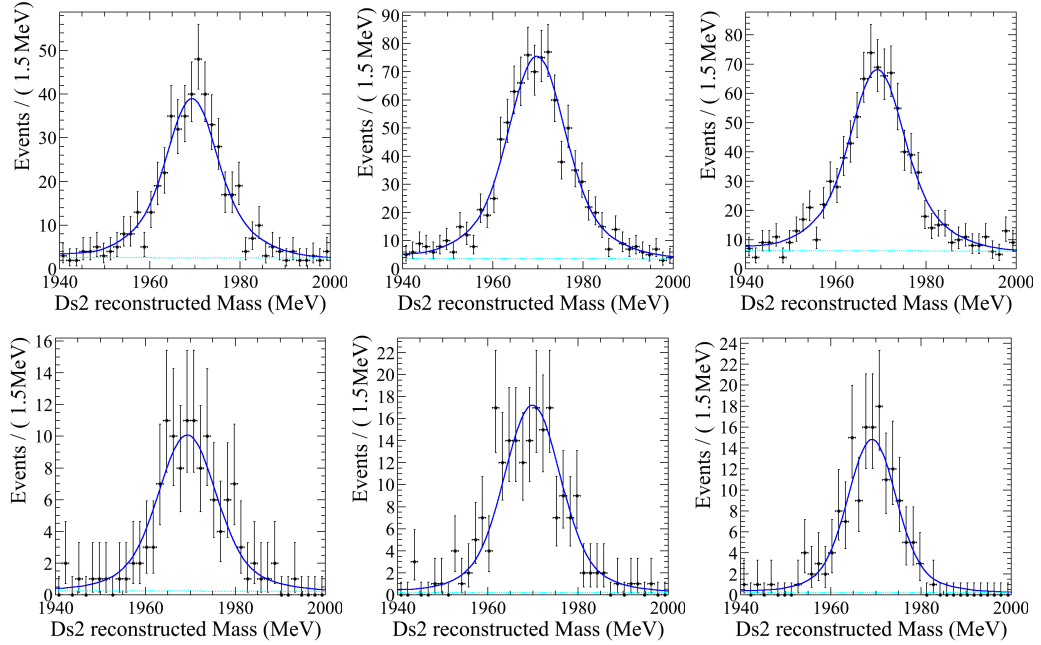


FIGURE 5.10: (Top) Left to right:  $D_s^\pm$  distributions for decays in the  $D_s^+ D_s^-$ ,  $D_s^{*\pm} D_s^\mp$ , and  $D_s^{*+} D_s^{*-}$  channels, after the full selection including the normal BDT cut is applied, and (Bottom) left to right: the same plots, but with a tighter BDT cut at 0.1.

The efficiency of the PID cuts is calculated using the LHCb `PIDCalib` package [103], and is therefore not present in this table. The `PIDCalib` package reweights the track PID according to variables that the PID is sensitive to, based on clean samples of kaons and pions from  $D^*$  decays. The reweighting variables used are the total momentum ( $p$ ), pseudorapidity ( $\eta$ ) (previously defined in equation 3.2), and number of tracks (nTrack). The signal distributions are set from data, after the full selection has been applied and the large majority of the background has been removed with a BDT cut. The selection is not restricted to the fiducial `PIDCalib` volume but, of the events used in the final analysis, only a small number ( $< 2\%$ ) of kaons have a  $p_T > 100 \text{ GeV}/c$ , and only a very small number of events have either a track outside the acceptance, or with more than 500 tracks. The efficiency of the PID cuts is found to be  $82.40 \pm 0.18\%$ . A systematic uncertainty of 3% is applied to account for the binning choice, which was estimated by rerunning the calibration with a different binning and comparing the result with the result using a default binning scheme.

The efficiency of the trigger selection applied is taken from simulation; the systematic uncertainty resulting from this method is considered in section 5.8.8.

Cut	$B_s^0 \rightarrow D_s^+ D_s^-$	$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$
Stripping Line	$0.01362 \pm 0.00008$	$0.0125 \pm 0.0001$	$0.0110 \pm 0.0001$
Mass Cuts	$0.894 \pm 0.006$	$0.878 \pm 0.010$	$0.883 \pm 0.010$
BDT Cut	$0.979 \pm 0.007$	$0.966 \pm 0.011$	$0.967 \pm 0.011$
$B^0 \rightarrow D_s^\pm D^\mp$ Veto	$0.487 \pm 0.005$	$0.503 \pm 0.008$	$0.489 \pm 0.008$
$\Lambda_b^0 \rightarrow \Lambda_c^+ D_s^-$ Veto	$0.963 \pm 0.010$	$0.963 \pm 0.016$	$0.959 \pm 0.016$
Trigger requirement	$0.966 \pm 0.007$	$0.967 \pm 0.011$	$0.966 \pm 0.011$
Total	$0.005596 \pm 0.000056$	$0.005120 \pm 0.000084$	$0.004408 \pm 0.000078$

TABLE 5.5: Efficiencies of the various cuts on the three channels of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ ; each efficiency is presented relative to the previous cut, with the full efficiency presented in the final row.

It should be noted that although the  $B^0 \rightarrow D_s^\pm D^\mp$  veto is somewhat inefficient, it was deemed necessary in order to avoid contaminating the low mass region of the final  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data distribution. The veto used here was sufficient to entirely remove the  $B^0 \rightarrow D_s^\pm D^\mp$  background as determined from MC, alleviating the need for a further



term in the full mass fit. Whilst the efficiency could be improved by using a looser veto, this would allow some of these background events into the final signal sample, and necessitate the inclusion of a small peak in the mass distribution to account for these events. This would certainly be possible at the main  $B_s^0 \rightarrow D_s^+ D_s^-$  peak, and is the preferred method in other analyses in this signal channel. However, the distribution of events coming from  $B^0 \rightarrow D_s^\pm D^\mp$  where one or both of the daughters come from a  $D_s^{*\pm}$  or  $D^{*\pm}$  is much harder to determine, which would affect the performance of the fit in the lower mass region of the final data sample. For this reason, and considering that the uncertainty on the final result is systematically dominated rather than statistically, the less efficient veto was used in order to avoid using a much more complicated final PDF.

### 5.5.2 Normalisation channel efficiency

The efficiencies of the selection criteria on the truth matched  $B^0 \rightarrow D_s^\pm D^\mp$  simulated events are presented in table 5.6. The PID cut efficiency is again calculated using the PIDCalib package, and is found to be  $(84.23 \pm 0.11)\%$ .

Cut	$B^0 \rightarrow D_s^\pm D^\mp$
Stripping Line	$0.01416 \pm 0.00009$
Mass Cuts	$0.885 \pm 0.006$
BDT Cut	$0.976 \pm 0.007$
$D^\pm$ Veto	$0.687 \pm 0.006$
$D_s^\pm$ Veto	$0.648 \pm 0.007$
$A_b^0 \rightarrow A_c^+ D_s^-$ Veto	$0.982 \pm 0.008$
Trigger requirement	$0.968 \pm 0.007$
Total	$0.005349 \pm 0.000056$

TABLE 5.6: Efficiencies of the various cuts on the  $B^0 \rightarrow D_s^\pm D^\mp$  channel; each efficiency is presented relative to the previous cut, with the full efficiency presented in the final row.

### 5.5.3 Generator efficiencies

The generator efficiency is the efficiency with which the MC events generated in a particular channel pass selection criteria that ensure the event is produced in the LHCb acceptance, and with greater than a minimum momentum. The generator level efficiencies of the three signal channels, and the normalisation channel, are shown in table 5.7. The uncertainties on these efficiencies are negligible, and thus are ignored.

Channel	Gen. Efficiencies
$B_s^0 \rightarrow D_s^+ D_s^-$	$0.1184 \pm 0.0003$
$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$0.1127 \pm 0.0005$
$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$	$0.1061 \pm 0.0005$
$B^0 \rightarrow D_s^\pm D^\mp$	$0.1071 \pm 0.0002$

TABLE 5.7: Generator level efficiencies for all signal and normalisation channels.

### 5.5.4 Relative efficiencies

The relative efficiency of the normalisation channel to each signal channel is shown in table 5.8, and these are calculated by multiplying together the generator, PID, trigger, and selection efficiencies. The efficiency of the  $B^0 \rightarrow D_s^\pm D^\mp$  decay relative to the inclusive  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decay efficiency is also presented; this is calculated by using a weighted average of the individual signal channel efficiencies, weighted according to the yields of each channel, as shown in equation 5.4 (see section 4.1 for further details of yields). Here,  $N_{D_s^+ D_s^-}$  is the yield in the  $B_s^0 \rightarrow D_s^+ D_s^-$  channel,  $N_{D_s^{*\pm} D_s^\mp}$  is the yield in the  $B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$  channel,  $N_{D_s^{*+} D_s^{*-}}$  is the yield in the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  channel,  $N_{sig}$  is the total signal yield, and  $\epsilon_x$  is the efficiency in the individual channel  $x$ :

$$\epsilon_{rel}^{B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}} = \frac{\left(\frac{N_{D_s^+ D_s^-}}{N_{sig}} \times \epsilon_{D_s^+ D_s^-}\right) + \left(\frac{N_{D_s^{*\pm} D_s^\mp}}{N_{sig}} \times \epsilon_{D_s^{*\pm} D_s^\mp}\right) + \left(\frac{N_{D_s^{*+} D_s^{*-}}}{N_{sig}} \times \epsilon_{D_s^{*+} D_s^{*-}}\right)}{\epsilon_{D_s^\pm D^\mp}}. \quad (5.4)$$

Channel	$\epsilon_{rel}$
$B_s^0 \rightarrow D_s^+ D_s^-$	$0.953 \pm 0.045$
$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$1.095 \pm 0.054$
$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$	$1.351 \pm 0.067$
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	$1.144 \pm 0.053$

TABLE 5.8: Relative efficiencies for all signal channels to  $B^0 \rightarrow D_s^\pm D^\mp$ , where the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  efficiency is a weighted average of the three individual efficiencies, weighted according to the relative yields in the individual channels.

## 5.6 Invariant mass distributions and fits

The fit to each data set is a simultaneous unbinned extended maximum likelihood fit to the invariant mass of the  $B_s^0$  and the two  $D_s$  candidates. A maximum likelihood fit seeks to maximise the value of the natural logarithm of the sum of the individual likelihood functions for all of the events in the distribution being fitted. As the observables of interest are the yields of the various components of the fit rather than the fraction of each component present, the fit is extended by including a Poisson term that depends on the total fitted event yield  $n'$ . The extended maximum likelihood function used for all of the fits in the rest of this analysis is shown in equation 5.5 [104]:

$$\mathcal{L} = \frac{e^{n'}}{N!} \prod_{i=1}^N \sum_{j=1}^m n_j P_j(m_i). \quad (5.5)$$

In this equation,  $N$  is the total number of events,  $j = 1, \dots, m$  represents the various components of the fit,  $n_j$  is the total yield of the component  $P_j$  of the fit,  $i = 1, \dots, N$  is the number of the event, and  $m_i$  is the invariant mass of event  $i$ . The total fitted event yield  $n'$  equals the sum of the yields of each individual component. For example, in the full  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  fit, the  $P_{1,\dots,j}$  components are the 3 signal channels, the  $D_s(2460)^+$  background, and the combinatorial background. Maximising the value of  $-\ln \mathcal{L}$  gives the most likely values of the yields for each one of the components.

The shapes for all of the fits are decided by fitting to simulated events for all signal decays, and by fitting the distribution of data events taken from a wrong sign version of

the stripping line (defined in section 5.6.5) in the case of combinatorial backgrounds. The shape of the  $B_s^0 \rightarrow D_s(2460)^+ D_s^-$  background is also determined by fitting simulated events. In all cases where a pull plot is included, the pull is calculated at each point by determining the  $\chi^2$  between the model and the data.

### 5.6.1 $B_s^0 \rightarrow D_s^+ D_s^-$

The fit to simulated events in this channel, along with the pull plot for the fit, is shown in figure 5.11. The peak is described by the combination of a Gaussian and a Crystal Ball function [105] with common means, where the Crystal Ball shape models the low mass tail of the  $B_s^0$ , caused by final state radiation. The Crystal Ball shape is defined as shown in equation 5.6, where  $x$  is the invariant mass,  $m$  is the invariant mass of the peak of the Gaussian portion of the function,  $\alpha$  determines where the tail begins,  $\sigma$  is the width, and  $n$  determines the power law tail of the function:

$$f(x) = \frac{\left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{1}{2}\alpha^2\right)}{\left(\frac{n}{|\alpha|} - |\alpha| - x\right)^n} \quad \text{when } x < -|\alpha|, \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right) \quad \text{when } x > -|\alpha|. \quad (5.6)$$

The  $n$  and  $\alpha$  parameters for the Crystal Ball shape found in this fit to simulated events are fixed in the data fit. Also fixed to the simulation values are the ratio of the Gaussian width to the Crystal Ball fit, and the fraction of the fit that is taken from the Crystal Ball function. The shared mean of the two functions and the width of the Crystal Ball component are floated in the final data fit.

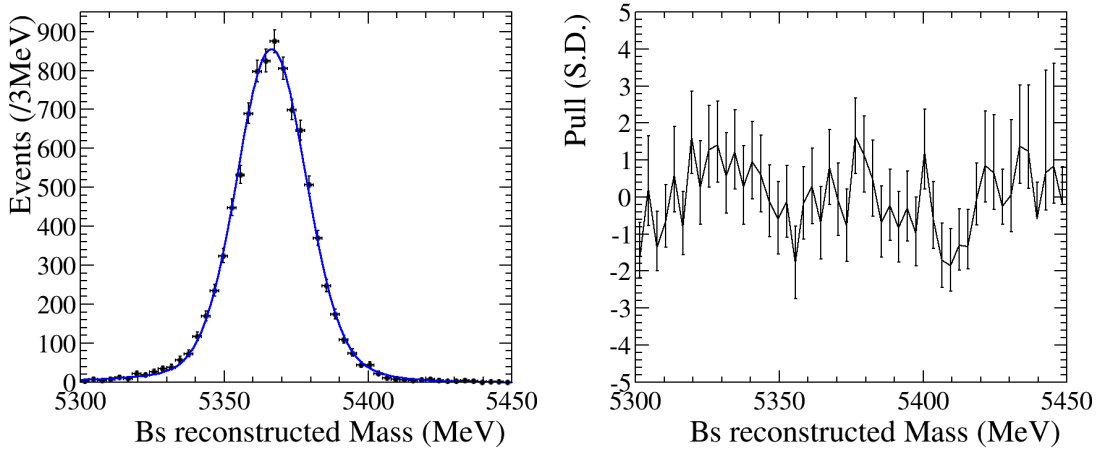


FIGURE 5.11: Mass distribution and fit for simulated events in the  $B_s^0 \rightarrow D_s^+ D_s^-$  decay channel with the pull plot for the fit.

### 5.6.2 $B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$

Figure 5.12 shows the fit to simulated events where there is one  $D_s^{*+}$  present. The fit uses a combination of two separate Gaussian functions, with different means and widths. The broadened structure is caused by the difference in mass distribution between events where the  $D_s^{*\pm}$  decays to  $D_s^\pm \gamma$  and those where it decays to  $D_s^\pm \pi^0$ . The mass distribution of the  $D_s^{*\pm}$  reconstructed without a  $\gamma$  or  $\pi^0$  depends on the spin of the missing particle; the  $\pi^0$  is spin-0, which results in a ‘double-horn’ structure for the mass distribution reconstructed without the  $\pi^0$ . The  $D_s^{*\pm}$  decays in simulation via a photon 94.2% of the time, and via a  $\pi^0$  the other 5.8%, as listed in the PDG [106]. The invariant mass distribution of simulated  $B_s^0 \rightarrow D_s^{*\pm} (\rightarrow D_s^\pm \pi^0) D_s^\mp$  events is shown in 5.13, clearly showing the horns that result from the  $\pi^0$  decays. The value for the ratio of the widths of the two Gaussians, and also the fraction of the first Gaussian in the function, are fixed in the final fit to data. The mean of each Gaussian is also fixed to the value found from this fit to simulation.

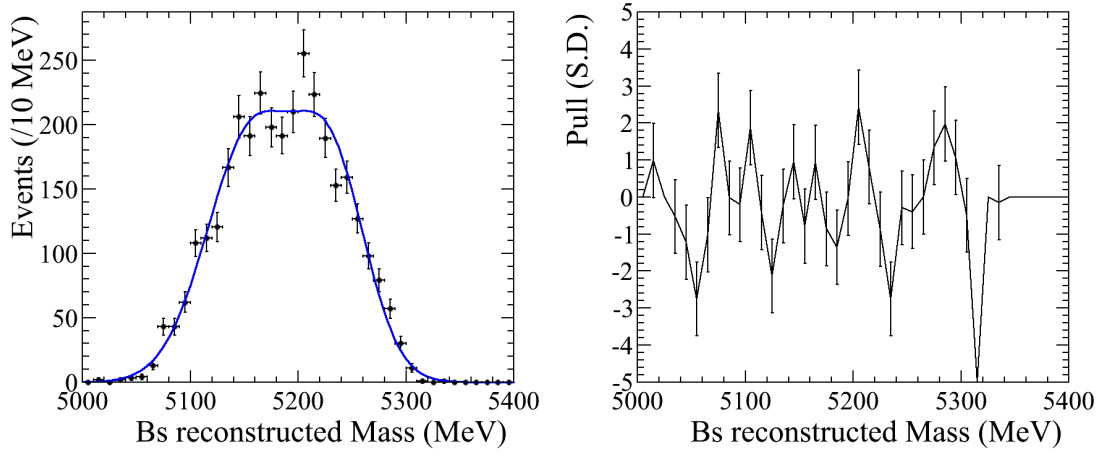


FIGURE 5.12: Mass distribution and fit for simulated events in the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  decay channel with the pull plot for the fit.

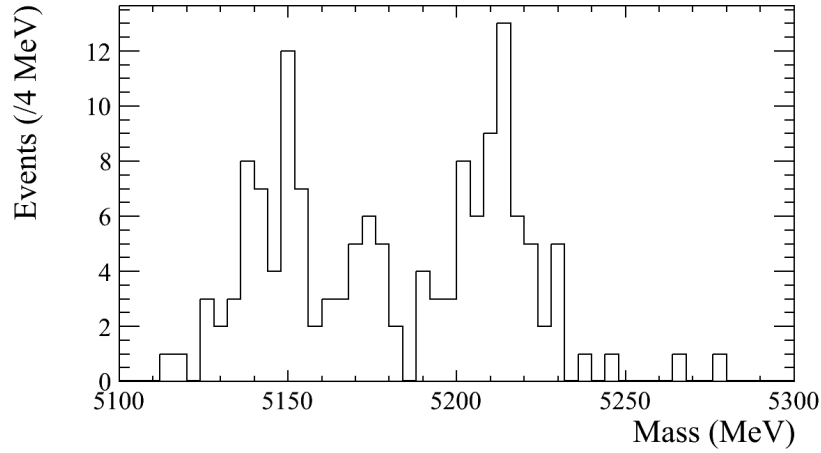


FIGURE 5.13: Mass distribution for simulated  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  events where the  $D_s^{*+}$  decays to a  $D_s^+$  and a  $\pi^0$ . The helicity related double horn structure is clearly seen. The double peak structure results from the angular distribution of the decay products due to the spins of the particles involved. The momentum of the  $\pi^0$  depends on the angle at which it is emitted, and as the  $\pi^0$  is not reconstructed the missing invariant mass also depends on the angle.

### 5.6.3 $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$

Events where there are two  $D_s^{*+}$  decays are found to be well described by a single Gaussian function, as shown in figure 5.14. The value for the mean of this Gaussian is fixed in the final data fit.

