

# Rare hadronic charged B meson decays at LHCb



Tom Hadavizadeh  
St Peter's College  
University of Oxford

A thesis submitted for the degree of  
*Doctor of Philosophy*

Trinity 2018

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Standard Model of Particle Physics . . . . .	3
1.1.1	Building blocks . . . . .	3
1.1.2	Parameters . . . . .	7
1.1.3	Triumphs . . . . .	8
1.1.4	Shortfalls . . . . .	9
1.2	Beyond the Standard Model . . . . .	11
1.2.1	Supersymmetry . . . . .	11
1.2.2	Leptoquarks . . . . .	12
1.2.3	Extra dimensions . . . . .	13
1.2.4	Axions . . . . .	13
<b>2</b>	<b>Theory</b>	<b>14</b>
2.1	The weak force in $b$ -hadron decays . . . . .	14
2.1.1	The weak force . . . . .	15
2.1.2	Weak interactions of quarks . . . . .	17
2.1.3	The CKM matrix . . . . .	17
2.1.4	$b$ -hadron physics . . . . .	19
2.1.5	QCD and hadronisation . . . . .	19
2.2	Annihilation topology decays . . . . .	21
2.2.1	Pure annihilation topology decays . . . . .	22
2.3	Rescattering . . . . .	23
2.4	Theoretical predictions for the $B^+ \rightarrow D_s^+ \phi$ decay . . . . .	24

2.4.1	Standard model predictions . . . . .	24
2.4.2	BSM models and predictions . . . . .	25
2.4.3	Previous measurements . . . . .	26
2.5	Theoretical predictions for the $B^+ \rightarrow D_s^+ K^+ K^-$ decay . . . . .	26
2.5.1	Standard model predictions . . . . .	28
<b>3</b>	<b>The LHCb experiment</b>	<b>29</b>
3.1	CERN and the LHC . . . . .	30
3.1.1	The accelerator complex . . . . .	31
3.1.2	Beam conditions . . . . .	33
3.2	The LHCb detector . . . . .	34
3.2.1	Magnet . . . . .	36
3.2.2	Vertex Locator . . . . .	37
3.2.3	Silicon Tracker . . . . .	41
3.2.4	Outer Tracker . . . . .	43
3.2.5	Ring imaging Cherenkov detectors . . . . .	45
3.2.6	Calorimeters . . . . .	49
3.2.7	Muon system . . . . .	52
3.2.8	Trigger . . . . .	54
3.2.9	Reconstruction . . . . .	57
3.3	Luminosity determination . . . . .	59
3.3.1	Van der Meer scans . . . . .	60
3.3.2	Beam gas imaging . . . . .	61
<b>4</b>	<b>Event selection</b>	<b>64</b>
4.1	Data samples and Simulations . . . . .	65
4.1.1	Data samples . . . . .	65
4.1.2	Simulation samples . . . . .	65
4.2	Online selection . . . . .	66
4.3	Offline selection . . . . .	69

4.3.1	Selection requirements . . . . .	70
4.3.2	Particle identification requirements . . . . .	79
4.3.3	Charmless and single-charm backgrounds . . . . .	79
4.3.4	Misidentified $D$ and $\Lambda_c^+$ hadrons . . . . .	81
4.3.5	Invariant mass vetoes . . . . .	83
4.3.6	Normalisation mode veto . . . . .	85
4.3.7	Multivariate analysis . . . . .	87
4.3.8	Impact parameter requirements . . . . .	103
4.4	Multiple candidates . . . . .	106
4.5	Refitting the decay chain . . . . .	107
4.6	Invariant mass distributions . . . . .	108
<b>5</b>	<b>Mass fit to <math>B^+ \rightarrow D_s^+ K^+ K^-</math> candidates</b>	<b>112</b>
5.1	Fit strategy . . . . .	113
5.2	Fit components . . . . .	114
5.2.1	Signal and normalisation decays . . . . .	114
5.2.2	Partially reconstructed backgrounds . . . . .	115
5.2.3	Combinatorial background . . . . .	121
5.3	Free and constrained parameters . . . . .	122
5.3.1	Parameter of interest . . . . .	122
5.3.2	Shape parameters . . . . .	122
5.3.3	Yields . . . . .	122
5.4	Fit validation . . . . .	123
5.5	Normalisation and signal fits . . . . .	125
5.6	Efficiency corrections . . . . .	128
5.6.1	Efficiencies from simulations . . . . .	128
5.6.2	Efficiencies requiring calibration samples . . . . .	131
5.6.3	Total efficiency . . . . .	134
5.7	Systematic uncertainties . . . . .	135
5.7.1	Relative efficiencies . . . . .	135