These decays are only weakly helicity dependent, and this is accounted for in the MC generation by using the **SVV-HELAMP** [54] decay package in **EVTGEN**, and specifying no helicity dependence, by using the values: **SVV-HELAMP** 1.0 0.0 1.0 0.0 1.0 0.0. Here, **SVV-HELAMP** means scalar to vector vector helicity amplitude, and the 6 arguments correspond to the 3 complex helicity eigenstates possible in the decay; the first 2 represent the magnitude and phase of the  $A_{+1}$  eigenstate, the second 2 represent the magnitude and phase of the  $A_0$  eigenstate, and the final 2 are the magnitude and phase of the  $A_{-1}$  eigenstate (for further details regarding the helicity eigenstates and their relation to polarisation eigenstates, please see section 5.2). The values here are standard **EVTGEN** values for specifying no helicity dependence of the final decay amplitudes.

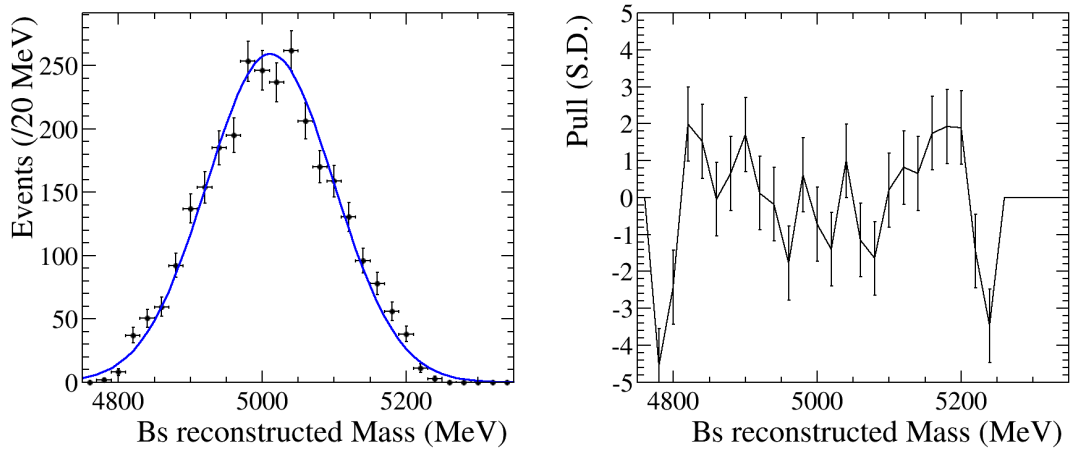


FIGURE 5.14: Mass distribution and fit for simulated events in the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  decay channel with the pull plot for the fit.

To verify that the mass PDF is indeed independent of the helicity of the decay, a small sample of MC events was produced with the **SVV-HELAMP** parameters altered to indicate a greater helicity dependence, using the values: **SVV-HELAMP** 0.4904 0.0 0.7204 0.0 0.4904 0.0. These values are taken from the **EVTGEN** values used to model the weak helicity

dependence displayed in the decay  $B^0 \rightarrow D_s^+ D^{*-}$ . The mass distribution was examined using the same fit model, and the results are shown in table 5.9. The two parameters agree within uncertainties, and therefore it was deemed that the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  partially reconstructed PDF is indeed helicity independent.

Helicity Dependence	Gaussian mean	Gaussian width
Weak	$5011.13 \pm 1.69$	$88.30 \pm 1.22$
Mild	$5019.18 \pm 5.81$	$89.67 \pm 4.26$

TABLE 5.9: Mean and width of the Gaussian used to fit the  $B_s^0 \rightarrow D_s^{*+} D_s^{*\mp}$  distribution in both the central MC produced with weak helicity dependence, and the local production with mild helicity dependence. The values are consistent within uncertainties in both cases.

#### 5.6.4 $D_s^\pm$ Peaks

The  $D_s^\pm$  peaks for the signal and also the  $D_s(2460)^+$  background are described by the same type of fit as used for the main  $B_s^0$  peak, with the combination of Gaussian and Crystal Ball functions with a shared mean, as shown in figure 5.15. Both peaks are required to use the same set of parameters governing the shape. This simulation fit is used to fix the  $n$  and  $\alpha$  values of the Crystal Ball, the ratio of the two widths, the shared mean and also the Crystal Ball fraction of the decay, in the final data fit.



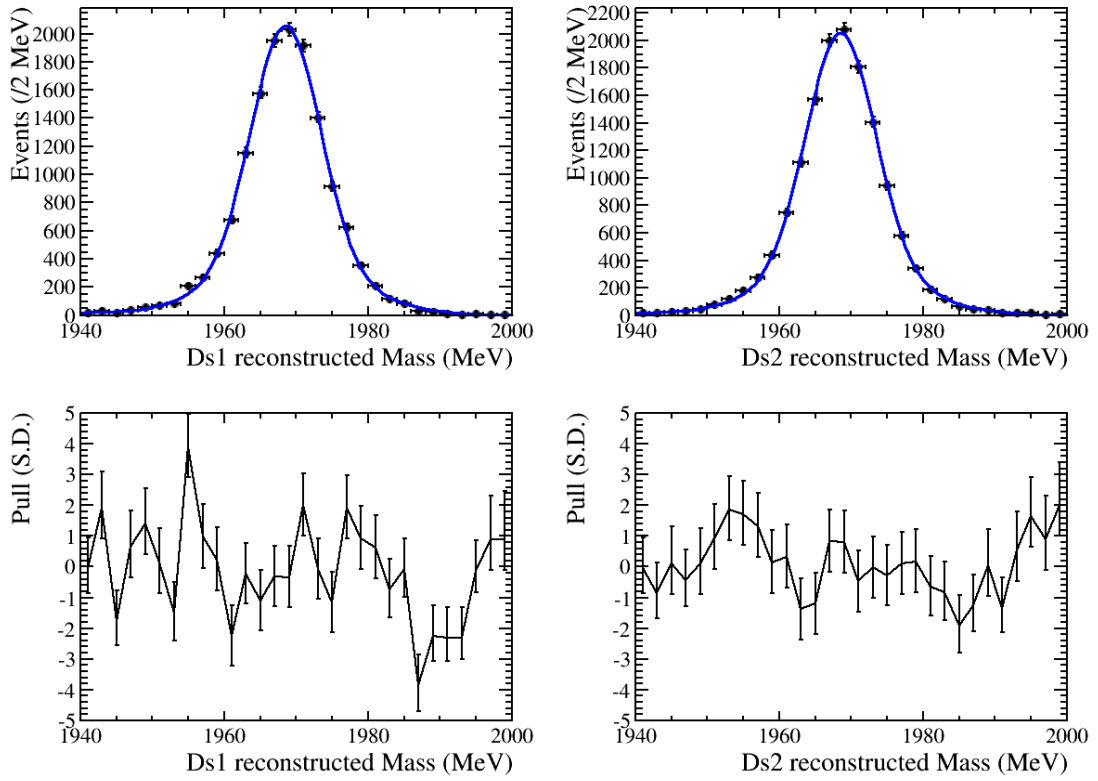


FIGURE 5.15: Mass distributions and fit for the  $D_s^\pm$  peaks in simulated events, with the pull plots for the fits. Ds1 and Ds2 are arbitrary labels for the two  $D_s^\pm$  candidates.

### 5.6.5 $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ combinatorial background

The full data fit also includes a background component that results from combinations of particles that are unrelated to one another, but appear to form two  $D_s^\pm$  mesons that come from a  $B_s^0$  decay. The combinatorial background is modelled by fitting to data events from a wrong sign version of the stripping line used for signal. This line is identical to the signal line but instead looks for  $B_s^0 \rightarrow D_s^\pm D_s^\pm$  events, which are clearly unphysical, but model the pairing of two unrelated  $D_s^\pm$ . The wrong sign line also has a pre-scaling of 0.1, and therefore allows only a tenth of the events to pass. The wrong sign data events that remain after all cuts have been applied are modelled using a first-order Chebyshev polynomial, as shown in figure 5.16. The shape parameters are floated in the full data fit, as is the yield. Whilst there are not many events remaining in the wrong sign data distribution after all cuts have been applied, the fit was also tested on the full wrong

sign distribution before any of the selection criteria, which left a much higher number of events in the distribution (shown in figure 5.17). It was found that all of the parameters agree within one standard deviation both before and after the cuts were applied.

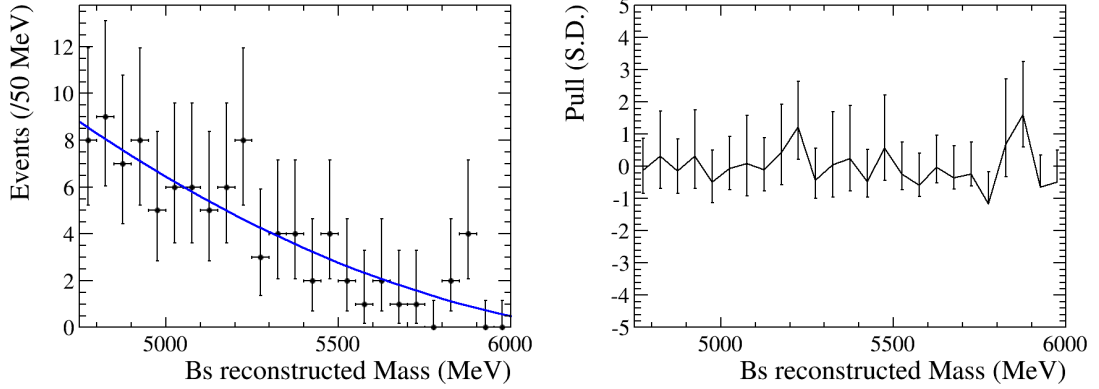


FIGURE 5.16: Mass fit and pull plot of the combinatorial background in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  channel, as taken from data from the wrong sign stripping line selection, after all cuts have been applied.

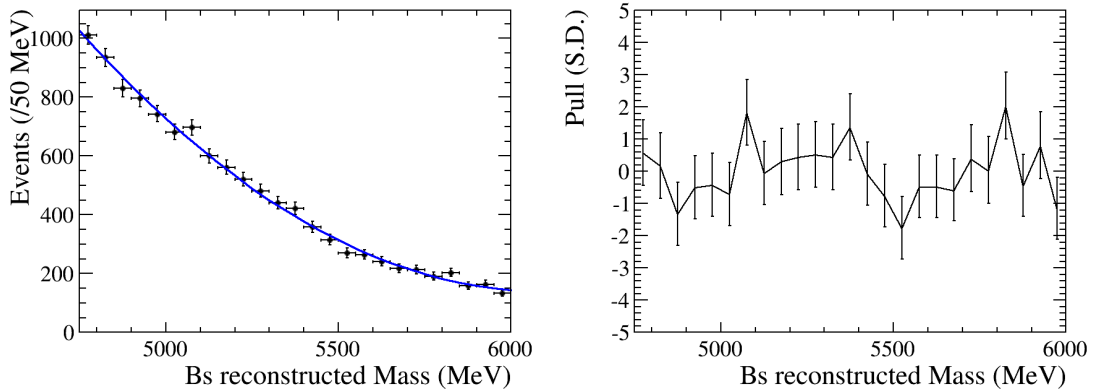


FIGURE 5.17: Mass fit and pull plot of the combinatorial background in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  channel, as taken from data from the wrong sign stripping line selection but before any of the offline selection is applied.

### 5.6.6 $B_s^0 \rightarrow D_s(2460)^+ D_s^-$ background

The shape of the  $B_s^0 \rightarrow D_s(2460)^+ D_s^-$  background is modelled with an Argus function [107], the result of which is shown in figure 5.18. The definition of the Argus function is shown in equation 5.7, where  $x$  represents the invariant mass,  $m$  is the mass cut off (the mass where the function has a value of 0),  $c$  is a constant that governs the rate at which the function decreases at the low end, and  $p$  changes the rate at which the function increases below the mass cut off:

$$f(x) = x \left( 1 - \left( \frac{x}{m} \right)^2 \right)^p \exp \left( c \left( 1 - \left( \frac{x}{m} \right)^2 \right) \right). \quad (5.7)$$

The parameters of this shape are fixed in the full data fit to the values found when fitting to simulated events, and the yield is allowed to vary. The  $D_s(2460)^+$  is generated to decay to the following states in the given fractions:  $D_s^{*+} \pi^0 = 58.82\%$ ,  $D_s^+ \gamma = 23.53\%$ ,  $D_s^+ \pi^+ \pi^- = 5.89\%$ ,  $D_s^{*+} \gamma = 5.88\%$ , and  $D_s^+ \pi^0 \gamma = 5.88\%$ . These fractions are as taken from the PDG, but with the branching fractions normalised to 100% as required for the MC generation.

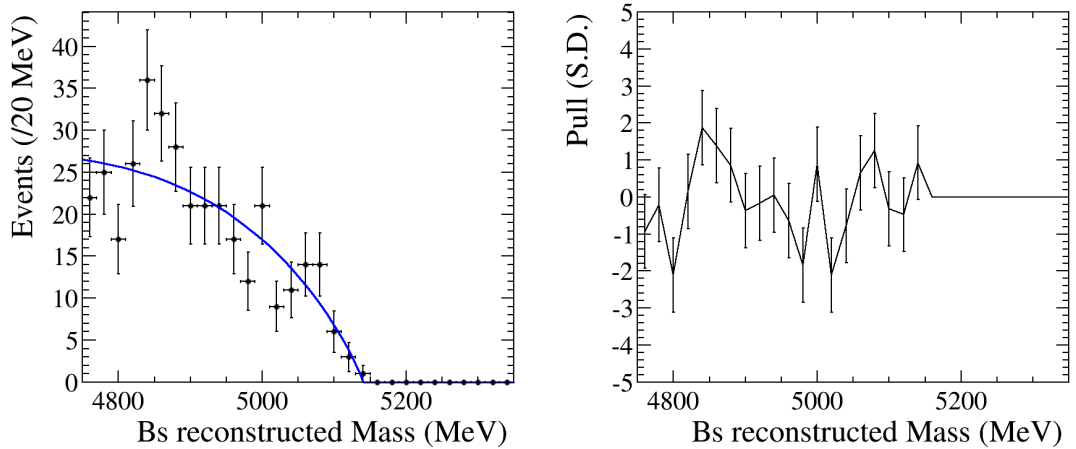


FIGURE 5.18: Mass fit and pull plot of the  $B_s^0 \rightarrow D_s(2460)^+ D_s^-$  background.

Whilst the mass cut off point for the function is perhaps overly aggressive, a fairly large systematic is applied to take into account the uncertainty on the mass distribution. This is covered in further detail in section 5.8.1. The final data plots produced when analysing the systematic for this background are available in appendix D.

Following studies of MC simulated  $B_s^0 \rightarrow D_s^+(2317)D_s^-$  decays, the background shape in figure 5.18 was deemed to also account for any potential background from this channel, albeit with the inclusion of a larger systematic in the  $D_s^{*+}D_s^-$  channel to account for the slightly higher mass cut off. This is again documented in detail in section 5.8.1.

### 5.6.7 $B^0 \rightarrow D_s^\pm D^\mp$

The reconstructed distribution in the main peak for the normalisation channel is also fitted with a combination of Gaussian and Crystal Ball functions, as with the main  $B_s^0$  peak and the  $D_s^\pm$  peaks. The fit and the resulting pull plot for the simulated events is shown in figure 5.19.

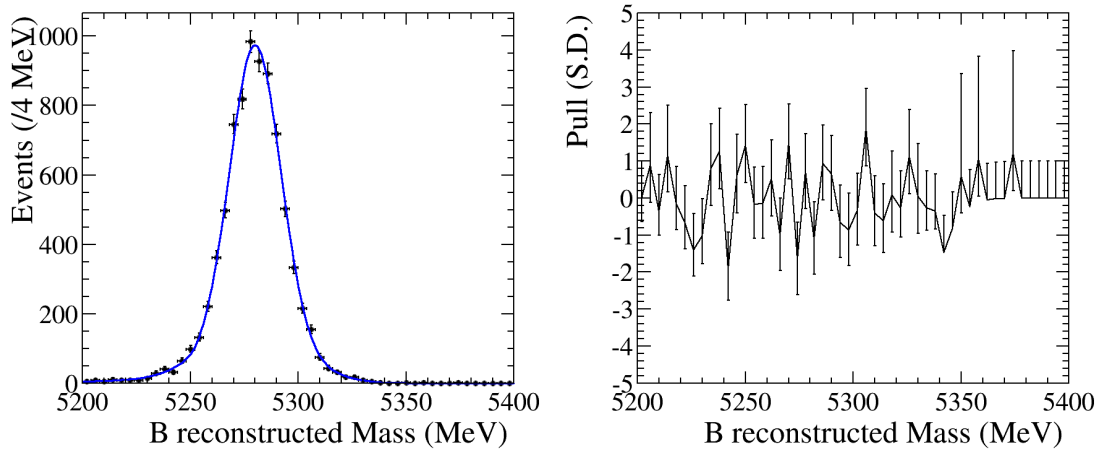


FIGURE 5.19: Mass fit and pull plot of the  $B^0 \rightarrow D_s^\pm D^\mp$  decay from simulated events.

### 5.6.8 $B^0 \rightarrow D_s^{*\pm} D^\mp$ and $B^0 \rightarrow D_s^\pm D^{*\mp}$

Compared to the central peak of the signal channel, which was fitted with a combination of two Gaussians, the central peak of the normalisation channel is instead fitted with

the sum of three Gaussian functions, with different means. This accounts for the fact that in the case of the normalisation mode, either the  $D_s^\pm$  or the  $D^\pm$  can pass through an excited state. The  $D^{\pm*}$  decays with a  $\pi^0$  much more often than the  $D_s^\pm$ , and hence the ‘horned’ structure resulting from the conservation of angular momentum is much more apparent in this case. The full fit to simulated events in both of these channels, and the resulting pull plot, is shown in figure 5.20.

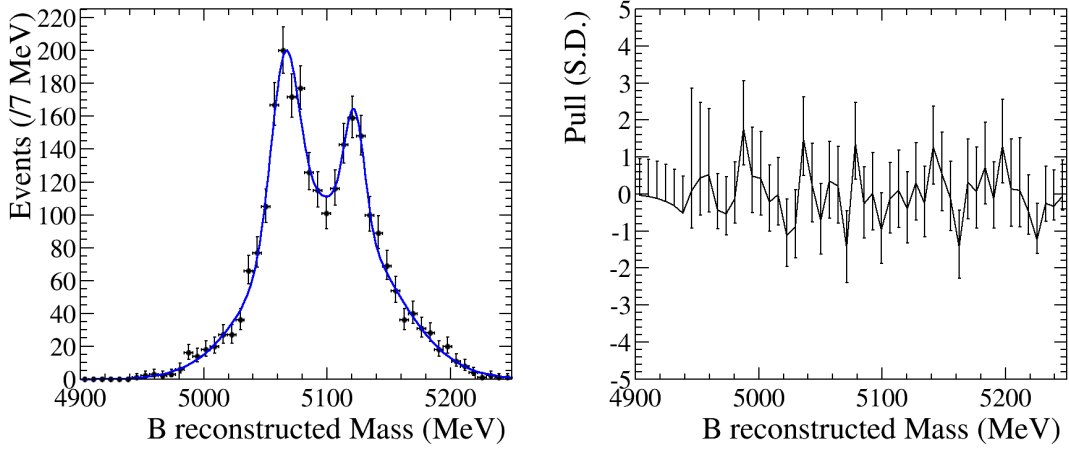


FIGURE 5.20: Mass fit and pull plot of the  $B^0 \rightarrow D_s^{\pm*} D^\mp$  and  $B^0 \rightarrow D_s^\pm D^{\mp*}$  decays from simulation.