5.7.2	Signal and normalisation PDFs . . . . .	137
5.7.3	Background PDFs . . . . .	138
5.7.4	Charmless contribution . . . . .	138
5.7.5	Total systematic uncertainty . . . . .	139
5.8	Results . . . . .	139
5.9	Outlook . . . . .	142
<b>6</b>	<b>Mass fit to <math>B^+ \rightarrow D_s^+ \phi</math> candidates</b>	<b>145</b>
6.1	Fit strategy . . . . .	146
6.1.1	Simultaneous categories . . . . .	147
6.1.2	$B^+ \rightarrow D_s^+ K^+ K^-$ model and assumptions . . . . .	149
6.2	Fit components . . . . .	155
6.2.1	Signal and normalisation decays . . . . .	156
6.2.2	Partially reconstructed backgrounds . . . . .	157
6.2.3	Combinatorial backgrounds . . . . .	165
6.3	Free and constrained parameters . . . . .	165
6.3.1	Parameter of interest . . . . .	165
6.3.2	Shape parameters . . . . .	166
6.3.3	Yields . . . . .	167
6.3.4	Fractions . . . . .	167
6.4	Fit validation . . . . .	169
6.5	Normalisation and signal fits . . . . .	172
6.6	Efficiency corrections . . . . .	175
6.6.1	Efficiencies from simulation . . . . .	175
6.6.2	Efficiencies requiring calibration samples . . . . .	178
6.6.3	Total efficiencies . . . . .	179
6.7	Systematic uncertainties . . . . .	180
6.7.1	Relative efficiencies . . . . .	180
6.7.2	Signal and normalisation PDFs . . . . .	182
6.7.3	Background PDFs . . . . .	182

6.7.4	Charmless contribution . . . . .	183
6.7.5	$B^+ \rightarrow D_s^+ K^+ K^-$ model assumption . . . . .	183
6.7.6	Total systematic uncertainty . . . . .	184
6.8	Results . . . . .	184
6.8.1	Limit setting . . . . .	186
6.8.2	Comparison to the previous measurement . . . . .	190
<b>A</b>	<b>Invariant mass fit to <math>B^+ \rightarrow D_s^+ \phi</math> candidates in each <math>D_s^+</math> decay mode category</b>	<b>191</b>
	<b>Bibliography</b>	<b>191</b>

# Chapter 1

## Introduction

### Contents

---

<b>1.1</b>	<b>The Standard Model of Particle Physics</b>	<b>3</b>
1.1.1	Building blocks	3
1.1.2	Parameters	7
1.1.3	Triumphs	8
1.1.4	Shortfalls	9
<b>1.2</b>	<b>Beyond the Standard Model</b>	<b>11</b>
1.2.1	Supersymmetry	11
1.2.2	Leptoquarks	12
1.2.3	Extra dimensions	13
1.2.4	Axions	13

---

Throughout history, society has endeavoured to glimpse into the heart of matter, breaking it down into its elementary building blocks. Although the term *atom* was originally used by the Ancient Greek philosophers to describe indivisible particles, modern particle physics took until the 19th century to take off; the first of today's elementary particles, the electron, was discovered by J.J. Thomson in 1897 [1]. Rutherford's discovery of the proton [2] and Chadwick's discovery of the neutron [3] all but completed the picture, as between them these three particles constitute essentially all stable matter in the universe.

In 1934 Yukawa proposed that the neutrons and protons were bound together using an interaction appropriately called the strong force [4]. This required a new particle to be hypothesised: the *meson* (meaning intermediate). The mass of this particle

would have to sit inbetween the light electron and heavier protons and neutrons. The electron correspondingly became part of a class of particles call *leptons* (meaning light) and the protons and neutrons part of *baryons* (meaning heavy). The search for the strong force mediator resulted in the joint discoveries of the pion and muon [5, 6], both initially thought to be mesons.

Throughout the 20th century further pieces of the puzzle were uncovered, including anti-electrons, neutrinos and a wealth of mesons and baryons. The seemingly bizarre complexity of particles was eventually explained by Gell-Mann [7] and Zweig [8] through the quark model. This proposed that mesons and baryons were in fact composite particles, made up of elementary constituents called *quarks*. These were considered point-like and could be combined rather simply: a quark and anti-quark combined to make a meson ( $q\bar{q}$ ), and three quarks or three anti-quarks combined to form baryons ( $qqq, \bar{q}\bar{q}\bar{q}$ )<sup>1</sup>.

The theoretical framework for the Standard Model of Particle Physics began with the unification of the electromagnetic and weak forces in the 1960s. The development of the strong interaction in the 1970s completed the theory in the form that it is understood today. During next 40 years the remaining particles predicted to be a part of this picture were rapidly discovered: the charm quark in 1974 [11, 12]; the  $\tau$  lepton in 1975 [13]; the  $b$ -quark in 1977 [14]; the  $W$  and  $Z$  bosons in 1983 [15, 16]; the  $t$ -quark in 1995 [17, 18] and finally the Higgs Boson in 2012 [19, 20]. A brief introduction to the Standard Model is detailed in this chapter: the successes and limitations are highlighted, along with potential theories that may supersede this model. A more detailed description of the relevant theoretical aspects that underpin the rare hadronic  $B$  meson decays are described in Chapter 2.

The experimental apparatus used to collect the data analysed in this thesis, the LHCb experiment, is described in Chapter 3, followed by the methodology used to process this data set in Chapter 4. Finally, the statistical methods used to extract measurements can be found in Chapters 5 and 6.

The research presented in this thesis has been published in Ref. [21].

---

<sup>1</sup>More exotic combinations ( $q\bar{q}q\bar{q}, qqqq\bar{q}\dots$ ) were also proposed and recently observed [9, 10].

## 1.1 The Standard Model of Particle Physics

The Standard Model (SM) provides explanation for a large number of natural phenomena to a high precision. There are currently a number of unresolved issues that are not explained, implying that this theory is not the complete description of the natural world. The theory was formulated in the 1970s, with the final particles being predicted and discovered by physicists around the world in the decades following.