### 5.6.9 $B^0 \rightarrow D_s^{*\pm} D^{*\mp}$

As was the case in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  fit, the lowest mass peak is fitted with a single Gaussian function. The fit for these simulated decays is shown in figure 5.21. It should be noted that practically all of this peak falls outside of the  $B^0$  selection mass window, with no events at the actual  $B^0$  peak.

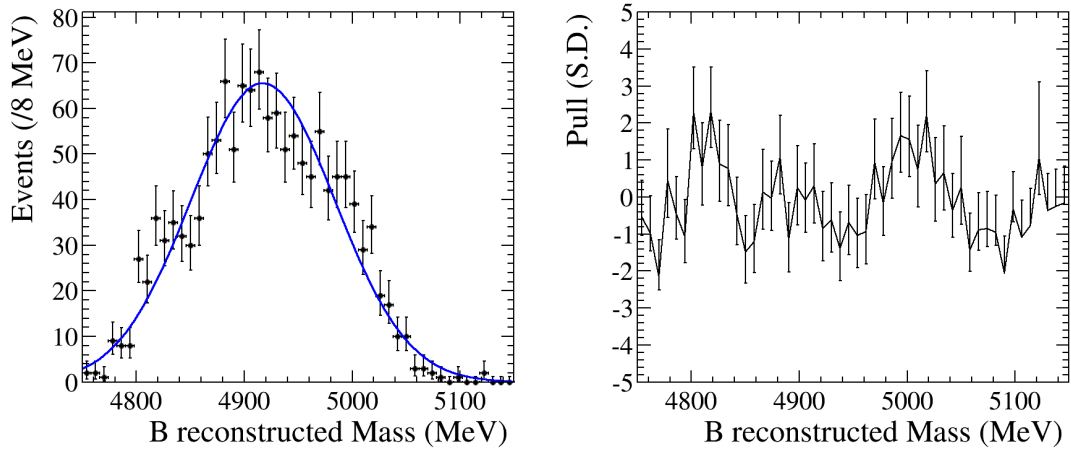
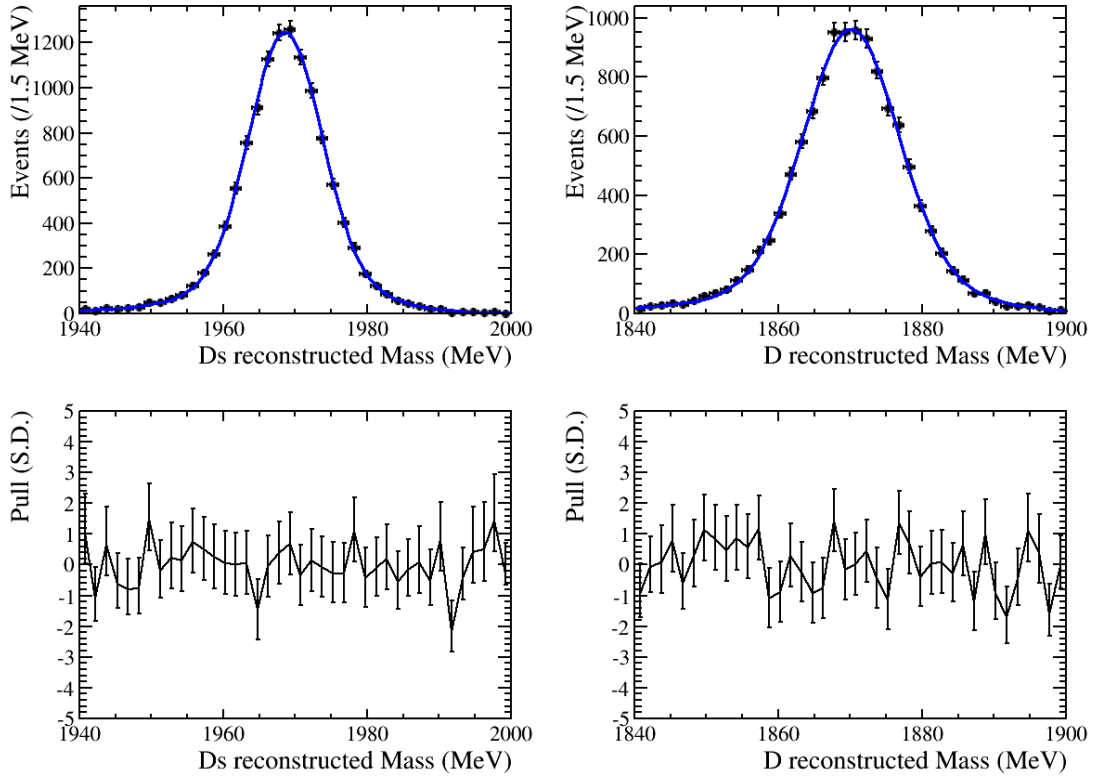


FIGURE 5.21: Mass fit and pull plot of the  $B^0 \rightarrow D_s^{\pm*} D^{\mp*}$  decay channel from simulation.

#### 5.6.10 $D_s^\pm$ and $D^\pm$ peaks

The two  $D$  meson peaks are again fitted with a combination of Gaussian and Crystal Ball functions, but this time the parameters are different between the two, owing to differences in the widths of the  $D_s^+$  and  $D^+$  particles. The distributions in the final fit to data events contain both signal events and background events from  $D^{*\pm}$  and  $D_s^{*\pm}$  decays. The fits and their corresponding pull plots are shown in figure 5.22.

FIGURE 5.22: Mass fits and pull plots of the  $D_s^\pm$  and  $D^\pm$  peaks.

### 5.6.11 $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$ combinatorial background

As with the signal channel, the combinatorial background in the normalisation channel is modelled using a Chebyshev polynomial, as shown in figure 5.23, fitted to data events taken from a wrong sign version of the stripping line used for the signal. The shape parameters are again floated in the full data fit, and the yield is allowed to vary. As only the yield in the main peak of the  $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$  channel is of interest for purposes of normalisation, it was not deemed necessary to examine the possibility of other low mass backgrounds in the channel.

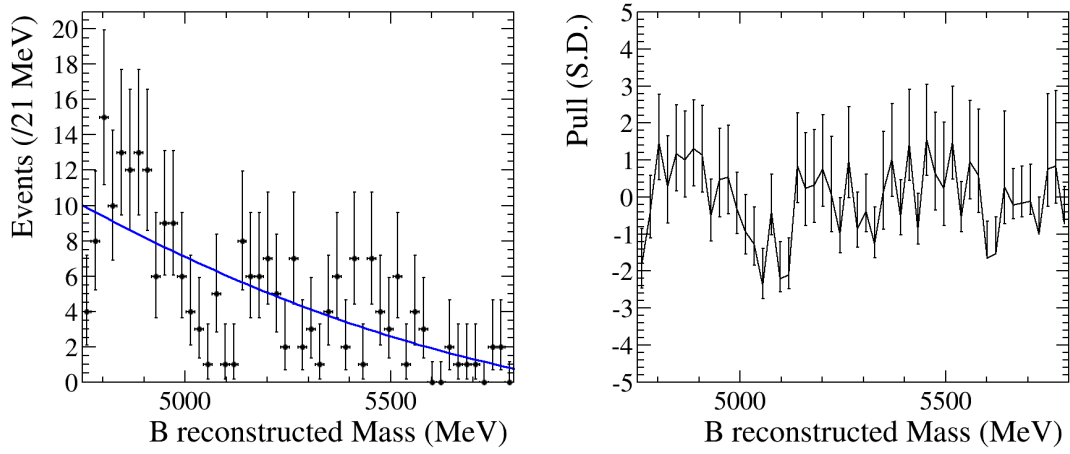


FIGURE 5.23: Mass fit and pull plot of the combinatorial background in the  $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$  channel.

## 5.7 Data Fits and Yields

The full data PDFs are created by combining the individual signal and background fits detailed in section 5.6, and are then used to extract the yields for each signal channel.

### 5.7.1 $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$

The full PDFs for the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data distribution, and the two  $D_s^{\pm}$  distributions, are shown in figures 5.24 and 5.25, where the final fit is a simultaneous unbinned maximum likelihood fit to all three particles. Fitting to all three particles simultaneously gives the fit more information to separate the signal and background events, by considering where the event lies in terms of  $B_s^0$  mass, and also both  $D_s^{\pm}$  masses. Therefore every event is represented by a single point in the  $B_s^0$  distribution, and a single point in both of the  $D_s^{\pm}$  distributions. A number of parameters in the full data fits are fixed to values taken from the individual simulation fits; these are detailed in the relevant sections. Some of the parameters are found by performing the fit to only the  $B_s^0$  mass distribution, or only the  $D_s^{\pm}$  distributions; the parameters taken from these partial fits are then fixed in the full simultaneous fit to the three distributions, to aid the stability of the fit. The yields, and a number of other parameters, are allowed to float. A detailed



break down of the parameters in the final data fit and whether they are fixed (and if so, how) or floated can be seen in appendix E. The yields for all three signal channels and the two backgrounds, along with the total signal yield, are shown in table 5.10. The errors on the total signal yield are determined by considering the covariances of the signal yields and are therefore lower than simply combining the errors in quadrature; in regions of the mass distribution where peaks overlap, the fit can determine that an event is part of a signal decay with a small error, but determining the exact identity of that decay is more difficult.

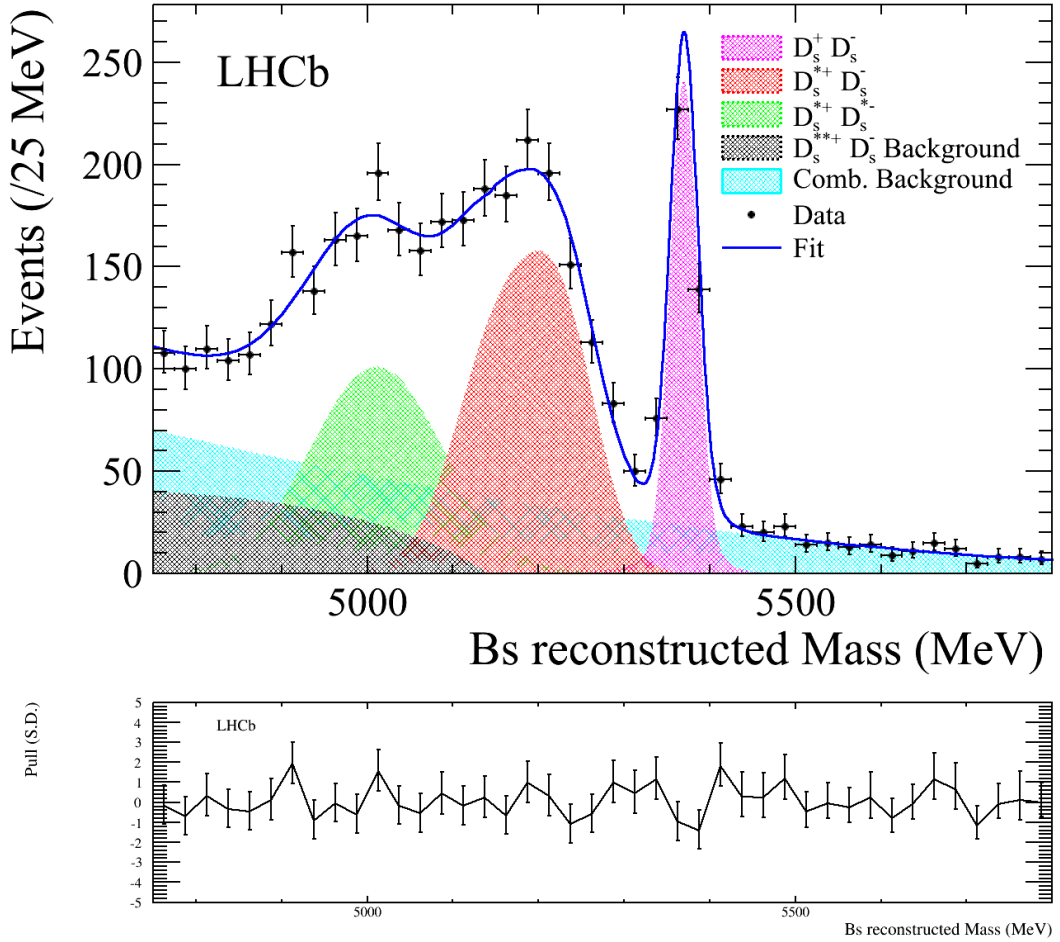


FIGURE 5.24: Mass fit for the full  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data set, showing the full  $B_s^0$  mass distribution with the pull plot underneath.

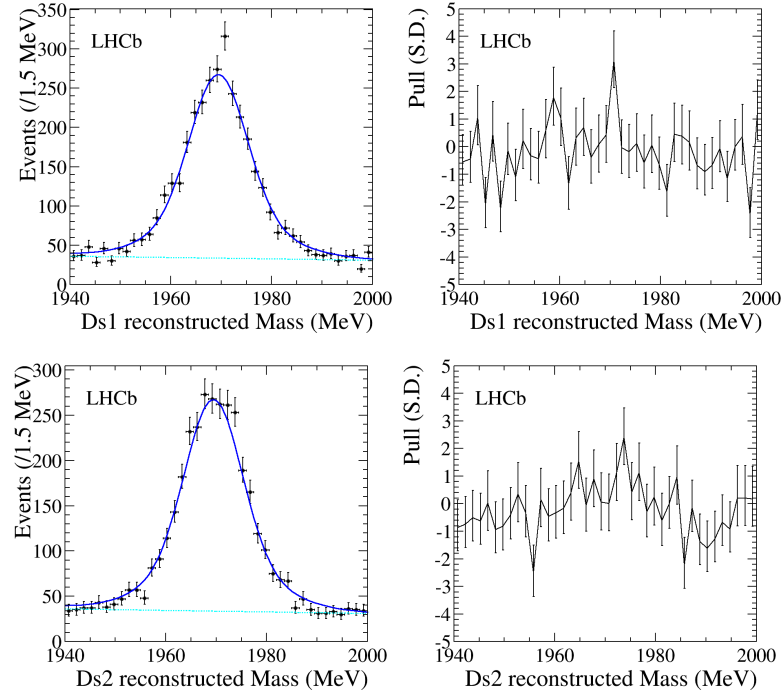


FIGURE 5.25: Mass fits for the full  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data set, showing the two  $D_s^\pm$  fits along with their pull plots. Ds1 and Ds2 are arbitrarily assigned labels for the two  $D_s^\pm$  candidates in each event. The cyan line in each case represents the combinatorial background component of the distribution.

Decay Mode	Yield
$B_s^0 \rightarrow D_s^+ D_s^-$	$412 \pm 23$
$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$1032 \pm 39$
$B_s^0 \rightarrow D_s^{\pm} D_s^{*\mp}$	$786 \pm 48$
Combinatorial Background	$1342 \pm 47$
$B_s^0 \rightarrow D_s(2460)^\pm D_s^\mp$	$432 \pm 42$
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	$2230 \pm 63$

TABLE 5.10: Yields in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data fit.

Due to the relative complexity of the fit, and the large number of parameters involved it was decided to investigate the stability of the fit and look for potential biases, by producing toy Monte Carlo samples from the data fit. The samples were produced with the same fixed parameters and same floated parameters as in the final data fit, with

the yields allowed to vary around their central values. The results of the toy studies are shown in detail in appendix B. All signal yields and their associated uncertainties were reproduced by the toy samples, and the  $-\log(\mathcal{L})$  of the fit to data was also well reproduced. No evidence of underlying bias was seen in the pull plots for each yield, where all mean and sigma values were consistent with 0 and 1 respectively.

### 5.7.2 $B^0 \rightarrow D_s^\pm D^\mp$

The full fit to the  $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$  data distribution is shown in figures 5.26 and 5.27. The parameters of this fit are again found by performing partial fits to either the  $B^0$  distribution, or the  $D_{(s)}^\pm$  distributions. The parameters are then set in the full simultaneous fits, with the yields and a few other parameters allowed to vary. The final data yields are presented in table 5.11. The relatively very small yield from the  $B^0 \rightarrow D_s^{\pm*} D^{\mp*}$  channel is because the vast majority of these decays are reconstructed with a mass lower than the 5050 MeV mass cut imposed on the  $B^0$  candidates, and only a small contribution from the tail of this distribution remains, as seen in (a) in figure 5.26. Only the main  $B^0 \rightarrow D_s^\pm D^\mp$  peak is used for normalisation purposes.

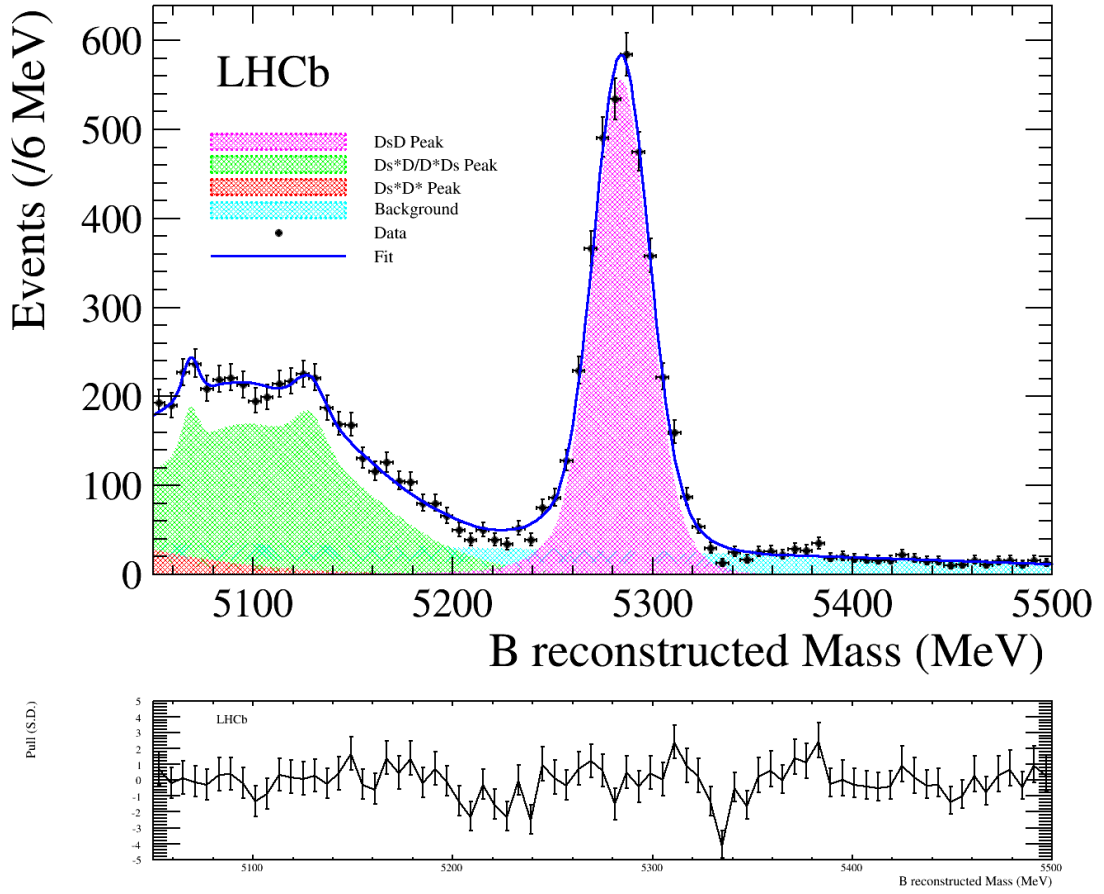
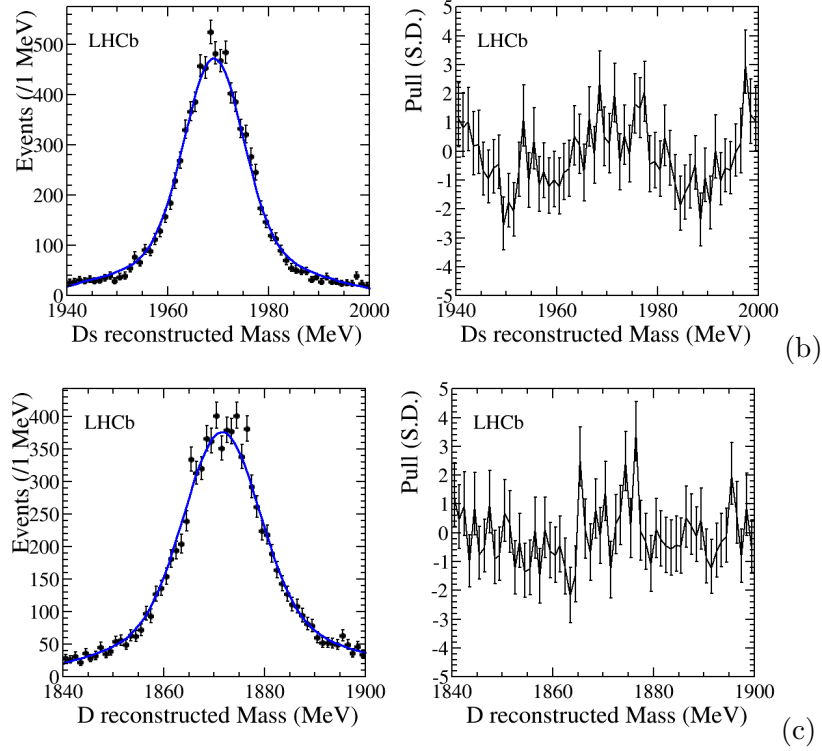


FIGURE 5.26: Mass fits for the full  $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$  data set, showing the full  $B^0$  fit with the pull plot beneath.