### 1.1.1 Building blocks

The SM is composed of a small number of fundamental particles; some are the components of matter and others are force-carriers that mediate interactions. The fundamental particles can be combined into composite combinations, bound by these fundamental interactions to form a rich tapestry of observable particles.

Mathematically, the theory is described by the unitary product group  $SU(3)_C \times SU(2)_L \times U(1)_Y$ , as the components are represented as gauge field theories. Throughout this thesis decays will be represented by Feynman Diagrams, pictorial representations of mathematical expressions [22]. The standard model is encapsulated by a Lagrangian density

$$\mathcal{L} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{QED}} + \mathcal{L}_{\text{Weak}} + \mathcal{L}_{\text{Higgs}}, \quad (1.1)$$

where the total quantity receives contributions from the fundamental forces as well as the spontaneous symmetry breaking mechanism allowing the particles to obtain masses. The particles of the SM can be separated into fundamental fermions (particles with half-integer spin) and fundamental bosons (particles with integer spin). These are described in the following sections.

#### Fundamental fermions

The fundamental spin-1/2 fermions that constitute matter can be divided into two categories: leptons and quarks. The fermions can be split into *families* or *generations* containing two leptons and two quarks. Each of these contains an up-type and down-type quark, a charged lepton and a neutrino. There are three of these *generations*

that are identical clones of one another, differing only by their rest masses; each subsequent generation is heavier than the previous. The interactions of these three generations via the fundamental forces are identical in the SM. These particles are listed in Table 1.1 along with their electromagnetic charge. The first generation,

Quarks			Leptons		
Name	Symbol	Q	Name	Symbol	Q
Up	$u$	+2/3	Electron neutrino	$\nu_e$	0
Down	$d$	-1/3	Electron	$e^-$	-1
Charm	$c$	+2/3	Muon neutrino	$\nu_\mu$	0
Strange	$s$	-1/3	Muon	$\mu^-$	-1
Top	$t$	+2/3	Tau neutrino	$\nu_\tau$	0
Bottom	$b$	-1/3	Tau	$\tau^-$	-1

Table 1.1: The fundamental fermions in the standard model. The electromagnetic charge,  $Q$ , is given in units of the absolute value of the electron charge:  $1.6 \times 10^{-19}$  C.

containing the  $u$ ,  $d$ ,  $e^-$  and  $\nu_e$  particles, is the least massive generation and most stable. The second and third generations exist only fleetingly, before decaying back down to the first generation. Therefore, all stable matter in the universe is comprised of first-generation particles. As detailed in Table 1.1, all fermions except neutrinos have a electromagnetic charge. Additionally, they have other quantum numbers that define their interaction with the other fundamental interactions.

## Antimatter

For each of the fundamental fermions in the standard model there exists a antimatter partner, as predicted by P. Dirac in 1928 [23]. The positive electron partner, called positron, was first discovered by C. Anderson in 1932 [24]. Antimatter particles only differ from their matter counterparts by their quantum numbers, such as charge or baryon and lepton number. The mass of a particles antimatter partner is identical. Matter and antimatter can recombine or *annihilate*, releasing the energy stored in their rest masses. Therefore, each of the fermions in Table 1 has a corresponding antimatter equivalent with the opposite electromagnetic charge.

## Fundamental interactions

The SM provides a mechanism for three fundamental forces: the electromagnetic, weak and strong forces. Each of these are mediated by force-carriers called gauge bosons. Details of these forces are listed in Table 1.2 along with gravity, the forth fundamental force.

Force	Boson	Symbol	Mass (GeV/c <sup>2</sup> )	Range (m)
Strong	gluon	$g$	0	$10^{-15}$
Electromagnetic	photon	$\gamma$	0	$\infty$
Weak	$W$ boson $Z$ boson	$W^\pm$ $Z$	$80.385 \pm 0.015$ $90.188 \pm 0.002$	$10^{-18}$
Gravity	<i>graviton?</i>		0	$\infty$

Table 1.2: The fundamental interactions in the standard model listed in order of strength.

The electromagnetic interaction is mediated by the massless photon. This affects all particles with a electromagnetic charge, including all fundamental fermions except neutrinos. As the mediator is massless the range of this force is infinite. The interaction is described fully by Quantum Electrodynamics (QED), an Abelian gauge field theory in which the photon doesn't undergo self-interactions as it is uncharged<sup>2</sup>.

The strong force is mediated by the massless gluon. This couples to colour charge, which comes in three variants; red, green and blue. The strong interaction is described by Quantum Chromodynamics (QCD), a non-Abelian gauge theory. This allows gluon self-coupling as the gluons carry colour charge themselves. This leads to complicated dynamics that limits the range of the strong force. Of the fundamental fermions, only the quarks have colour charge. Leptons don't interact strongly.

The weak force is mediated by the massive  $W^\pm$  and  $Z$  bosons. The range of this force is limited to the order of  $1/m_{W,Z}$  and appears much weaker than the electromagnetic force for interactions with energies much below  $\sim m_{W,Z}$ . The weak interaction is also described a non-Abelian gauge theory in which the gauge bosons can self-interact.

---

<sup>2</sup>Light-by-light scattering has been observed when mediated by quark loops [25]

## The origin of mass

The final particle of the SM is the Higgs boson, a key component to allow the fundamental particles to gain mass. Discovered in 2012 at the Large Hadron Collider [19, 20] it constitutes the final particle predicted to be part of the SM.

In the standard model all fermions would naively be expected to be massless, as the Lagrangian density terms that would allow them to gain a mass are excluded due to symmetry constraints. The exact symmetry can be broken by a mechanism proposed by R. Brout, F. Englert and P. Higgs [26, 27] in which a new scalar field is introduced called the Higgs field.

The masses of particles arise as a result of the self coupling of the Higgs field. The field potential contains quadratic and quartic terms in the field strength, allowing a situation where stable minima exist, displaced from the origin (which is an unstable maximum). The perturbations of this field occur around any one of the degenerate minima, requiring the symmetry of the system to be spontaneously broken. This leads the vacuum expectation value,  $v$ , of the field to be non-zero.

Gauge boson masses are a result of the kinetic energy of the Higgs field. The fermion masses arise from the Higgs-fermion field interactions directly. The Higgs boson itself gains mass from the Higgs field kinematics and potential.

The masses of the SM fermions as determined using experiment input are shown in Table 1.3. In the standard model neutrinos are massless particles, however they have

Particle	Mass
$u$	$2.2^{+0.6}_{-0.4} \text{ MeV}/c^2$
$d$	$4.7^{+0.5}_{-0.4} \text{ MeV}/c^2$
$e^-$	$0.5109989461 \pm 0.0000000031 \text{ MeV}/c^2$
$c$	$1280 \pm 30 \text{ MeV}/c^2$
$s$	$96^{+8}_{-4} \text{ MeV}/c^2$
$\mu^-$	$105.6583745 \pm 0.0000024 \text{ MeV}/c^2$
$t$	$173 \pm 0.6 \text{ GeV}/c^2$
$b$	$4.18^{+0.04}_{-0.03} \text{ GeV}/c^2$
$\tau^-$	$1.77686 \pm 0.00012 \text{ GeV}/c^2$

Table 1.3: The masses of the charged fundamental fermions, from Ref. [28].

been observed to oscillate between flavours, requiring that they are in fact massive. It is clear that the charged lepton masses can be determined much more precisely than the quarks. The quark masses can't be measured directly, instead they are determined indirectly using input from hadronic measurements.

Whilst very significant, the fermion masses themselves are not responsible for the majority of mass in the universe. The proton is a composite particle of two  $u$ -quarks and one  $d$ -quark with a mass of  $m_p = 938 \text{ MeV}/c^2$ . This is about 100 times larger than the sum of the constituent quark masses. The remainder of the mass originates from the binding dynamics of the strong interaction. The proton can be thought of as containing a sea of quarks and gluons that mediate the attraction between the valence quarks. The invariant mass of the whole system gives rise to the proton mass.

### **$CP$ violation**

The combined symmetries of charge conjugation (swapping particles for antiparticles) and parity inversion (mirroring space) are referred to as  $CP$ -symmetry. The SM contains mechanisms that allow this  $CP$ -symmetry to be violated in decays: processes can happen at different rates for matter and antimatter. The weak interaction introduces all of the  $CP$  violation observed in the SM. Both the quark and lepton sectors have the freedom to mix between different generations, introducing a complex weak phase that can cause asymmetries in interference between competing processes. In principle processes mediated by the strong interaction could exhibit  $CP$ -violation, however this has not been observed to occur.

#### **1.1.2 Parameters**

The SM contains 26 free parameters. The theory doesn't predict the value of these numbers, instead they must be worked out experimentally.

**Masses:** the masses of the 12 fundamental fermions (alternatively parametrised as the Higgs-fermion couplings).

**Couplings:** the three coupling strengths of the electromagnetic, strong and weak forces.

**Higgs potential:** two parameters determine the Higgs potential: the vacuum expectation value and the Higgs Boson mass.

**Mixing parameters:** these eight parameters allow the weak interaction eigenstates and mass eigenstates to differ in both the quark and lepton sector.

**Strong  $CP$  phase:** this could allow  $CP$  violation in the strong interaction, but is experimentally measured to be extremely small.

The relatively large number of free parameters in the SM may hint to this model only being a subset of a more complete theory. Trends are observed between the quantities, for example the masses of the fermions in each generation are each fairly similar in size. Some larger theory may be responsible for these trends.

### 1.1.3 Triumphs

The SM predicts a number of processes to an extraordinary accuracy.

#### Magnetic moments

The intrinsic magnetic moment for electrons was predicted by Dirac to be  $g = 2$ . In QED this quantity receives corrections as a result of higher order effects that make the value slightly larger than two. The electrons anomalous coupling  $a_e = (g - 2)/2$  has been measured [29] and predicted theoretically [30] to be

$$\begin{aligned} a_e(\text{theory}) &= (1159652182.031 \pm 0.015 \pm 0.015 \pm 0.720) \times 10^{-12} \\ a_e(\text{exp.}) &= (1159652180.73 \pm 0.28) \times 10^{-12}. \end{aligned} \tag{1.2}$$

The remarkable level of agreement demonstrates the predictive accuracy of the SM.