Decay Mode	Yield
$B^0 \rightarrow D_s^\pm D^\mp$	$3615 \pm 64$
$B^0 \rightarrow D_s^{\pm*} D^\mp / B^0 \rightarrow D_s^\pm D^{\mp*}$	$3555 \pm 110$
$B^0 \rightarrow D_s^{\pm*} D^{\mp*}$	$195 \pm 86$
Combinatorial Background	$1643 \pm 56$

TABLE 5.11: Yields in the  $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$  data fit.FIGURE 5.27: Mass fits for the full  $B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$  data set, showing the  $D_s^\pm$  and  $D^\pm$  fits along with their pull plots.

## 5.8 Systematic Uncertainties

The majority of systematic uncertainties cancel in the ratio. Five of the six final state particles are common to the decay of both the  $B^0$  and  $B_s^0$ , and only the switch from a  $\pi^+$  to a  $K^+$  is likely to play a role. Other contributions arise from the various fit shapes used, the selection efficiencies, and the uncertainty on  $f_s/f_d$ ; these uncertainties are all detailed in the remainder of this chapter.

Besides the switch from a  $\pi^+$  to a  $K^+$  in the final states, the signal and normalisation channels also differ slightly due to the longer lifetime of the  $D^+$  compared to the  $D_s^+$ . This was investigated as a potential source of systematic uncertainty, by applying an upper lifetime cut to the  $D^+$  of 0.002 ns in order to cut the distribution at the same point the  $D_s^+$  lifetime distribution becomes negligible. This cut was found to reduce the yield in the normalisation channel to  $(87.6 \pm 2.3)\%$  of the original value, but the efficiency of this cut relative to the previous selection was found to be  $(87.1 \pm 1.3)\%$ . Therefore any possible systematic arising from this difference in lifetimes was deemed to be negligible and was ignored.

### 5.8.1 $D_s(2460)$ background shape

As the shape for this background is fixed in the final data fit, with just the yield allowed to float, it was necessary to examine the effect that varying the parameters had on the overall signal yields. The important parameters for the Argus shape used are the mass cut off, and a constant governing the rate at which the function increases below the mass cut off. The values of the parameters and their errors, as taken from the fit to simulation, are shown in table 5.12.

Parameter	Value	Error
Mass cut off	5140.22	7.70
Constant	0.886	0.361

TABLE 5.12: Parameters for the Argus fit to MC events.

As the two parameters had a correlation coefficient of 0.17, it was decided to separately vary the parameters within one sigma, and take the average of the largest plus and minus effects as the systematic error. The effects of varying the parameters on the yields of the signal channels are shown in table 5.13.

Decay Mode	Actual Yield	Cut off +	Cut off -	Constant +	Constant -
$B_s^0 \rightarrow D_s^+ D_s^-$	$412 \pm 23$	$412 \pm 23$	$412 \pm 23$	$412 \pm 23$	$413 \pm 23$
$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$1032 \pm 39$	$1020 \pm 38$	$1034 \pm 39$	$1030 \pm 39$	$1029 \pm 38$
$B_s^0 \rightarrow D_s^{*\pm} D_s^{*\mp}$	$786 \pm 48$	$777 \pm 48$	$794 \pm 47$	$848 \pm 45$	$695 \pm 53$
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	$2230 \pm 63$	$2220 \pm 63$	$2241 \pm 63$	$2290 \pm 61$	$2136 \pm 66$

TABLE 5.13: Variations of yields in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data fit, when the Argus parameters are varied by their error from simulation.

The final data plots when the constant is varied can be seen in Appendix D.

It was also deemed necessary to consider potential contributions from partially reconstructed  $B_s^0 \rightarrow D_s(2317) D_s^+$  decays. A small number of MC events were generated in this channel, with the  $D_s(2317)$  forced to decay 100% of the time to  $D_s^+ \pi^0$  (as this is currently the only observed decay of the particle). Figure 5.28 shows the events remaining after the stripping line selection (black) and after the full selection has been applied (red). The selection efficiency of the events is similar to that for the  $D_s(2460)$  background MC sample, but there is no evidence from the plot in figure 5.28 to suggest that the remaining distribution varies greatly from the  $D_s(2460)$  distribution, or that these decays would not be accounted for in the  $D_s(2460)$  distribution already included in the final fit. As the branching fraction is unknown, and owing to a lack of knowledge of the other  $D_s(2317)$  decay modes, the  $D_s(2460)$  background shape is extended to 5180 MeV to assess the potential systematic uncertainty on the yields due to this possible background. The yields after this extension are shown in table 5.14, and the systematic for the  $D_s(2460)$  is extended to encompass this.

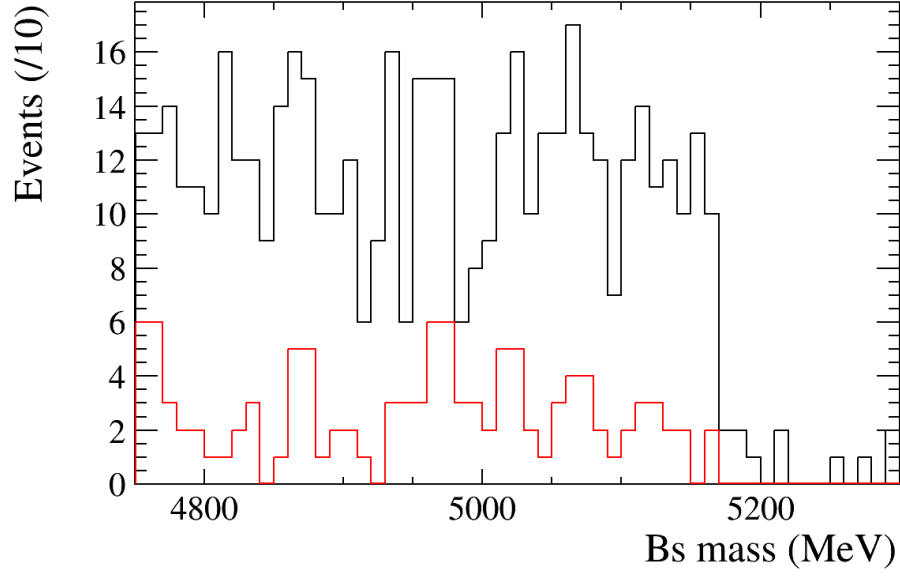


FIGURE 5.28: Remaining  $B_s^0 \rightarrow D_s(2317)D_s^+$  MC signal events after stripping line selection (black) and after the full selection have been applied (red). The  $D_s(2460)$  is conservatively extended to 5180 MeV to assess a possible systematic from these decays.

Decay Mode	Yield
$B_s^0 \rightarrow D_s^+ D_s^-$	$413 \pm 23$
$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$1015 \pm 38$
$B_s^0 \rightarrow D_s^{*\pm} D_s^{*\mp}$	$748 \pm 50$
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	$2176 \pm 64$

TABLE 5.14: Variations of yields in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data fit, when the Argus cut off is extended to 5180 MeV to account for potential background from decays of the form  $B_s^0 \rightarrow D_s(2317)D_s^+$ .

### 5.8.2 Combinatorial background shape

To ensure that the signal yields are not biased by the choice of shape for the combinatorial background, the fit type is switched from a Chebyshev polynomial to an exponential in the final fit, where both shapes are allowed to float in the fit. The two sets of yields are shown in table 5.15.



Decay Mode	Chebyshev Yield	Exponential Yield
$B_s^0 \rightarrow D_s^+ D_s^-$	$412 \pm 23$	$418 \pm 23$
$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$1032 \pm 39$	$1044 \pm 38$
$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$	$786 \pm 48$	$801 \pm 48$
Combinatorial Background	$1342 \pm 47$	$1324 \pm 46$
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	$2230 \pm 63$	$2263 \pm 63$

TABLE 5.15: Effect of using an exponential combinatorial background fit on the signal channel yields.

It is clear that the effect of switching to an exponential fit for the combinatorial background is small. The difference between the yields is used to assign a small systematic in each case.

### 5.8.3 Mass PDFs

In order to account for the uncertainty on the parameters in the description of the  $B_s^0$  mass distribution, and also to propagate forward the uncertainties on the individual parameters that are fixed, to values from either MC or the partial data fit, in the final mass fit, a systematic uncertainty is applied based on the variation of these parameters. The first parameter in the PDF has its value increased by one standard deviation, whilst all other parameters are kept at the values used in the final fit. The effect of this variation on the individual channel yields, and also on the overall inclusive yield, is recorded, and then the procedure is repeated but with the parameter set to one standard deviation less than its value in the final fit. This is repeated for every parameter in the fit individually. In each case, the average of the largest deviation observed above and below the yield from the final fit is taken to be the systematic error for that channel. A summary of the parameters that have the largest effects on each yield, along with the size of the effect, is shown in table 5.16.

Decay Mode	Largest Increase Parameter	+Yield	Largest Decrease Parameter	-Yield	Average
$B_s^0 \rightarrow D_s^+ D_s^-$	$D_s^+$ Crystal Ball width	6	$D_s^+$ Crystal Ball width	6	$\pm 6$
$B_s^0 \rightarrow D_s^{*+} D_s^-$	$D_s^{*+} D_s^-$ Gaussian width	28	$D_s^{*+} D_s^-$ Gaussian width	30	$\pm 29$
$B_s^0 \rightarrow D_s^{*+} D_s^{*-}$	$D_s^{*+} D_s^{*-}$ Gaussian width	66	$D_s^{*+} D_s^{*-}$ Gaussian width	65	$\pm 66$
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$	$D_s^{*+} D_s^{*-}$ Gaussian width	36	$D_s^{*+} D_s^{*-}$ Gaussian width	38	$\pm 37$

TABLE 5.16: Largest and smallest deviations from the final yields when varying every parameter in the PDF by plus and minus one standard deviation.

#### 5.8.4 $f_s/f_d$ and $D$ meson branching ratio

The uncertainties on both the value of  $f_s/f_d$  and the ratio of the  $D^+$  to  $D_s^+$  branching fractions both contribute to the overall systematic uncertainty in the analysis. The values, along with uncertainties, are shown in table 5.17.

Parameter	Value	Error
$f_s/f_d$	0.259	0.015
$\frac{\mathcal{B}(D^+ \rightarrow K^- \pi^+ \pi^+)}{\mathcal{B}(D_s^+ \rightarrow K^- K^+ \pi^+)}$	1.684	0.056

TABLE 5.17: Uncertainties on  $f_s/f_d$  and the ratio of the  $D^+$  to  $D_s^+$  branching fractions.

As shown in [108], the uncertainty on the value of  $f_s/f_d$  is dependent on the uncertainty of the ratio of the  $D^+$  to  $D_s^+$  branching fractions. As the inverse is used in the branching ratio calculation, the two errors cancel. Therefore the contribution from the ratio of the  $D^+$  to  $D_s^+$  branching fractions is removed, and is also removed in quadrature from the uncertainty on  $f_s/f_d$ ; this means the value taken for  $f_s/f_d = 0.259 \pm 0.012$ .

#### 5.8.5 Simulation size

There is an uncertainty on the selection efficiencies calculated from simulation, owing to the finite size of the simulated samples. The uncertainties are calculated by considering the square root of the sample size remaining after all cuts, relative to the size itself. This value is then added in quadrature to the equivalent value for the normalisation channel for each of the signal channels. The values found for the systematic uncertainties here are shown in the summary table in 5.19.

### 5.8.6 Data sample size

As the relative PID and trigger efficiencies are calculated from the data, both of these values have a systematic uncertainty that is dependent on the size of the data sample. These values are added in quadrature, and listed in the row ‘Sample size’ in table 5.19.

### 5.8.7 PIDCalib binning

There is a systematic uncertainty applied based on the binning scheme used in the PIDcalib method for calculating the efficiency of the PID cuts used in the selection, as detailed in 5.5.1. This is estimated by considering the efficiency given by the method when using a different binning scheme, and is found to be 3.0%.

### 5.8.8 Trigger efficiency

A small systematic uncertainty is included to account for potential differences arising from the trigger requirement between MC and data. This is calculated by considering the maximum difference between the ratio of L0HadronTOS  $B_s^0$  and  $B^0$  candidates from data to the three signal MC samples. The fraction of  $B_s^0$  candidates in the final signal data sample that are L0HadronTOS is calculated relative to the fraction of  $B^0$  candidates in the final normalisation data sample that are L0HadronTOS. This ratio is also calculated for the three individual signal channels in MC. The three numbers from MC are compared to the single number from data, and the maximum difference is taken to be the systematic uncertainty arising from the trigger efficiency. Using this method, the uncertainty is calculated to be 1.5%.

### 5.8.9 Further tests

This section details further tests that were performed, but that proved to have no systematic uncertainty associated with them.

*Branching ratio dependence on BDT cut:* the branching ratios of the four channels were remeasured when the BDT cut was varied from  $-0.2$  to  $0.0$ , by considering the yield and selection efficiency of the signal and normalisation channels at each BDT cut. The

branching ratios as a function of BDT cut are shown in figure 5.29; the error bars in this figure are systematic only, and the gradients are consistent with zero within one standard deviation in each case. Thus it was deemed unnecessary to include a systematic uncertainty resulting from the BDT cut choice.

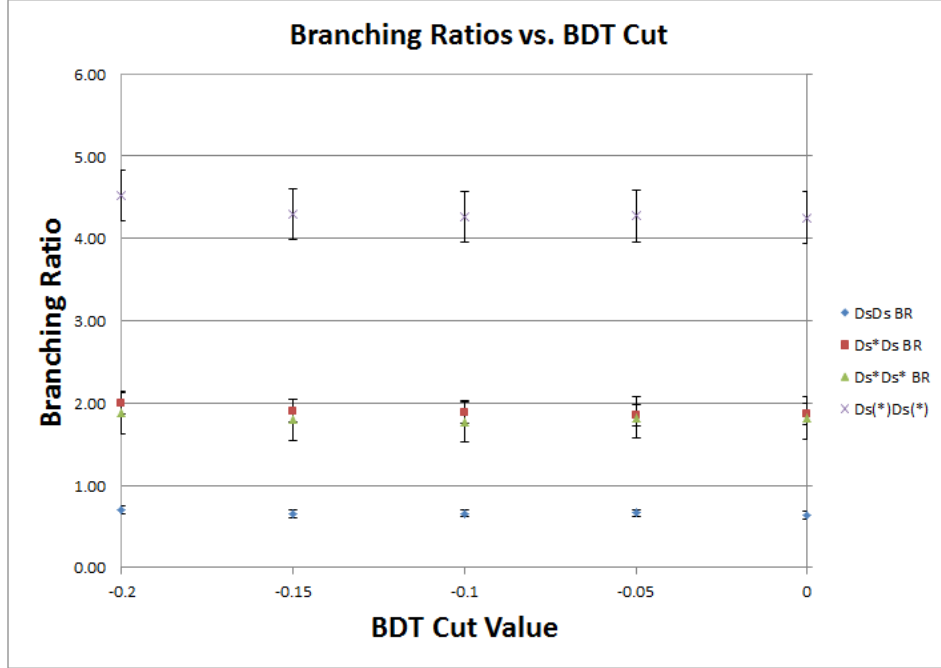


FIGURE 5.29: The variation of the branching ratio of the four channels with the BDT cut used in the selection. The error bars are systematic only. The lack of dependence on the BDT cut value is clear for each of the four channels.

*Multiple candidate events:* A multiple candidate event is one where two or more  $B_s^0$  candidates pass the full selection criteria in the same event. There is a very small probability of both  $b$  quarks in the event forming  $B_s^0$  mesons and decaying to a  $D_s^+$  and  $D_s^-$ . Instead, the vast majority of multiple candidate events occur when a  $K^\pm$  or  $\pi^\pm$  is switched for another  $K^\pm$  or  $\pi^\pm$  candidate, and the full event still satisfies all selection criteria. The presence of multiple candidate events was examined in both the signal and normalisation channels, in both the data and signal MC. The percentage of events containing multiple candidates in each case is shown in table 5.18.

Channel	Multiple candidate events (%)
$B_s^0 \rightarrow D_s^+ D_s^-$ MC	1.40
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ data	3.38
$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ data with tighter BDT	2.38
$B^0 \rightarrow D_s^\pm D^\mp$ MC	0.78
$B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$ data	2.67
$B^0 \rightarrow D_s^{\pm(*)} D^{\mp(*)}$ data with tighter BDT	1.37

TABLE 5.18: Showing the percentage of events containing multiple candidates in the signal and normalisation channels, both in data and MC. The tighter BDT entries refer to the results when using a BDT cut of 0.1, which ensures the background events remaining in the data distributions are minimal.

The important results are the ratios between the normalisation and signal channels; in MC, this is seen to be 0.557. In data, with the actual BDT cut used, this is found to be 0.790. However, when using a tighter BDT cut of 0.1 to remove the combinatorial background, this ratio is found to drop to 0.576, which is compatible with the value found in MC. Therefore this ratio is well modelled by the MC, and it is determined that any possible systematic effect resulting from multiple candidates is negligible.

#### 5.8.10 Total systematic uncertainties

The relative contributions to the systematic uncertainty, and the overall totals after combining the individual errors in quadrature, are shown in table 5.19.

Systematic	$B_s^0 \rightarrow D_s^+ D_s^-$	$B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$	$B_s^0 \rightarrow D_s^{*\pm} D_s^{*\mp}$	$B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$
Simulation size	1.4%	1.9%	2.1%	1.9%
$D_s(2460)$	0.0%	1.6%	9.7%	3.5%
Comb. Background	1.5%	1.2%	1.9%	1.5%
$f_s/f_d$	4.7%	4.7%	4.7%	4.7%
Sample size	0.3%	0.3%	0.3%	0.3%
Mass PDF	1.5%	2.8%	8.4%	1.7%
Trigger Efficiency	1.5%	1.5%	1.5%	1.5%
PIDCalib binning	3.0%	3.0%	3.0%	3.0%
Total	6.3%	7.0%	14.4%	7.4%

TABLE 5.19: Sources of systematic uncertainty in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  data fit.

## 5.9 Results

The relative branching fraction of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decays to  $B^0 \rightarrow D_s^\pm D^\mp$  is calculated using equation 5.8:

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s^{+(*)} D_s^{-(*)})}{\mathcal{B}(B^0 \rightarrow D_s^\pm D^\mp)} = \frac{f_d}{f_s} \cdot \epsilon_{rel}^{B^0/B_s^0} \cdot \frac{\mathcal{B}(D^+ \rightarrow K^- \pi^+ \pi^+)}{\mathcal{B}(D_s^+ \rightarrow K^- K^+ \pi^+)} \cdot \frac{N_{B_s^0}}{N_{B^0}}. \quad (5.8)$$

The value for the ratio of fragmentation fractions is taken as  $0.259 \pm 0.012$  [108]. The  $D$  meson branching fractions are  $\mathcal{B}(D^+ \rightarrow K^- \pi^+ \pi^+) = (9.13 \pm 0.19)\%$  [106] for the  $D^\pm$ , and  $\mathcal{B}(D_s^+ \rightarrow K^- K^+ \pi^+) = (5.42 \pm 0.14)\%$  [108] for the  $D_s^\pm$ . The results for the three individual channels, along with the result for the inclusive decay, are shown below:

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s^+ D_s^-)}{\mathcal{B}(B^0 \rightarrow D_s^\pm D^\mp)} = 0.71 \pm 0.04 \text{ (stat)} \pm 0.05 \text{ (syst)}$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s^{*\pm} D_s^{\mp})}{\mathcal{B}(B^0 \rightarrow D_s^{\pm} D^{\mp})} = 1.89 \pm 0.08 \text{ (stat)} \pm 0.13 \text{ (syst)}$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s^{*\pm} D_s^{*\mp})}{\mathcal{B}(B^0 \rightarrow D_s^{\pm} D^{\mp})} = 1.77 \pm 0.11 \text{ (stat)} \pm 0.26 \text{ (syst)}$$

$$\frac{\mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-})}{\mathcal{B}(B^0 \rightarrow D_s^{\pm} D^{\mp})} = 4.26 \pm 0.14 \text{ (stat)} \pm 0.31 \text{ (syst)}.$$

Using the current world best measurement of the  $B^0 \rightarrow D_s^{\pm} D^{\mp}$  branching fraction of  $(7.2 \pm 0.8) \cdot 10^{-3}$ , the branching fractions of the various decay modes are shown below, where the uncertainties are statistical, systematic and due to the normalisation:

$$\mathcal{B}(B_s^0 \rightarrow D_s^+ D_s^-) = (5.10 \pm 0.29 \pm 0.39 \pm 0.57) \cdot 10^{-3}$$

$$\mathcal{B}(B_s^0 \rightarrow D_s^{*\pm} D_s^{\mp}) = (1.36 \pm 0.06 \pm 0.10 \pm 0.15)\%$$

$$\mathcal{B}(B_s^0 \rightarrow D_s^{*\pm} D_s^{*\mp}) = (1.28 \pm 0.08 \pm 0.18 \pm 0.14)\%$$

$$\mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.07 \pm 0.10 \pm 0.23 \pm 0.34)\%.$$

## 5.10 Conclusions

The branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  has been measured using the 2011 LHCb dataset, consisting of approximately  $1.1 \text{ fb}^{-1}$  of proton-proton collisions. The branching

fraction was found to be

$$\mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.07 \pm 0.10 \pm 0.23 \pm 0.34)\%$$

where the first uncertainty is statistical, the second is systematic, and the third is a result of the normalisation to the channel  $B^0 \rightarrow D_s^\pm D^\mp$ . The dominant systematic effects come from the uncertainty in the value of  $f_s/f_d$  in the  $B_s^0 \rightarrow D_s^+ D_s^-$ ,  $B_s^0 \rightarrow D_s^{*\pm} D_s^\mp$ , and  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  channels, and from the limited understanding of the shape of the  $B_s^0 \rightarrow D_s(2460)^\pm D_s^\mp$  background in the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  channel. It should also be noted that, for three of the four results, the dominant uncertainty arises from the normalisation channel, and therefore the results can be improved with an improved measurement of the  $B^0 \rightarrow D_s^+ D^-$  branching fraction. In addition, this is the first time that the background arising from the  $B_s^0 \rightarrow D_s(2460)^\pm D_s^\mp$  decay has been accounted for in the measurement of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decay. Comparing this new LHCb result with those from CDF and Belle, and also the value from the PDG [106]:

$$B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.38 \pm 0.25 \pm 0.30 \pm 0.56)\% \quad (\text{CDF})$$

$$B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (4.32_{-0.39-1.03}^{+0.42+1.04})\% \quad (\text{Belle})$$

$$B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (4.5 \pm 1.4)\% \quad (\text{PDG})$$

$$B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.07 \pm 0.10 \pm 0.23 \pm 0.34)\% \quad (\text{LHCb 2011 dataset}).$$



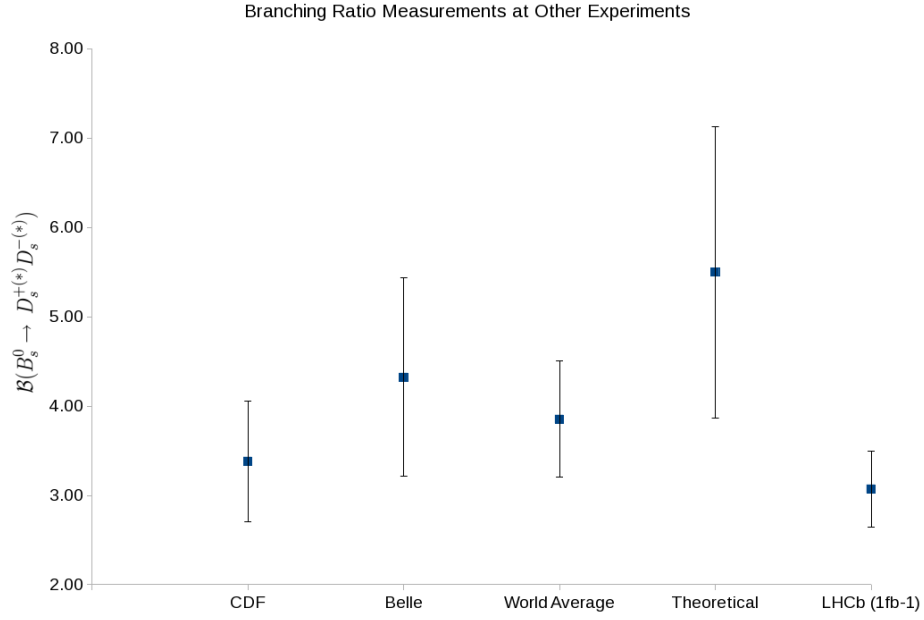


FIGURE 5.30: The previous measurements of the branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ , and the theoretical prediction, along with the LHCb  $1 \text{ fb}^{-1}$  result. This result is consistent with all previous measurements and the theoretical prediction, and is the most precise result to date. The average result displayed here is an average of the two previous results, rather than the PDG value.

Figure 5.30 reproduces the plot in figure 5.5, with the new LHCb result added. This result is found to be consistent with, and more precise than, these two most recent previous measurements and also the average of the two previous values.

The results for the individual branching fractions can also be compared to the results from Belle. The Belle results for the three individual branching fractions, as taken from [92], were:

$$\mathcal{B}(B_s^0 \rightarrow D_s^+ D_s^-) = (5.80_{-0.9}^{+1.1} \pm 1.3) \cdot 10^{-3}$$

$$\mathcal{B}(B_s^0 \rightarrow D_s^{*\pm} D_s^\mp) = (1.76_{-0.22}^{+0.23} \pm 0.40)\%$$

$$\mathcal{B}(B_s^0 \rightarrow D_s^{*\pm} D_s^{*\mp}) = (1.98_{-0.31-0.50}^{+0.33+0.52})\%.$$

The branching fractions of the three individual channels presented in this chapter are therefore consistent with, and more precise than, the previous individual measurements from Belle.

Using equation 5.1, the  $B(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-})$  result can be interpreted in terms of  $\Delta\Gamma_s/\Gamma_s$ . Substituting in the value for the branching fraction of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ ,  $\Delta\Gamma_s/\Gamma_s$  is found to be:

$$\frac{\Delta\Gamma_s}{\Gamma_s} = 0.0633 \pm 0.0087.$$

Comparing to the current world average from the PDG, which is calculated from two separate analyses of  $B_s^0 \rightarrow J/\psi \phi$  decays from LHCb [18] and D0 [109]:

$$\frac{\Delta\Gamma_s}{\Gamma_s} = 0.138 \pm 0.012.$$

The result from  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decays is found to be inconsistent with the world average value, at approximately half the size. This suggests that the theoretical update in [90] is indeed correct;  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  do not saturate the  $CP$ -Even modes, and it appears that 3-body final states will contribute a similar level to  $\Delta\Gamma_s/\Gamma_s$ . As suggested in [92] and [96], the likelihood is that the  $CP$ -Odd contamination is only around 6%, and the main issue is that 3-body  $B$  decays contribute a sizeable portion of the  $CP$ -Even decays.

In order to fully determine this, it will be necessary to make future measurements of the branching fractions of these other 3-body modes. For modes that are not  $CP$  eigenstates, angular analyses will be needed in order to disentangle the different  $CP$  components. Taken with these other measurements, it will be possible to determine a value for  $\Delta\Gamma_s/\Gamma_s$  that can be compared to the values determined from clean  $B_s^0 \rightarrow J/\psi \phi$  measurements, in order to fully assess the compatibility with Standard Model predictions for the  $B_s^0$  mixing parameters.

## Chapter 6

# Summary

This thesis has presented studies of the effect of radiation damage on the performance of the Vertex Locator detector of the LHCb experiment, specifically on the current-voltage (I-V) characteristics of the silicon sensors used in the detector. The increase in the bulk current drawn by sensors at full depletion with the luminosity delivered to the detector is found to be well described theoretically within experimental uncertainties, after annealing of the silicon has been properly assessed. It is seen that the detector will be able to function properly in spite of the current level of radiation damage, and the predicted future radiation damage it will endure, for the remainder of its lifespan.

In order to properly assess the precision of the predictions, it will be necessary to separate the sensor leakage current into the bulk and surface components, as the prediction is valid only for the bulk current. It will also be necessary to better measure the temperature of the silicon during operation.

This thesis has also presented a measurement of the branching fraction of the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decay, using approximately  $1 \text{ fb}^{-1}$  of proton-proton collisions recorded by LHCb during 2011. The branching fraction is measured relative to the decay  $B^0 \rightarrow D_s^\pm D^\mp$ , and is found to be:

$$\mathcal{B}(B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}) = (3.07 \pm 0.10(stat.) \pm 0.23(sys.) \pm 0.34(norm.))\%.$$

This measurement is found to be consistent with, and more precise than, all previous measurements of the branching fraction, and the current world average. It is also consistent with theoretical predictions. Comparison with direct determinations of the value of  $\Delta\Gamma_s/\Gamma_s$  suggest that updated theoretical suggestions, that the branching ratio of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  cannot be used to approximate  $\Delta\Gamma_s/\Gamma_s$ , are indeed correct. Future measurements of other final state  $B_s^0$  decays are needed before this measurement can be used to accurately determine the value of  $\Delta\Gamma_s/\Gamma_s$ .

The three assumptions on which the original theory was based, as shown in section 5.2, all require review in light of more recent theoretical updates and experimental results. Calculations suggest that the assumption that  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decays are predominantly  $CP$ -Even is approximately true, with a  $CP$ -Odd component of around 6%. However, an angular analysis of the  $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$  decays in particular is required in order to confirm this. It is more likely that the main reason for the difference between the value of  $\Delta\Gamma_s/\Gamma_s$  inferred from  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  decays and the value measured from  $B_s^0 \rightarrow J/\psi \phi$  decays results from other 3-body  $CP$ -Even decays that were neglected in the original theory, but actually contribute a sizeable amount to the total width difference between  $CP$ -Even and  $CP$ -Odd  $B_s^0$  decays. Angular analyses of these 3-body  $B_s^0$  decays will be needed in order to measure the  $CP$ -Even components, which can then be used with the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  branching fraction to give a measurement of  $\Delta\Gamma_s/\Gamma_s$  that is comparable with the  $B_s^0 \rightarrow J/\psi \phi$  result.

## Appendix A

### BDT Variables MC vs Data

The BDT is trained using simulated events for the signal sample. This appendix demonstrates the agreement between the MC and data distributions for the variables used in the training and testing of the BDT. The MC distributions are taken from truth matched signal  $B_s^0 \rightarrow D_s^+ D_s^-$  decays, and the data distributions are taken from the events that remain after the full selection, including a BDT cut at -0.1, has been applied. Therefore the data distributions shown include combinatorial background events as well as signal events. In all instances, the data distribution is shown in red, and the MC distribution is shown in black. The only clear discrepancy that is seen is in the  $B_s^0$  impact parameter  $\chi^2$  to its primary vertex. This appears to be due to the combinatorial background having a very different distribution in this variable compared with signal. This is apparent when considering the distribution with a much tighter BDT cut, and therefore virtually no combinatorial background, as demonstrated in figure A.2. Tightening the BDT cut also removes the smaller discrepancies seen in the  $D_s^\pm$  lifetime, as shown in A.6. In variables where the combinatorial background has a very different distribution to the signal, the comparison between signal MC and this data sample is not sufficient to demonstrate the agreement between signal distribution in data and MC. Using a tight BDT cut to remove the combinatorial background allows for a more accurate comparison between the signal distributions, as the background contamination in the data samples is almost all removed.

All variables shown and discussed in this appendix have been defined in brief in section 5.3, and are also described in full detail in appendix F.

Also shown is the BDT response in data and signal MC, in figure A.5.

Also shown are the linear correlation matrices for both signal and background events, in figures A.7 and A.8.

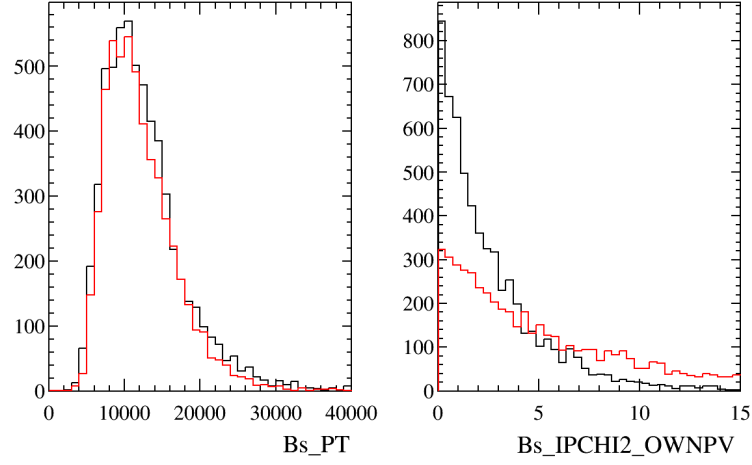


FIGURE A.1: Plots showing the agreement between MC and data distributions for the variables used in the BDT from the  $B_s^0$ . Data is shown in red, and MC is in black.

The normalisation is arbitrary.

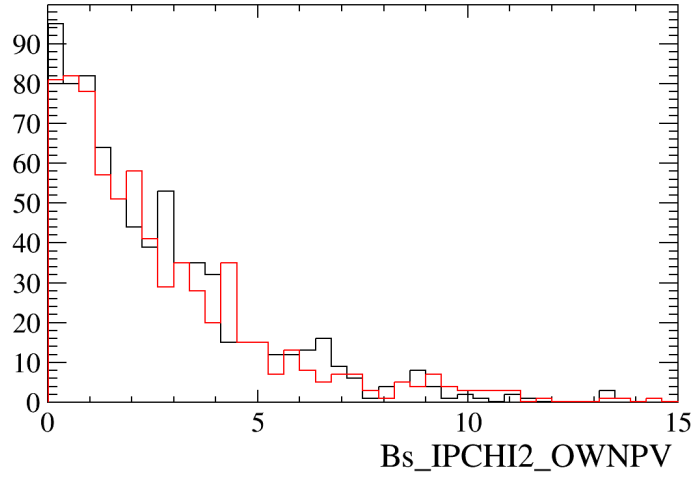


FIGURE A.2: Plot showing the agreement between signal MC and data signal for the  $B_s^0$  impact parameter  $\chi^2$  to its own primary vertex, after tightening the BDT cut to remove the combinatorial background component of the full data distribution. Data is shown in red, and MC is in black. The normalisation is arbitrary.

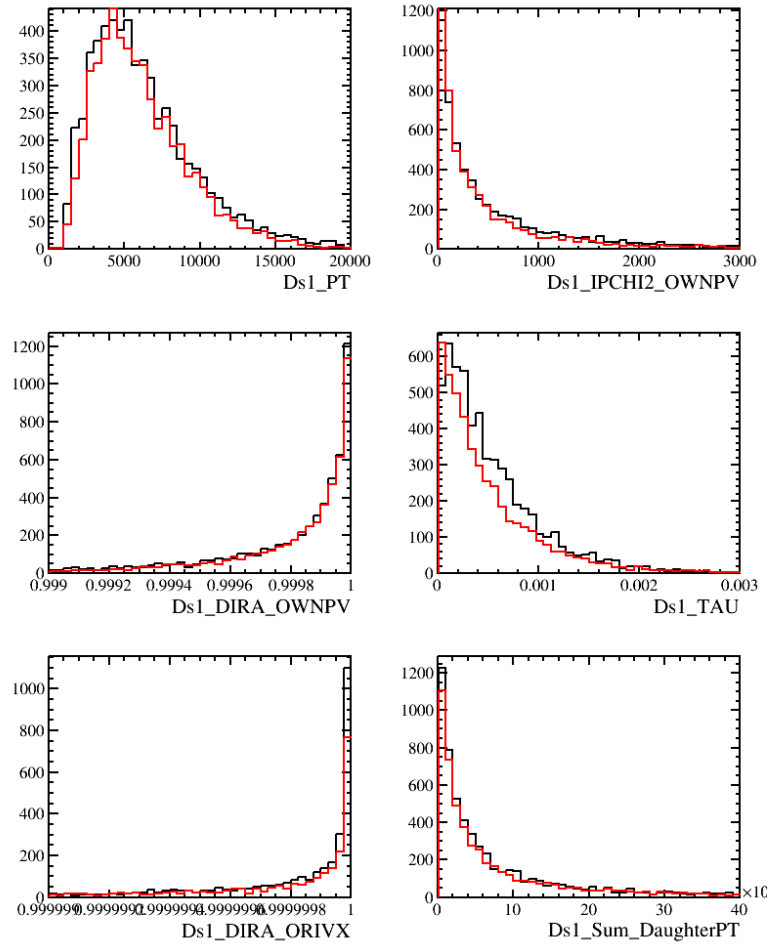


FIGURE A.3: Plots showing the agreement between MC and data distributions for the variables used in the BDT from the first  $D_s^\pm$ . Data is shown in red, and MC is in black. The normalisation is arbitrary.