### Precision tests of electroweak theory

The SM makes predictions for the properties of the electroweak gauge bosons. Tests were performed on the  $Z$  boson measuring the total decay width, a quantity sensitive to the number of light neutrino species. Experiments conducted at the Large Electron Positron collider (LEP) and SLAC Linear Collider determined the number of species to be  $2.9840 \pm 0.0082$  [31], consistent with the SM picture. Additionally, these experiments used vast collections of  $Z$  bosons to measure branching fractions, polarisations and asymmetries to high precision.

#### 1.1.4 Shortfalls

In spite of the success of SM predictions a number of areas are not explained by the theory.

##### Baryon asymmetry

The universe is observed to be matter dominated. Macroscopic quantities of antimatter elsewhere in the universe would result in matter-antimatter boundaries that radiate energy from annihilation. This would lead to intense sources of high energy radiation visible to gamma ray telescopes; a phenomenon that has not been observed [32]. Three conditions that can lead to asymmetries between matter and antimatter were postulated by A.D. Sakharov in 1966 [33]. These require baryon number violation,  $C$ - and  $CP$ -symmetry violation and interactions out of thermal equilibrium. The SM provides all of these, however the source of  $CP$ -violation is far too small to account for the observed cosmic asymmetry. For processes at the electroweak energy scale, the quark masses (excluded the  $t$ -quark) are comparatively small, leading to a tiny level of  $CP$  violation. Sources beyond the scope of the SM are therefore expected to contribute to this asymmetry.

## Lepton flavour anomalies

Recent measurements at the LHC and  $B$ -factories hint towards anomalous decay rates in decays to different generations of leptons [34, 35]. In the standard model the couplings to the lepton flavours are equal, the only differences in decay rates arise as a result of the different masses.

## Neutrino masses

The neutrinos in the SM are massless. This has been demonstrated to be incorrect by the observation of neutrino flavour oscillations that require non-zero masses [36, 37]. Models that go beyond the SM explore the possibilities of neutrinos being Dirac fermions (similar to SM fermions) or Majorana fermions (their own antiparticles).

## Gravity

The SM doesn't incorporate a gauge field theory of gravity. The interaction can be hypothesised to be quantised by a massless spin-2 graviton. However, a complete integration of general relativity and the standard model would require a theory of quantum gravity.

## Dark matter and dark energy

Astrophysical evidence suggests that baryonic matter only makes up small fraction of the total matter content of the universe. The distribution of galactic rotation curves suggests a large halo of matter is present in the galaxy that only interacts gravitationally [38]. Additionally phenomena such as the bullet cluster and gravitational lensing point towards a higher density of matter than can be observed with electromagnetic radiation [39, 40].

The acceleration of universe's expansion is attributed to dark energy; a currently poorly understood type of energy distinct from dark matter and baryonic matter.

### Strong $CP$ violation

The string interaction  $CP$  violating phase is constrained to be extremely small from measurements of the neutrons electric dipole moment. The current limits constrain the strong  $CP$  violating phase to be below  $\mathcal{O}(10^{-10})$  [41]

### Hierarchy problem

The typical energies of different interactions in the SM are on vastly difference scales, for example the electroweak interaction scale is  $\mathcal{O}(10^3)$  GeV, whilst the GUT scale (the energy at which the electroweak and strong force could be unified) is of order  $\mathcal{O}(10^{16})$  GeV. Additionally, gravity is expected to be comparable at the Planck scale  $\mathcal{O}(10^{19})$  GeV. These different mass scales make it difficult to understand why the Higgs mass is so much less than these energies; quantum corrections would be expected to raise the mass to these scales. Additional theories are required to cancel these contributions.

## 1.2 Beyond the Standard Model

As the SM is deficient in it's description of the natural world, a large number of theories proposing extensions have been hypothesised. A small selection of Beyond the Standard Model (BSM) theories are summarised here.

### 1.2.1 Supersymmetry

The Supersummetry (SUSY) theory unifies fermions and bosons by proposing a symmetry that can interchange the two. This requires creating an additional set of particles for each SM fermion and boson; the supersymmetric partners. This symmetry is broken to allow the SM and SUSY partners to have different masses. This theory predicts dark matter candidates, fixes the hierarchy problem and unifies the gauge interactions. However, the broken symmetry results in a large number of free parameters, leading to a huge phase-space in which to search. So far no evidence of SUSY has been observed.

### **R-parity violating**

Within SUSY, the original SM particles and the newly introduced SUSY particles can be distinguished by a quantum number called *R*-parity, for which SM (SUSY) particles have eigenvalues +1 (−1). In SUSY models that conserved *R*-parity, the lightest SUSY particle must be stable, as it would not be able to decay to just SM particles. This results in a convenient dark matter candidate, and prevents rapid proton decay that could be observed if baryon and lepton number conservation can be violated.

Some models exist in which *R*-parity violation (RPV) occurs. If only one of baryon or lepton number is also violated then the rapid decay of the proton can be avoided.

### **Two Higgs doublets**

A number of extensions to the SM exist which include additional Higgs doublets. The SM contains a single Higgs doublet, corresponding to the scalar Higgs boson and the three Goldstone bosons [42, 43] that allow the  $W^\pm$  and  $Z$  gauge bosons to gain a longitudinal degree of freedom and hence mass. Introducing a second Higgs doublet results in four more physical states; this gives a total of three neutral Higgs bosons and two charged. SUSY requires two Higgs doublets with up-type and down-type fermions coupling to different doublets.

#### **1.2.2 Leptoquarks**

A number of theories result in the creation of *leptoquarks*: additional particles that facilitate quarks and leptons to be interchanged. In these theories the SM strong SU(3) and electroweak SU(2)  $\times$  U(1) symmetries are embedded in a larger symmetry, for example SU(5). The gauge bosons of this symmetry can couple to leptons and quarks. Leptoquarks can allow lepton flavour universality to be violated [44].

### 1.2.3 Extra dimensions

Attempts by T. Kaluza and O. Klein to unify the electromagnetic interaction and gravity by extending space-time into a fifth dimension [45, 46]. Many different models exist that propose the geometry of this additional dimension. These provide solutions to the hierarchy problem and propose additional particles that could be observed.

### 1.2.4 Axions

The problems surrounding the lack of strong  $CP$  violation can be circumvented by introducing a new field to the SM as proposed by R. Peccei and H. Quinn [47]. This results in a new light or massless gauge boson called the Axion, however this has not been observed.

# Chapter 2

## Theory

### Contents

---

<b>2.1</b>	<b>The weak force in <math>b</math>-hadron decays</b>	<b>14</b>
2.1.1	The weak force	15
2.1.2	Weak interactions of quarks	17
2.1.3	The CKM matrix	17
2.1.4	$b$ -hadron physics	19
2.1.5	QCD and hadronisation	19
<b>2.2</b>	<b>Annihilation topology decays</b>	<b>21</b>
2.2.1	Pure annihilation topology decays	22
<b>2.3</b>	<b>Rescattering</b>	<b>23</b>
<b>2.4</b>	<b>Theoretical predictions for the <math>B^+ \rightarrow D_s^+ \phi</math> decay</b>	<b>24</b>
2.4.1	Standard model predictions	24
2.4.2	BSM models and predictions	25
2.4.3	Previous measurements	26
<b>2.5</b>	<b>Theoretical predictions for the <math>B^+ \rightarrow D_s^+ K^+ K^-</math> decay</b>	<b>26</b>
2.5.1	Standard model predictions	28

---

This chapter contains an overview of the theory relevant to rare hadronic  $B$  meson decays.

### 2.1 The weak force in $b$ -hadron decays

Ground state  $b$ -hadrons can only decay via the weak interaction as the  $b$ -quark must change flavour to one of the first or second generation quarks. Exclusive decays of  $B$

mesons were first observed by the CLEO experiment in 1983 [48]. Although the weak interaction dictates the core of the process, the contributions from the strong force are unavoidable. In the SM quarks abide by *confinement*, preventing bare quarks from propagating unhindered. Instead quarks are only ever observed in bound states, or hadrons, with other quarks or anti-quarks. Therefore the observation of the weakly decaying  $b$ -quark is necessarily accompanied by the strong interaction governing the hadronisation of the initial and final state quarks.

### 2.1.1 The weak force

The weak interaction differs remarkably from the electromagnetic and strong interactions. It is the only of the three to be mediated by massive gauge bosons and uniquely violates parity; the action of mirroring space. Additionally, there are two types of gauge boson: the charged  $W^\pm$  bosons that result in *charged-current* interactions; and the neutral  $Z$  boson, responsible for *neutral-current* interactions.

Confusingly, the weak interaction is actually stronger than the electromagnetic force, however the large mass of the mediators suppresses the interaction at low energy. The *charged-current* weak interaction was originally described by Fermi using a four point coupling  $G_F$ . This works well for low energy interactions in which the momentum transfer is much less than the  $W^\pm$  boson mass,  $|q^2| \ll m_{W^\pm}^2$ . In the higher mass range this breaks down leading to unitarity violation in processes such as  $e^+e^- \rightarrow W^+W^-$ . This is resolved by the addition of the neutral  $Z$  boson and the *neutral-current* interactions that restore unitarity.

#### Parity violation

Parity is conserved in QCD and QED so it was naturally assumed to follow suit in the weak interaction. In 1957 it was shown by C.S. Wu and collaborators that parity was indeed violated in the beta decay of cobalt-60 [49]. The rate of electrons emitted during the decay were measured as a function of the polar angle. In the parity conservation scenario the same rates would have been expected at angles of  $\theta$

and  $180^\circ - \theta$  to an applied magnetic field; however electrons are preferably emitted against the field.

The weak interaction can be described as a V-A interaction (vector-axialvector), corresponding to a maximally parity violating interaction; the  $W^\pm$  boson only couples to left-handed chiral particle states. In the ultra-relativistic limit this would mean that the weak force only couples to left-handed helicity particles and right-handed helicity antiparticles. This is relaxed when accounting for the non-zero masses of fermions (when the helicity eigenstates are no longer equal to the chiral eigenstates), but leads to helicity suppression for certain processes, for example  $\pi^- \rightarrow e^-\bar{\nu}_e$  with respect to  $\pi^- \rightarrow \mu^-\bar{\nu}_\mu$ . The relative branching fractions for these two processes acts as strong experimental evidence for the V-A interaction.

### Electroweak unification

Although the weak and electromagnetic forces can be described by QED and the Fermi interaction at low energies, this becomes insufficient at high energies. Here, the weak and electromagnetic forces can be unified through the single Electroweak gauge theory with a  $SU(2)_L \times U(1)_Y$  symmetry.