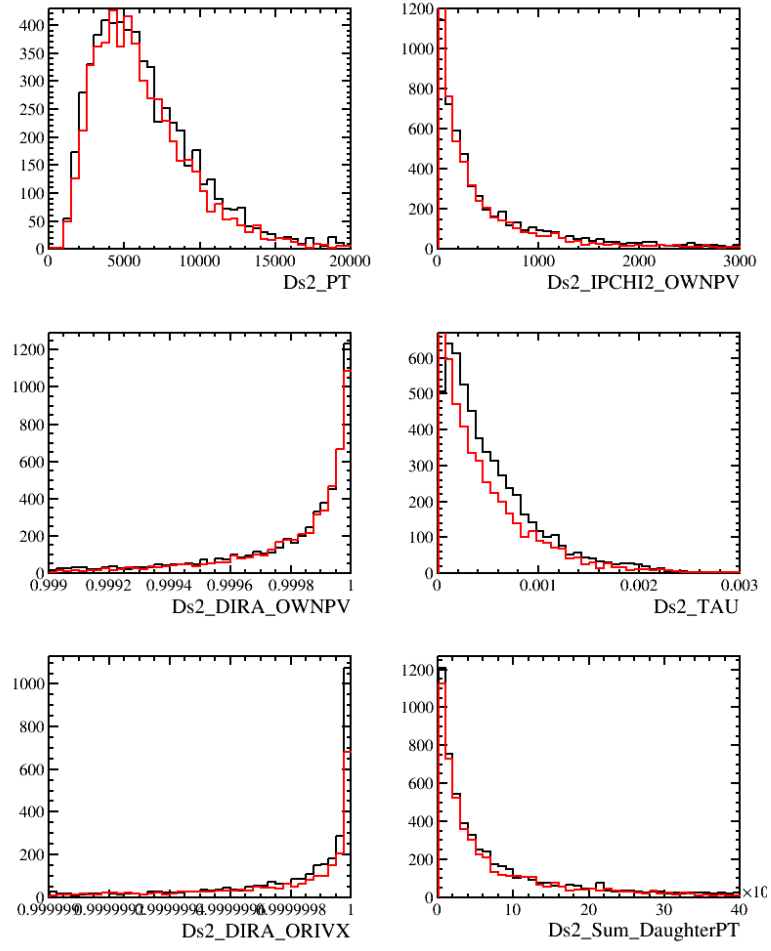


FIGURE A.4: Plots showing the agreement between MC and data distributions for the variables used in the BDT from the second  $D_s^\pm$ . Data is shown in red, and MC is in black. The normalisation is arbitrary.



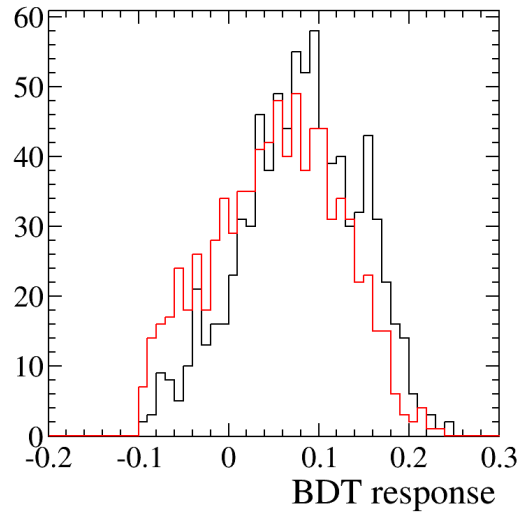


FIGURE A.5: Plots showing the data and MC distributions for the BDT response value. The selection is tightened by using a harsher PID cut in this case, as it could not be done by increasing the BDT cut. The BDT cut is the best way to remove background however, so this is not a fully background removed data set. Data is shown in red, and MC is in black. The normalisation is arbitrary.

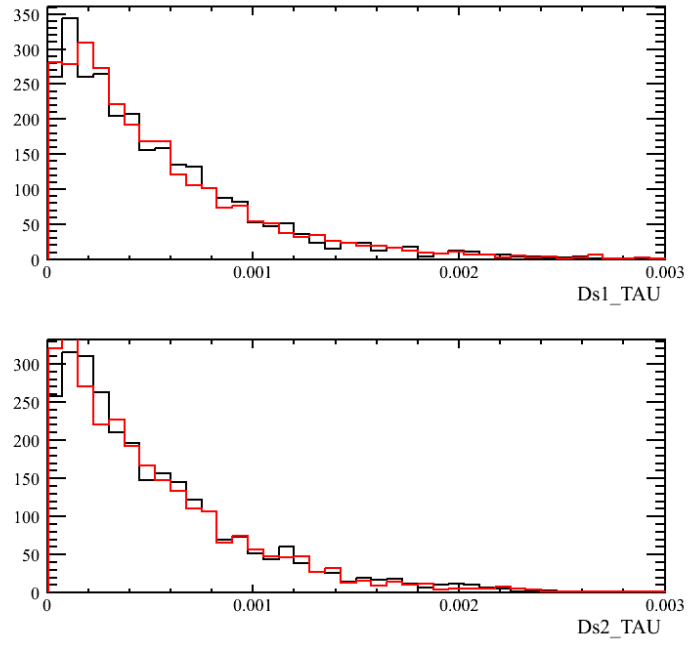


FIGURE A.6: Plots showing the agreement between signal MC and data signal for the lifetime distributions of the  $D_s^\pm$ , after tightening the BDT cut to remove the combinatorial background component of the full data distribution. Data is shown in red, and MC is in black. The normalisation is arbitrary.

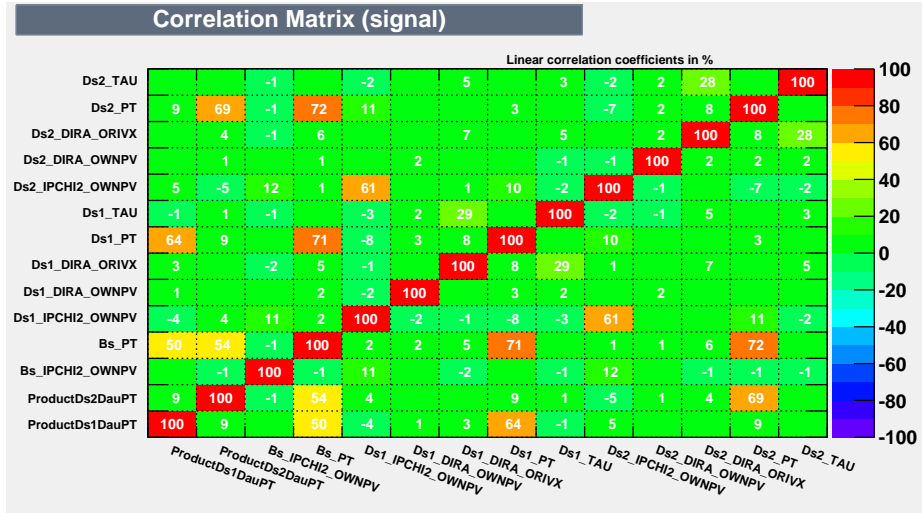


FIGURE A.7: The correlation matrix for the BDT input variables for the signal events.

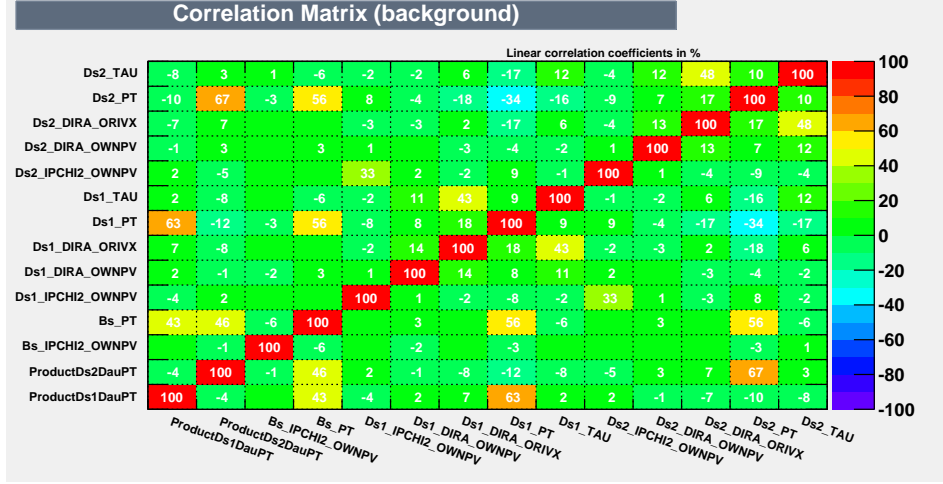


FIGURE A.8: The correlation matrix for the BDT input variables for the background events.

## Appendix B

### Toy MC Studies

The final fit to the data for the signal channel is a fairly complex fit with a number of parameters. For this reason, it was decided to produce toy Monte Carlo samples to analyse the stability of the fit, and check for potential biases. The toys were generated using the same fixed parameters and same floated parameters as in the data fit, with the yields allowed to vary around the central values as taken from the data fit. For each of the yields, the yield and its error as measured by the toys are plotted in this appendix, along with the pull plot for each case. The yields and their associated errors can be compared with those measured in the data fit, shown in Fig. 5.10. Also shown is the  $-\log(\mathcal{L})$  distribution from the toys, with the actual value found from the fit to data, which was 27680, shown as a red arrow on the plot. The pull plots display the number of standard deviations between the point and the mean value.

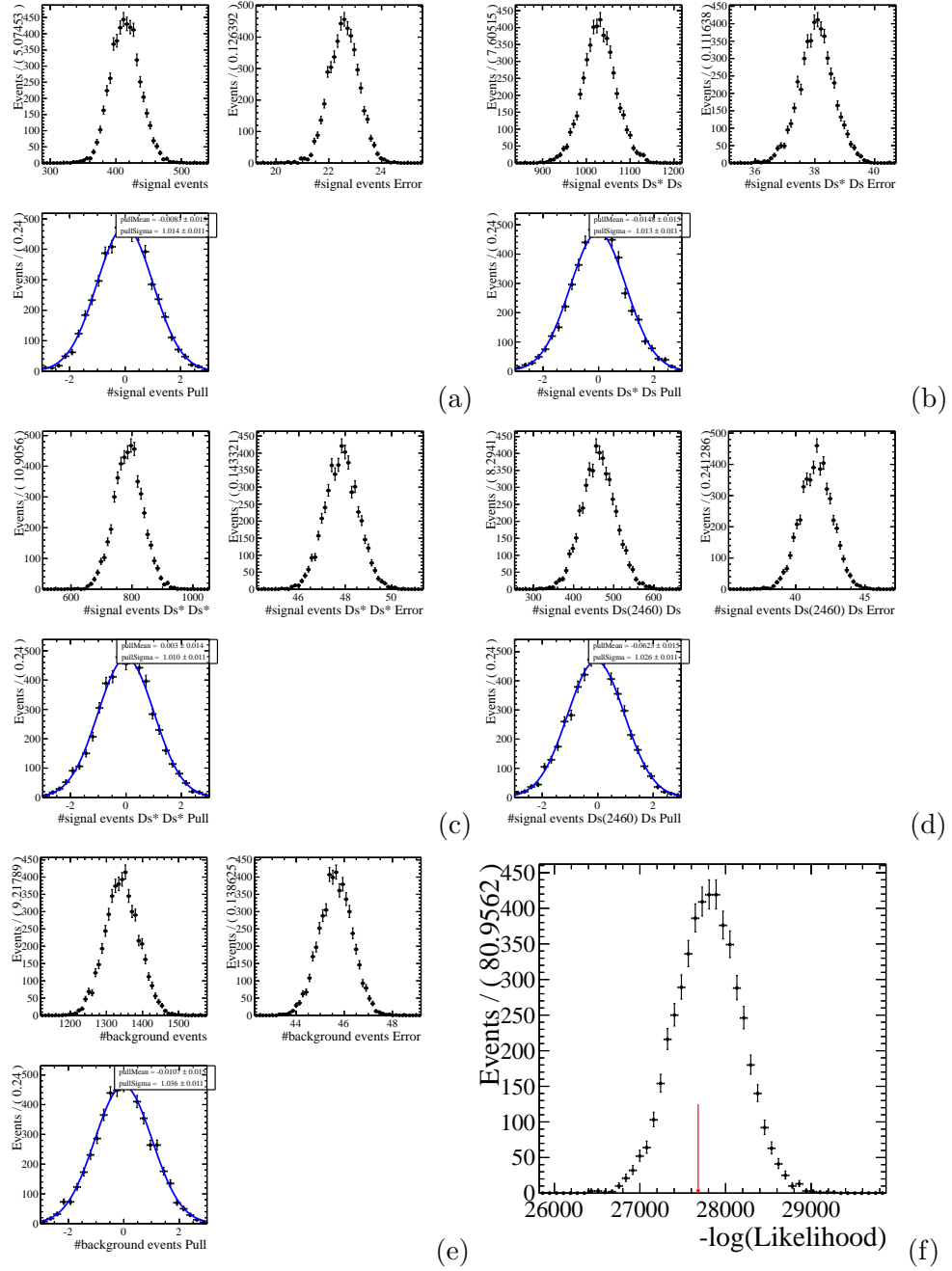


FIGURE B.1: Result of toy studies, showing the distributions for (a)  $D_s^+ D_s^-$  signal yield, (b)  $D_s^{*+} D_s^-$  signal yield, (c)  $D_s^{*+} D_s^{*-}$  signal yield, (d)  $D_s(2460)^+ D_s^-$  background yield (e) combinatorial background yield, (f)  $-\log(\mathcal{L})$ , where the arrow represents the value produced in the final data fit.

## Appendix C

# Data Fits with Varying BDT Cut

The plots in this section show the effect that varying the BDT cut has on the final data fit. Higher values for the cut correspond to a tighter cut, and greater removal of the combinatorial background, as shown in the plots. However this comes with a loss of signal efficiency, which becomes detrimental after a certain point. The signal and background yields were used to choose the optimal BDT cut. The plots demonstrate the power of the BDT to remove combinatorial background; in the plot with the tightest cut, it is almost entirely removed. By considering the yields, the optimal cut was set at  $-0.1$ ; this is the case in the final data plot shown in 5.24

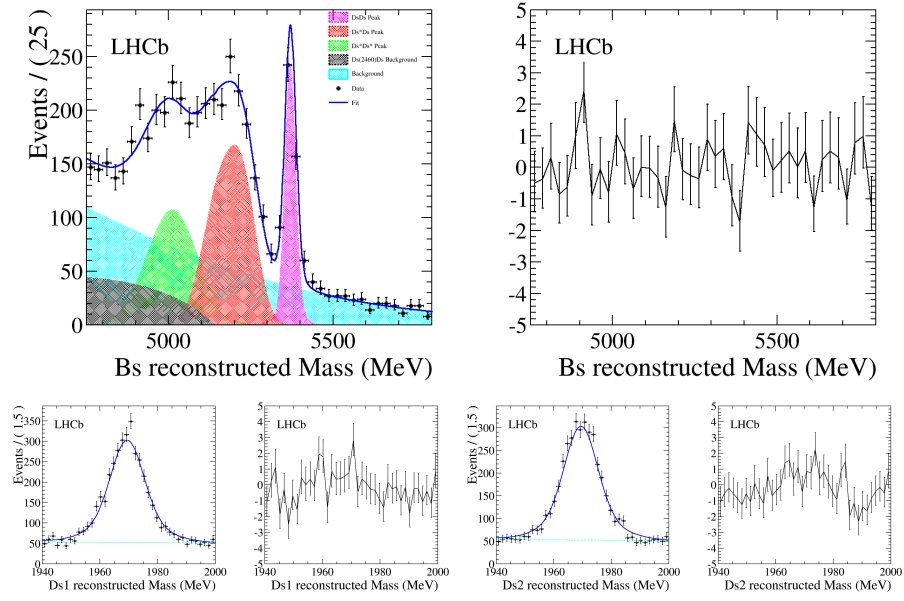


FIGURE C.1: The final  $B_s^0 \rightarrow D_s^+ D_s^-$  and  $D_s^\pm$  data plots using a BDT cut of greater than  $-0.2$ .

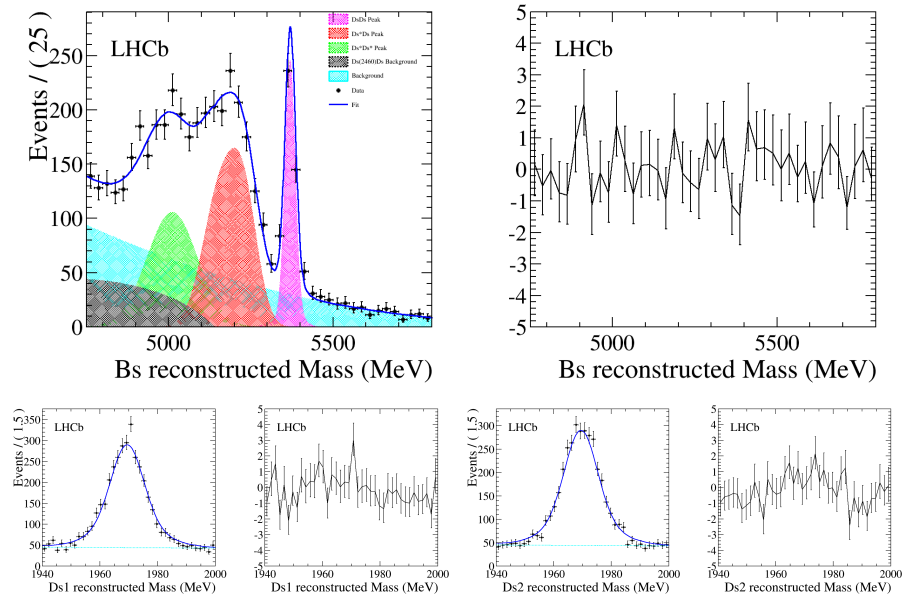


FIGURE C.2: The final  $B_s^0 \rightarrow D_s^+ D_s^-$  and  $D_s^\pm$  data plots using a BDT cut of greater than  $-0.15$ .

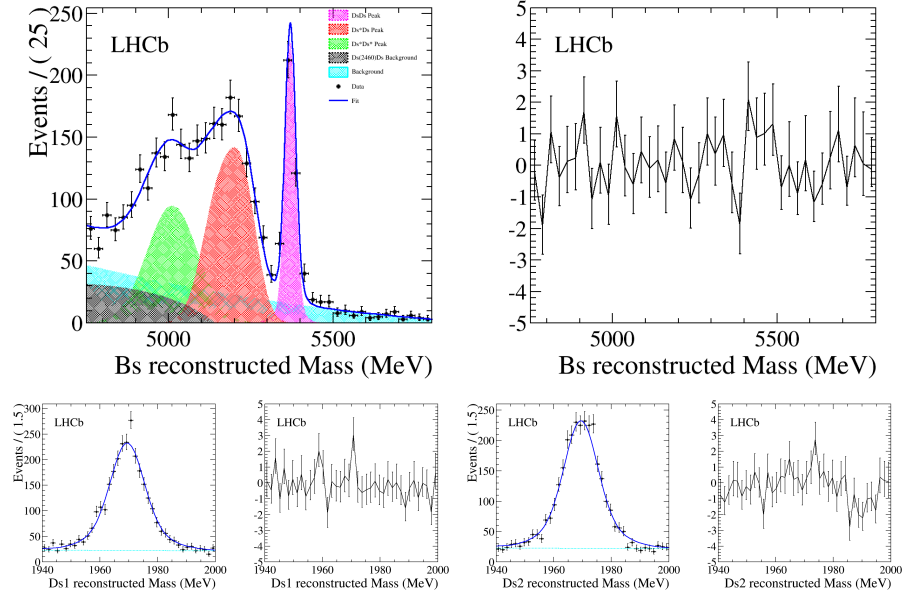


FIGURE C.3: The final  $B_s^0 \rightarrow D_s^+ D_s^-$  and  $D_s^\pm$  data plots using a BDT cut of greater than  $-0.05$ .