The  $Z$  boson introduced to prevent unitarity violation is experimentally observed to couple to both left-handed and right-handed particles, in contrast to the  $W^\pm$  boson. This seemingly contradictory situation was resolved by Glashow, Salam and Weinberg [50–52] who developed a unification of the electromagnetic and weak interactions. This begins by replacing the physical electromagnetic  $U(1)_Q$  interaction with a similar  $U(1)_Y$  interaction that couples to weak hyper-charge,  $Y$ . This gauge symmetry requires a new gauge boson  $B$ . The weak  $SU(2)_L$  interaction generates three gauge bosons. The third of these, the neutral  $W^3$ , can mix with the new  $B$  gauge boson to generate the two physical bosons  $\gamma$  and  $Z$

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix}, \quad (2.1)$$

where  $\theta_W$  is the weak mixing angle. This allows the  $Z$  boson to couple to both left- and right-handed particles as it is a mixture of the parity-violating  $W^3$  boson and parity-conserving  $B$  boson.

### 2.1.2 Weak interactions of quarks

It is experimentally observed that the rates of kaon and pions decays do not proceed at the same rate; the decay rate of  $s$ -quark containing particles are suppressed,

$$\frac{\Gamma(K^+ \rightarrow \mu^+ \nu_\mu)}{\Gamma(\pi^+ \rightarrow \mu^+ \nu_\mu)} \sim 0.257. \quad (2.2)$$

It was observed by N. Cabibbo that in many quantum mechanical systems, different *good* quantum numbers are appropriate depending on the dynamics. He proposed that the eigenstates of the weak interaction may be different to the mass eigenstates that govern the propagation of particles [53]. The Cabibbo angle,  $\theta_c$ , was introduced to mix the mass eigenstates into the weakly interacting eigenstate,

$$d' = \cos \theta_c d + \sin \theta_c s. \quad (2.3)$$

However, this on its own was not enough to explain the lack of flavour-changing neutral currents such as  $K^0 \rightarrow \mu^+ \mu^-$ . This was resolved by S.L. Glashow, J. Iliopoulos, and L. Maiani with the introduction of the GIM mechanism [54]. This suppressed neutral current processes by proposing a fourth quark, the charm quark. This was paired with the strange quark and allowed Cabibbo's mixing matrix to be elegantly completed

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} = \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \quad (2.4)$$

In the SM the weak interaction facilitates the changing of quark flavour via the *charged-current* interaction  $W^\pm \rightarrow q\bar{q}'$  whilst the *neutral-current* interaction must conserve flavour,  $Z \rightarrow q\bar{q}$ .

### 2.1.3 The CKM matrix

Cabibbo's two-generation mixing matrix can be easily extended to the three-generation scenario in the SM. The extended matrix referred to at the Cabibbo-Koboyashi-Maskawa (CKM) matrix [55] is expressed generally as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.5)$$

where the transitions between quark flavours are determined by the complex quantities  $V_{qq'}$ . In the SM this matrix is required to be a unitary matrix obeying the condition

$$U_{\text{CKM}}^\dagger U_{\text{CKM}} = I. \quad (2.6)$$

Therefore, the extension of the mixing matrix not only allows all three generations to be considered, but also increases the number of terms that parametrise the mixing matrix from one to four. Of these four parameters, three can be represented as angles and one must be interpreted as a complex phase. Importantly, this phase allows  $CP$  violation to be present in the SM and was one of the main motivations of M. Koboyashi and T. Maskawa in proposing the matrix. The matrix can be represented in the Wolfenstein parametrisation using four parameters  $\lambda$ ,  $A$ ,  $\rho$  and  $\eta$  [56]. When represented with factors up to  $\mathcal{O}(\lambda^3)$ , the CKM matrix is given by

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.7)$$

This representation is helpful as it dictates that for  $CP$  violation to be present in the SM,  $\eta$  must be non-zero. Additionally, the contribution from the imaginary parts is predominately in  $V_{ub}$  and  $V_{td}$  (the terms  $V_{cd}$  and  $V_{ts}$  receive smaller contributions from  $\eta$  at  $\mathcal{O}(\lambda^5)$ ).

The magnitudes of the CKM matrix elements can be inferred from various processes and have been measured to have the following sizes [28]

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{pmatrix}. \quad (2.8)$$

The values of the diagonal elements are determined to be close to unity, with the elements getting smaller further away from the diagonal.

The first observation of  $CP$  violation occurred in 1964 by V. Fitch and J. Cronin [57]. A beam of long-lived neutral kaons was produced and the decay to  $2\pi$  was observed. Neutral kaons were known to decay to  $2\pi$  and  $3\pi$  final states. The  $2\pi$  state is a +1 eigenstate of the  $CP$  operation, whilst the  $3\pi$  is a -1 eigenstate. If the long-lived and short-lived neutral kaon eigenstates were also  $CP$  eigenstates then only the  $K_s^0 \rightarrow \pi\pi$  and  $K_L^0 \rightarrow \pi\pi\pi^0$  decays would be allowed. The observation of  $K_L^0 \rightarrow \pi^+\pi^-$  decays demonstrated that the physical kaon states did not conserve  $CP$  symmetry.