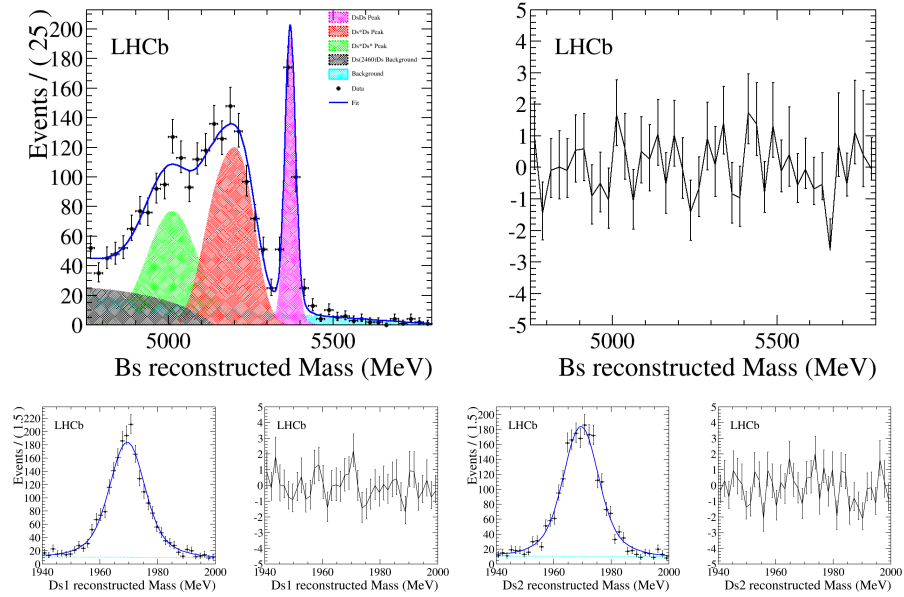


FIGURE C.4: The final  $B_s^0 \rightarrow D_s^+ D_s^-$  and  $D_s^\pm$  data plots using a BDT cut of greater than  $0.0$ .



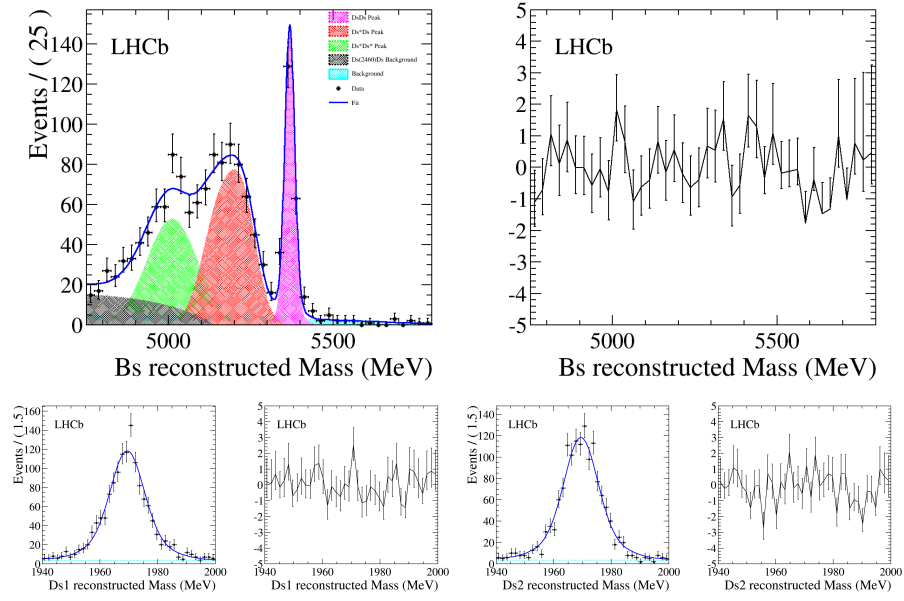


FIGURE C.5: The final  $B_s^0 \rightarrow D_s^+ D_s^-$  and  $D_s^\pm$  data plots using a BDT cut of greater than 0.05.

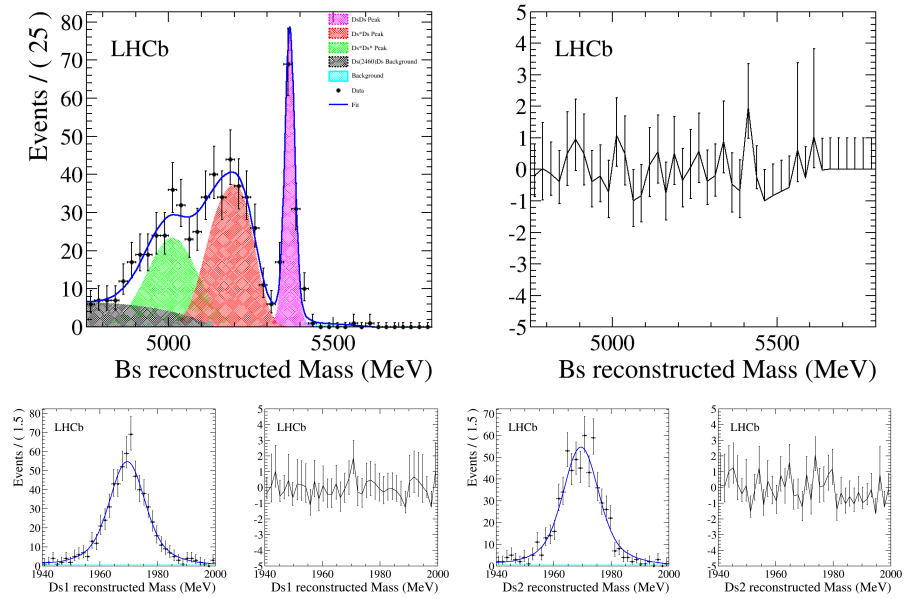


FIGURE C.6: The final  $B_s^0 \rightarrow D_s^+ D_s^-$  and  $D_s^\pm$  data plots using a BDT cut of greater than 0.1.

## Appendix D

### $D_s(2460)^+$ Systematic Study Plots

This appendix contains the two final mass plots when a parameter for the Argus function used to fit the contribution from the  $D_s(2460)^\pm$  is varied to plus or minus one standard deviation of the value found when fitting to simulation. This parameter governs the rate at which the function increases below the mass cut-off, and, due to the size of the uncertainty obtained from fitting to simulation, is the dominant source of systematic uncertainty in the yield.

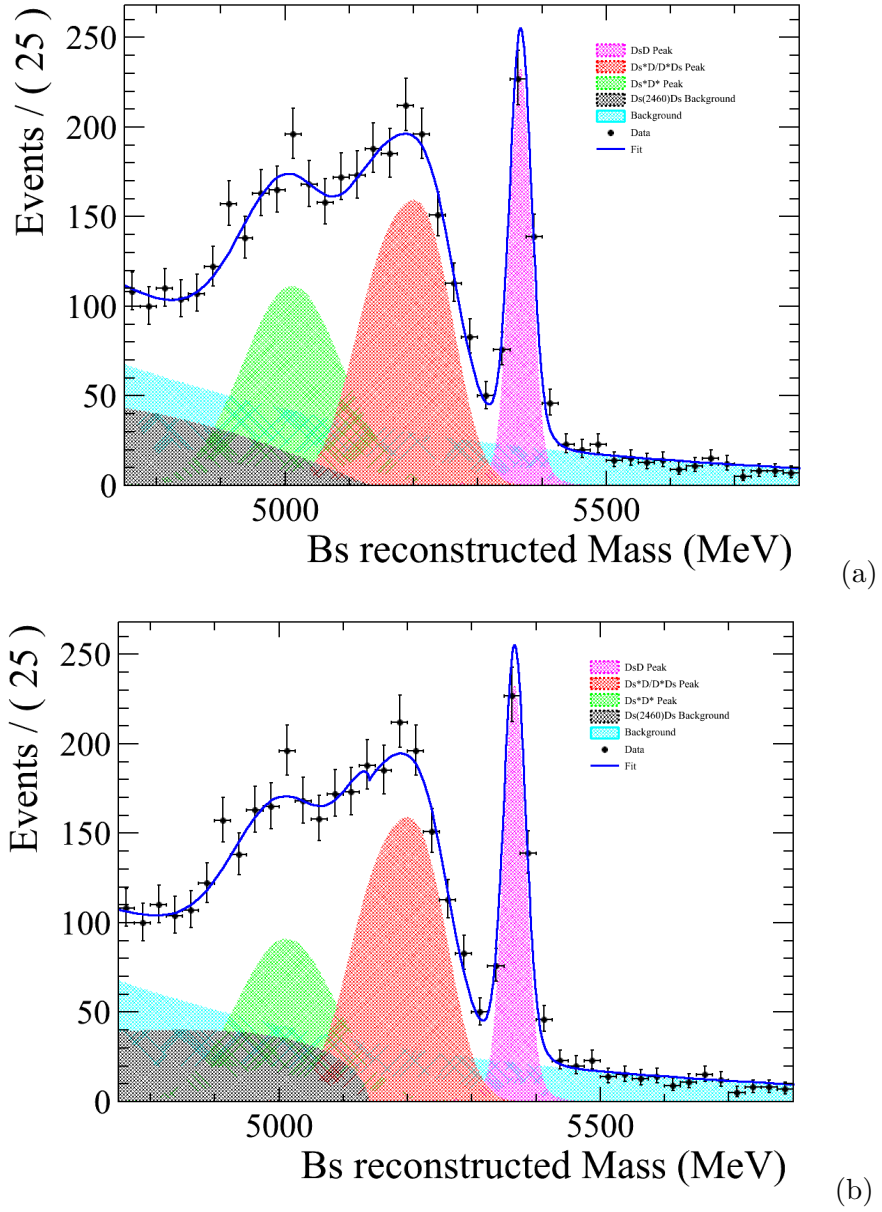


FIGURE D.1: The two final mass distributions when the Argus parameter governing the rate of increase in the function below the mass cut off is varied (a) to plus one sigma of its simulation value, and (b) to minus one sigma from the simulation value. The discontinuity in this plot at around 5120 MeV appears to be unphysical, resulting from the sharp cut off in the  $D_s(2460)^+$  distribution in this case. The cut off is unlikely to be this sharp in reality.

# Appendix E

## Full Data Fit Parameters

Parameter	Fixed/Floated	Value
$D_s^+ D_s^-$ mean	Floated	$5369.80 \pm 1.00$
$D_s^+ D_s^- \sigma_{CB}$	Floated	$32.34 \pm 1.86$
$D_s^+ D_s^- \alpha_{CB}$	Fixed to MC	$1.07 \pm 0.18$
$D_s^+ D_s^- n_{CB}$	Fixed to MC	$80.0 \pm 60.0$
$D_s^+ D_s^- \frac{\sigma_G}{\sigma_{CB}}$	Fixed to MC	$0.485 \pm 0.024$
$D_s^+ D_s^- \text{frac}_{CB}$	Fixed to MC	$0.143 \pm 0.027$
$D_s^{*+} D_s^-$ Yield	Floated	$412 \pm 23$
$D_s^{*+} D_s^-$ Mean 1	Fixed to MC	$5156.91 \pm 2.89$
$D_s^{*+} D_s^-$ Mean 2	Fixed to MC	$5231.39 \pm 2.99$
$D_s^{*+} D_s^- \sigma_1$	Fixed to $B_s^0$ only data fit	$52.58 \pm 3.49$
$D_s^{*+} D_s^- \frac{\sigma_1}{\sigma_2}$	Fixed to MC	$0.787 \pm 0.047$
$D_s^{*+} D_s^- \text{frac}_1$	Fixed to MC	$0.613 \pm 0.041$
$D_s^{*+} D_s^-$ Yield	Floated	$1032 \pm 39$
$D_s^{*+} D_s^{*-}$ Mean	Fixed to MC	$5011.13 \pm 1.69$
$D_s^{*+} D_s^{*-} \sigma$	Fixed to $B_s^0$ only data fit	$77.48 \pm 8.05$
$D_s^{*+} D_s^{*-}$ Yield	Floated	$786 \pm 48$
$D_s^+$ (2460) ARGUS $m$	Fixed to MC	$0.886 \pm 0.361$
$D_s^+$ (2460) ARGUS Cut off	Fixed to MC	$5140.22 \pm 7.70$
$D_s^+$ (2460) ARGUS parameter	Fixed to MC	$-4.19 \pm 4.96$
$D_s^+$ (2460) Yield	Floated	$432 \pm 42$
$D_s^+$ mean	Fixed to $D_s^+$ only data fit	$1969.43 \pm 0.12$
$D_s^+ \sigma_{CB}$	Fixed to $D_s^+$ only data fit	$11.70 \pm 0.56$
$D_s^+ \alpha_{CB}$	Fixed to MC	$1.99 \pm 0.26$
$D_s^+ n_{CB}$	Fixed to MC	$0.49 \pm 0.78$
$D_s^+ \frac{\sigma_G}{\sigma_{CB}}$	Fixed to MC	$0.479 \pm 0.010$
$D_s^+ \text{frac}_{CB}$	Fixed to MC	$0.271 \pm 0.021$
$B_s^0$ background c0	Floated	$-0.947 \pm 0.037$
$B_s^0$ background c1	Floated	$0.148 \pm 0.042$
$D_s^+$ background c0	Fixed to $D_s^+$ only data fit	$-0.022 \pm 0.051$
Combinatorial background yield	Floated	$1342 \pm 47$

TABLE E.1: Every parameter in the final signal channel data fit, whether the parameter is floated or fixed in the final fit (and, if fixed, what to), and the final value of the parameter.

Parameter	Fixed/Floated	Value
$D_s^\pm D^\mp$ mean	Floated	$5284.25 \pm 0.28$
$D_s^\pm D^\mp \sigma_{CB}$	Fixed to $B$ only data fit	$13.60 \pm 0.25$
$D_s^\pm D^\mp \alpha_{CB}$	Fixed to MC	$2.33 \pm 0.17$
$D_s^\pm D^\mp n_{CB}$	Fixed to MC	$1.02 \pm 0.39$
$D_s^\pm D^\mp \frac{\sigma_G}{\sigma_{CB}}$	Fixed to MC	$1.784 \pm 0.077$
$D_s^\pm D^\mp \text{frac}_{CB}$	Fixed to MC	$0.785 \pm 0.055$
$D_s^\pm D^\mp$ Yield	Floated	$3615 \pm 64$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp}$ Mean 1	Fixed to $B$ only data fit	$5128.45 \pm 2.90$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp}$ Mean 2	Fixed to $B$ only data fit	$5099.87 \pm 3.12$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp}$ Mean 3	Fixed to $B$ only data fit	$5068.72 \pm 1.79$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp} \sigma_1$	Fixed to $B$ only data fit	$9.69 \pm 7.75$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp} \frac{\sigma_1}{\sigma_2}$	Fixed to $B$ only data fit	$0.394 \pm 0.217$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp} \sigma_3$	Fixed to $B$ only data fit	$54.46 \pm 1.44$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp} \text{frac}_1$	Fixed to $B$ only data fit	$0.020 \pm 0.009$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp} \text{frac}_2$	Fixed to $B$ only data fit	$0.046 \pm 0.022$
$D_s^{*\pm} D^\mp / D_s^\pm D^{*\mp}$ Yield	Floated	$3556 \pm 110$
$D_s^{*\pm} D^{*\mp}$ Mean	Fixed to MC	$4916.99 \pm 1.87$
$D_s^{*\pm} D^{*\mp} \sigma$	Fixed to $B$ only data fit	$90.40 \pm 5.63$
$D_s^{*\pm} D^{*\mp}$ Yield	Floated	$195 \pm 86$
$D_s^+$ mean	Fixed to $D_s^+$ only data fit	$1969.36 \pm 0.09$
$D_s^+ \sigma_{CB}$	Fixed to $D_s^+$ only data fit	$11.65 \pm 0.18$
$D_s^+ \alpha_{CB}$	Fixed to MC	$1.23 \pm 0.14$
$D_s^+ n_{CB}$	Fixed to MC	$9.99 \pm 5.74$
$D_s^+ \frac{\sigma_G}{\sigma_{CB}}$	Fixed to MC	$0.498 \pm 0.016$
$D_s^+ \text{frac}_{CB}$	Fixed to MC	$0.744 \pm 0.032$
$D^+$ mean	Fixed to $D^+$ only data fit	$1871.70 \pm 0.13$
$D^+ \sigma_{CB}$	Fixed to $D^+$ only data fit	$14.76 \pm 0.28$
$D^+ \alpha_{CB}$	Fixed to MC	$1.10 \pm 0.09$
$D^+ n_{CB}$	Fixed to MC	$29.95 \pm 15.36$
$D^+ \frac{\sigma_G}{\sigma_{CB}}$	Fixed to MC	$0.490 \pm 0.014$
$D^+ \text{frac}_{CB}$	Fixed to MC	$0.753 \pm 0.027$
$B$ background c0	Floated	$-0.285 \pm 0.055$
$B$ background c1	Floated	$-0.117 \pm 0.056$
$D_s^+$ background c0	Fixed to $D_s^+$ only data fit	$0.038 \pm 0.069$
$D^+$ background c0	Fixed to $D^+$ only data fit	$0.299 \pm 0.065$
Combinatorial background yield	Floated	$1643 \pm 56$

TABLE E.2: Every parameter in the final normalisation channel data fit, whether the parameter is floated or fixed in the final fit (and, if fixed, what to), and the final value of the parameter.

# Appendix F

## Variable descriptions

The following appendix is a description of all variables used in the  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  branching fraction analysis described in chapter 5.

### F.1 BDT Variables

$B_s^0$  *Impact Parameter  $\chi^2$  to own Primary Vertex (IP\_CHI2\_OWNPV)*: The impact parameter is the perpendicular distance between the reconstructed path of the  $B_s^0$  and the identified primary vertex. A lower  $\chi^2$  value is desired as it ensures the  $B_s^0$  did come from the primary vertex.

$B_s^0$  *Transverse Momentum (PT)*: The portion of the  $B_s^0$  momentum that is perpendicular to the beamline. Higher  $p_T$  particles are easier to subsequently reconstruct through their decay products.

$D_s^\pm$  *Impact Parameter  $\chi^2$  to own Primary Vertex (IP\_CHI2\_OWNPV)*: A higher value is preferred for this  $\chi^2$  as this makes it less likely that the  $D_s^\pm$  actually came from the primary vertex, rather than being produced as a result of the  $B_s^0$  decay as required.

$D_s^\pm$  *Transverse Momentum (PT)*: The portion of the  $D_s^\pm$  momentum that is perpendicular to the beamline. Again,  $D_s^\pm$  with higher  $p_T$  are easier to reconstruct from their decay products.

$D_s^\pm$  *Direction Angle to own Primary Vertex (DIRA\_OWNPV)*: The cosine of the angle between the  $D_s^\pm$  momentum vector and the primary vertex in the event. To ensure the  $D_s^\pm$  is not from the primary vertex, and is travelling forwards in the detector.

$D_s^\pm$  **Direction Angle to own Origin Vertex (DIRA\_ORIVX)**: The cosine of the angle between the  $D_s^\pm$  momentum vector and the production vertex for that  $D_s^\pm$ . To ensure the  $D_s^\pm$  is produced at its associated vertex, and is travelling forwards in the detector.

$D_s^\pm$  **Lifetime (TAU)**: The lifetime of the  $D_s^\pm$ . An extended lifetime is characteristic of the decays of particles containing  $b$  and  $c$  quarks. The longer the particle lives, the further it travels in the detector, and the more accurately its decay can be assessed. This helps reduce combinatorial background from ghost tracks.

$D_s^\pm$  **Product of Daughter Transverse Momentum (ProductDsDauPT)**: The product of the transverse momentum of the  $K^+$ ,  $K^-$ , and  $\pi^\pm$  daughters of the  $D_s^\pm$ . Tracks associated with low transverse momentum are harder to measure accurately, and thus the best case scenario is when the  $D_s^\pm$  transverse momentum is shared equally amongst the daughters. This scenario maximises the product of the daughter transverse momentum.

## F.2 Other Variables

$D_s^\pm$  **Distance Of Closest Approach (DOCA)**: The DOCA of two particles is the closest distance that they pass to each other in the event reconstruction. In the stripping line this is set to less than 0.5 mm; this is to ensure the two  $D_s^\pm$  candidates originate from the same point.

**Track**  $\chi^2/\text{ndf}$  : This is a track quality variable. The lower the value, the better quality of track, and the lower chance that the track is actually a fake track.

**Vertex**  $\chi^2/\text{ndf}$  : This variable assesses the quality of a specified vertex. The lower the value, the better the quality of the vertex, and the greater the likelihood that the vertex is real and not just a collection of particles that are not associated.

**Linear Particle ID Variables (PIDK/PID $\pi$ )**: There are two separate types of variable used for particle identification in LHCb analyses. These PID variables make use of log likelihood  $\ln \mathcal{L}$  information provided by the detector's various sub-systems to compute a likelihood for each particle mass hypothesis, relative to the pion hypothesis.

**Multivariate Particle ID Variables (ProbNNk/ProbNN $\pi$ )**: The second set of particle identification variables in LHCb make use of a Neural Net to improve on the  $\ln \mathcal{L}$  variables by including more information, and considering correlations between detector systems.

# Bibliography

- [1] Dirac, P. A. M. The Quantum Theory of the Electron. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 117(778):610–624, 1928.
- [2] O. Johns. *Analytical Mechanics for Relativity and Quantum Mechanics*. Oxford Graduate Texts. OUP Oxford, 2011.
- [3] K.A. Olive et al. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014.
- [4] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.
- [5] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys.Rev.Lett.*, 13:508–509, 1964.
- [6] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–323, 1964.
- [7] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.
- [8] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.
- [9] Wim de Boer. The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology. *arXiv*, hep-ph:1309.0721, 2013.
- [10] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken symmetries. *Phys. Rev.*, 127:965–970, Aug 1962.
- [11] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys.Rev.Lett.*, 81:1562–1567, 1998.
- [12] Marina Artuso, Elisabetta Barberio, and Sheldon Stone. *B Meson Decays*. *PMC Phys.*, A3:3, 2009.
- [13] H. Albrecht et al. Observation of  $B^0$  - anti- $B^0$  Mixing. *Phys.Lett.*, B192:245–252, 1987.



- [14] A. Abulencia et al. Observation of  $B^0(s)$  - anti- $B^0(s)$  Oscillations. *Phys.Rev.Lett.*, 97:242003, 2006.
- [15] R. Aaij et al. Effective lifetime measurements in the  $B_s^0 K^+ K^-$ ,  $B^0 K^+ K^-$  and  $B_s^{0+} K^-$  decays. *Phys.Lett.*, B736:446–454, 2014.
- [16] R. Aaij et al. Measurement of the  $B_s$  effective lifetime in the  $J/\psi f_0(980)$  final state. *Phys.Rev.Lett.*, 109:152002, 2012.
- [17] Ashton B. Carter and A. I. Sanda. CP nonconservation in cascade decays of  $b$  mesons. *Phys. Rev. Lett.*, 45:952–954, Sep 1980.
- [18] R. Aaij et al. Measurement of the CP-violating phase  $\phi_s$  in the decay  $B_s^0 \rightarrow J/\psi \phi$ . *Phys.Rev.Lett.*, 108:101803, 2012.
- [19] Y. Amhis et al. Averages of  $b$ -hadron,  $c$ -hadron, and  $\tau$ -lepton properties as of summer 2014. 2014.
- [20] Alexander Lenz and Ulrich Nierste. Numerical Updates of Lifetimes and Mixing Parameters of B Mesons. *arXiv*, hep-ph:1102.4274, 2011.
- [21] J. Charles, O. Deschamps, S. Descotes-Genon, R. Itoh, H. Lacker, et al. Predictions of selected flavour observables within the Standard Model. *Phys.Rev.*, D84:033005, 2011.
- [22] J.H. Christenson, J.W. Cronin, V.L. Fitch, and R. Turlay. Evidence for the 2 pi Decay of the  $K(2^0)$  Meson. *Phys.Rev.Lett.*, 13:138–140, 1964.
- [23] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.
- [24] Andrzej J. Buras, Markus E. Lautenbacher, and Gaby Ostermaier. Waiting for the top quark mass,  $K^+ \rightarrow \pi^+$  neutrino anti-neutrino,  $B(s)0$  - anti- $B(s)0$  mixing and CP asymmetries in B decays. *Phys.Rev.*, D50:3433–3446, 1994.
- [25] CKMfitter collaboration. CKMfitter: Preliminary results as of Summer 2014, 2014. [http://ckmfitter.in2p3.fr/www/results/plots\\_ckm14/ckm\\_res\\_ckm14.html](http://ckmfitter.in2p3.fr/www/results/plots_ckm14/ckm_res_ckm14.html).
- [26] A. Ceccucci, Z. Ligeti, and Y. Sakai. CKMfitter: Preliminary results as of Summer 2014, 2014. <http://pdg.lbl.gov/2014/reviews/rpp2014-rev-ckm-matrix.pdf>.
- [27] M. Battaglia, A.J. Buras, P. Gambino, A. Stocchi, D. Abbaneo, et al. The CKM matrix and the unitarity triangle. Workshop, CERN, Geneva, Switzerland, 13-16 Feb 2002: Proceedings. 2003.
- [28] B.R. Martin and G. Shaw. *Particle Physics*. Manchester Physics Series. Wiley, 2008.
- [29] E. Halkiadakis. Proceedings for TASI 2009 Summer School on 'Physics of the Large and the Small': Introduction to the LHC experiments. 2010.

- [30] AC Team. Diagram of an LHC dipole magnet. Schma d'un aimant dipole du LHC. June 1999.
- [31] M Lamont. Input from Evian 2011. page 6 p, 2012.
- [32] A Augusto Alves et al. The LHCb Detector at the LHC. *J. Instrum.*, 3(LHCb-DP-2008-001. CERN-LHCb-DP-2008-001):S08005, 2008. Also published by CERN Geneva in 2010.
- [33] Eddy Jans. The LHCb detector. *arXiv*, hep-ex:0910.1740, 2009.
- [34] H Burkhardt and P Grafstrm. Absolute Luminosity from Machine Parameters. Technical Report LHC-PROJECT-Report-1019. CERN-LHC-PROJECT-Report-1019, 2007.
- [35] Paula Collins. The LHCb Upgrade. Technical Report arXiv:1108.1403, Aug 2011. Comments: Proceedings of FPCP 2011.
- [36] Antunes-Nobrega et al. *LHCb reoptimized detector design and performance: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2003.
- [37] Mark Tobin. Performance of the LHCb Tracking Detectors. Technical Report LHCb-PROC-2013-015. CERN-LHCb-PROC-2013-015, CERN, Geneva, Mar 2013.
- [38] LHCb: Inner tracker technical design report. (CERN-LHCC-2002-029.), 2002.
- [39] D van Eijk et al. Radiation hardness of the LHCb Outer Tracker. *Nucl. Instrum. Methods Phys. Res., A*, 685(LHCb-DP-2012-001. CERN-LHCb-DP-2012-001):62–69, 2012.
- [40] M. Adinolfi et al. Performance of the LHCb RICH detector at the LHC. *Eur. Phys. J.*, C73:2431, 2013.
- [41] Christian Lippmann. Particle identification. *Nucl.Instrum.Meth.*, A666:148–172, 2012.
- [42] S Amato et al. *LHCb RICH: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.
- [43] S Amato et al. *LHCb magnet: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.
- [44] S Amato et al. *LHCb calorimeters: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.
- [45] Pascal Perret and Xavier Vilass-Cardona. Performance of the LHCb calorimeters during the period 2010-2012. *J.Phys.Conf.Ser.*, 587(1):012012, 2015.
- [46] R. Aaij et al. Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  branching fraction and search for  $B^0 \rightarrow \mu^+ \mu^-$  decays at the LHCb experiment. *Phys.Rev.Lett.*, 111:101805, 2013.
- [47] P R Barbosa-Marinho et al. *LHCb muon system: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.

- [48] F. Archilli, W. Baldini, G. Bencivenni, N. Bondar, W. Bonivento, et al. Performance of the Muon Identification at LHCb. *JINST*, 8:P10020, 2013.
- [49] LHCb collaboration. The LHCb High Level Trigger description, 2011. <http://lhcb-trig.web.cern.ch/lhcb-trig/HLT/HltDescription.htm>.
- [50] M Clemencic et al. The LHCb simulation application, GAUSS: design, evolution and experience. *J. Phys. Conf. Ser.*, 331:032023, 2011.
- [51] I. Belyaev et al. Handling of the generation of primary events in GAUSS, the LHCb simulation framework. *Nuclear Science Symposium Conference Record (NSS/MIC)*, IEEE:1155, 2010.
- [52] Torbjörn Sjöstrand, Stephen Mrenna, and Peter” Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.
- [53] Torbjörn Sjöstrand, Stephen Mrenna, and Peter” Skands. A brief introduction to PYTHIA 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008.
- [54] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.
- [55] Piotr Golonka and Zbigniew Was. PHOTOS Monte Carlo: a precision tool for QED corrections in  $Z$  and  $W$  decays. *Eur.Phys.J.*, C45:97–107, 2006.
- [56] S. Agostinelli et al. Geant4: a simulation toolkit. *Nucl. Instrum. Meth.*, A506:250, 2003.
- [57] John Allison, K. Amako, J. Apostolakis, H. Araujo, P.A. Dubois, et al. Geant4 developments and applications. *IEEE Trans.Nucl.Sci.*, 53:270, 2006.
- [58] LHCb collaboration. The BOOLE Project, 2014. [lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/bool/](http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/bool/).
- [59] LHCb collaboration. The MOORE Project, 2014. [lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/](http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/).
- [60] LHCb collaboration. The BRUNEL Project, 2014. [lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/](http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/).
- [61] LHCb collaboration. The DAVINCI Project, 2014. [lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/](http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/).
- [62] Rene Brun and Fons Rademakers. Root an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(12):81 – 86, 1997. New Computing Techniques in Physics Research V.
- [63] B.G. Yacobi. *Semiconductor Materials: An Introduction to Basic Principles*. Microdevices Series. Kluwer Academic/Plenum Publishers, 2003.

- [64] Helmuth Spieler. *Semiconductor Detector Systems*. Oxford University Press, 2005.
- [65] Stephen Farry. *A measurement of Z production using tau final states with the LHCb detector*. PhD thesis, U. Coll., Dublin, Aug 2012. Presented 12 Nov 2012.
- [66] P R Barbosa-Marinho et al. *LHCb VELO (VERtex LOCator): Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2001.
- [67] Roel Aaij et al. LHCb Detector Performance. *Int. J. Mod. Phys. A*, 30(arXiv:1412.6352. CERN-PH-EP-2014-290. LHCb-DP-2014-002. CERN-LHCb-DP-2014-002):1530022. 82 p, Dec 2014.
- [68] Erik Aras Papadelis. Characterisation and Commissioning of the LHCb VELO Detector. 2009.
- [69] CERN. LHCb - Large Hadron Collider beauty experiment, 2008. <http://lhcb-public.web.cern.ch/lhcb-public/en/Detector/VELO-en.html>.
- [70] Graham Beck and Georg Viehhauser. Analytic model of thermal runaway in silicon detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 618(13):131 – 138, 2010.
- [71] R Aaij et al. Performance of the LHCb Vertex Locator. *J. Instrum.*, 9(arXiv:1405.7808. CERN-LHCb-DP-2014-001. LHCb-DP-2014-001):P09007. 61 p, May 2014. Comments: 61 pages, 33 figures.
- [72] E. Fretwurst, G. Lindstrom, I. Pintilie, J. Stahl, and I. Pintilie. Radiation damage in silicon detectors caused by hadronic and electromagnetic irradiation. *arXiv*, physics:0211118, 2002.
- [73] M. Huhtinen. Simulation of non-ionising energy loss and defect formation in silicon. *Nucl.Instrum.Meth.*, A491:194–215, 2002.
- [74] J.R. Srouf, C.J. Marshall, and P.W. Marshall. Review of displacement damage effects in silicon devices. *IEEE Transactions on Nuclear Science*, 50:653–670, 2003.
- [75] A. Vasilescu (INPE Bucharest) and G. Lindstroem (University of Hamburg). Displacement damage in silicon, on-line compilation, 2000. <http://rd50.web.cern.ch/rd50/NIEL/default.html>.
- [76] M. Moll, E. Fretwurst, and G. Lindstrm. Leakage current of hadron irradiated silicon detectors material dependence. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 426(1):87 – 93, 1999.
- [77] H.W. Kraner, Z. Li, and K.U. Posnecker. Fast neutron damage in silicon detectors. *Nucl.Instrum.Meth.*, A279:266–271, 1989.

- [78] A. Affolder, K. Akiba, M. Alexander, S. Ali, M. Artuso, et al. Radiation damage in the LHCb Vertex Locator. *JINST*, 8:P08002, 2013.
- [79] Michael Moll. *Radiation damage in silicon particle detectors: microscopic defects and macroscopic properties*. PhD thesis, Hamburg Univ., Hamburg, 1999. Presented on 30 Nov 1999.
- [80] Henry Brown, Themis Bowcock, and David Hutchcroft. *Monitoring Radiation Damage in the Vertex Locator and Top Pair Production in LHCb*. PhD thesis, Liverpool U., Sep 2013. Presented 07 Nov 2013.
- [81] Alexandre Chilingarov. Intrinsic concentration and generation current temperature dependence, 2011. <http://eprints.lancs.ac.uk/39969/>.
- [82] G.L. Casse, D. Hutchcroft, D. Jones, and G.D. Patel. VELO module production: Vacuum tank tests. (CERN-LHCB-2007-082), 2008.
- [83] LHCb Collaboration. LHCb Tracker Upgrade Technical Design Report. Technical Report CERN-LHCC-2014-001. LHCB-TDR-015, CERN, Geneva, Feb 2014.
- [84] Andrew Hickling et al. Use of IT (current vs temperature) scans to study radiation damage in the LHCb VELO. Technical Report LHCb-PUB-2011-021. CERN-LHCb-PUB-2011-021, CERN, Geneva, Oct 2011.
- [85] R. Aaij et al. Measurement of the  $\bar{B}_s^0 \rightarrow D_s^- D_s^+$  and  $\bar{B}_s^0 \rightarrow D^- D_s^+$  effective lifetimes. *Phys.Rev.Lett.*, 112(11):111802, 2014.
- [86] Robert Fleischer and Robert Knegjens. Effective Lifetimes of  $B_s$  Decays and their Constraints on the  $B_s^0$ - $\bar{B}_s^0$  Mixing Parameters. *Eur.Phys.J.*, C71:1789, 2011.
- [87] Robert Fleischer and Robert Knegjens. In Pursuit of New Physics With  $B_s^0 \rightarrow K^+ K^-$ . *Eur.Phys.J.*, C71:1532, 2011.
- [88] Roel Aaij et al. Measurement of the  $CP$ -violating phase  $\phi_s$  in  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  decays. *Phys.Rev.Lett.*, 113(21):211801, 2014.
- [89] R. Aleksan, A. Le Yaouanc, L. Oliver, O. Pne, and J.-C. Raynal. Estimation of for the bsbs system. exclusive decays and the parton model. *Physics Letters B*, 316(4):567 – 577, 1993.
- [90] Alexander Lenz. Theoretical update of  $B$ -Mixing and Lifetimes. *arXiv*, hep-ph:1205.1444, 2012.
- [91] A. J. Bevan et al. The Physics of the  $B$  Factories. *Eur. Phys. J.*, C74:3026, 2014.
- [92] S. Esen et al. Precise measurement of the branching fractions for  $B_s \rightarrow D_s^{(*)+} D_s^{(*)-}$  and first measurement of the  $D_s^{*+} D_s^{*-}$  polarization using  $e^+ e^-$  collisions. *Phys.Rev.*, D87(3):031101, 2013.

- [93] Jonathan L. Rosner. Determination of pseudoscalar-charmed-meson decay constants from  $b$ -meson decays. *Phys. Rev. D*, 42:3732–3740, Dec 1990.
- [94] Isard Dunietz, Robert Fleischer, and Ulrich Nierste. In pursuit of new physics with  $B_s$  decays. *Phys.Rev.*, D63:114015, 2001.
- [95] Alexander Lenz and Ulrich Nierste. Theoretical update of  $B_s - \bar{B}_s$  mixing. *JHEP*, 0706:072, 2007.
- [96] Chun-Khiang Chua, Wei-Shu Hou, and Chia-Hsien Shen. Long-Distance Contribution to  $\Delta\Gamma_s/\Gamma_s$  of the  $B_s^0 - \bar{B}_s^0$  System. *Phys.Rev.*, D84:074037, 2011.
- [97] Fabian Krinner, Alexander Lenz, and Thomas Rauh. The inclusive decay revisited. *Nuclear Physics B*, 876(1):31 – 54, 2013.
- [98] T. Aaltonen et al. Measurement of  $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$  Branching Ratios. *Phys.Rev.Lett.*, 108:201801, 2012.
- [99] M Williams, Vladimir V Gligorov, C Thomas, H Dijkstra, J Nardulli, and P Spradlin. The HLT2 Topological Lines. Technical Report LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan 2011.
- [100] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and regression trees*. Wadsworth international group, Belmont, California, USA, 1984.
- [101] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl.Instrum.Meth.*, A543:577–584, 2005.
- [102] Robert E. Schapire and Yoav Freund. A decision-theoretic generalization of on-line learning and an application to boosting. *Jour. Comp. and Syst. Sc.*, 55:119, 1997.
- [103] LHCb collaboration. The PIDCalib Package, 2015. <https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage>.
- [104] Adrian Bevan. *Statistical Data Analysis for the Physical Sciences*. Cambridge University Press, 2013.
- [105] Tomasz Skwarnicki. *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*. PhD thesis, Institute of Nuclear Physics, Krakow, 1986. DESY-F31-86-02.
- [106] J. Beringer et al. Review of particle physics. *Phys. Rev.*, D86:010001, 2012. and 2013 partial update for the 2014 edition.
- [107] H. Albrecht et al. Search for hadronic  $b_u$  decays. *Physics Letters B*, 241(2):278 – 282, 1990.

- 
- [108] Storaci B. Updated average  $f_s/f_d$   $b$ -hadron production fraction ratio for 7 TeV  $pp$  collisions. 2013.
- [109] Victor Mukhamedovich Abazov et al. Measurement of the CP-violating phase  $\phi_s^{J/\psi\phi}$  using the flavor-tagged decay  $B_s^0 \rightarrow J/\psi\phi$  in 8 fb $^{-1}$  of  $p\bar{p}$  collisions. *Phys.Rev.*, D85:032006, 2012.