Mesons containing  $c$ -quarks have not been observed to violate  $CP$  asymmetry. The SM predictions for  $CP$  violation in the charm sector are very small as a result of the CKM elements that govern these processes.

#### 2.1.4 $b$ -hadron physics

Hadrons containing  $b$ -quarks are a rich environment to study CKM physics. There are a wide variety of topologies that allow different parameters to be measured. In the  $b$ -meson sector  $CP$  violation has been observed in decays of  $B^+$  [58, 59],  $B^0$  [60, 61],  $B_s^0$  [62] mesons. Neutral  $B$  mesons are observed to oscillate between the particle and antiparticle states, allowing the extraction of ratios of CKM matrix elements.

This thesis is concerned with hadronic two- and three-body decays of charged  $B$  mesons. The relevant topologies of hadronic two-body  $B$  meson decays are shown in Fig. 2.1. These different possibilities receive varying levels of suppression as a result of the topology. The quark accompanying the  $b$ -quark can act as a spectator (as in Figs. 2.1a, 2.1b, and 2.1c) or can be involved in the decay process (as in Figs. 2.1e, 2.1d, 2.1f). In situations in which more than one of these topologies can contribute, interference can occur between the competing amplitudes. If the processes have different weak phases, *i.e.* contain different CKM matrix elements, the interference can lead to  $CP$  violation.

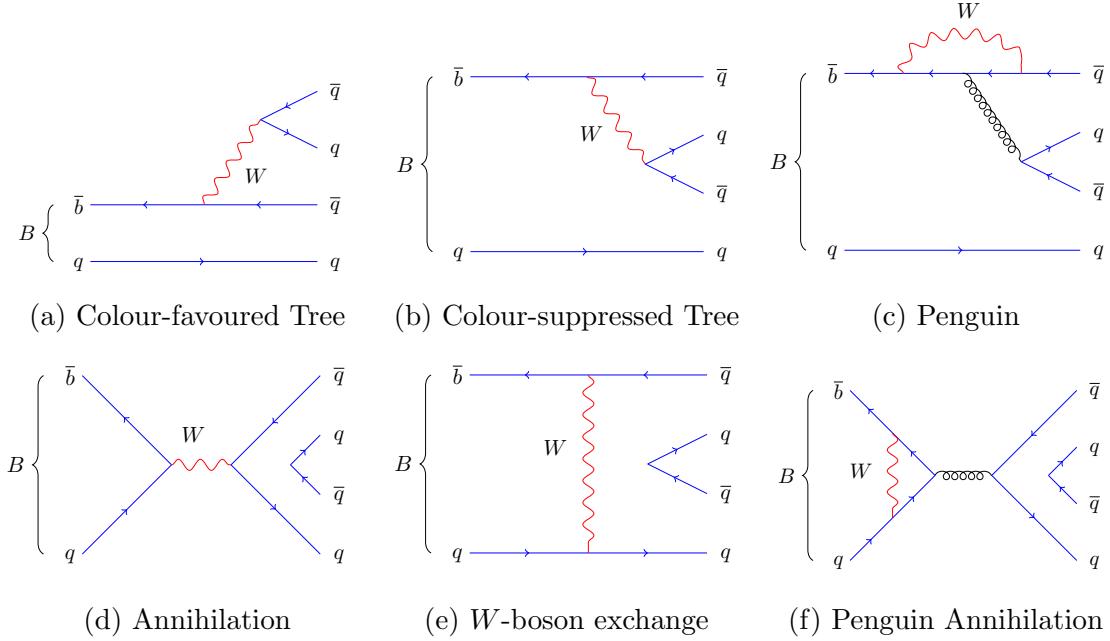


Figure 2.1: Two-body weak decay topologies of  $b$ -mesons. The  $W$ -boson exchange and penguin annihilation topologies only contribute to neutral  $B$  meson decays, whilst the annihilation topology only contributes to charged  $B$  mesons.

### 2.1.5 QCD and hadronisation

Although the core decay of ground state  $B^+$  mesons is mediated by the weak interaction, the prediction of decay rates requires input from the strong interaction. The prediction of processes involving QCD is difficult because quarks and gluons cannot propagate freely as a result of *confinement*. This means hadronisation will occur, leading to the physically observed bound states. It is still possible to make comparisons of experimental observations with QCD predictions by exploiting a number of features of the theory.

Unlike electromagnetism, at small length scales the strong interaction coupling is weak, whilst at larger length scales it becomes stronger. This is referred to as Asymptotic Freedom [63, 64] and allows high energy (small length scale) processes to be calculated perturbatively. Low energy interactions remain difficult as infinite numbers of terms would be required to compute exactly.

It is possible to impose an arbitrary energy scale into the calculations. General expressions for processes can then be related to physical observables. For example,

the fractional content of different species within hadrons can be determined as Parton Density Functions, which can then be evolved to work at any arbitrary scale via dedicated renormalisation group equations.

Factorisation can be exploited to split high energy perturbative QCD (pQCD) from the low energy processes below the QCD scale  $\mathcal{O}(200 \text{ MeV})$ . This allows the high energy components to be computed perturbatively, with the low energy components taken from other calculations. One method of computing the non-perturbative expressions is Lattice QCD. This involves quantising space-time using a finite grid. The field equations are evaluated at these discrete points. The spacing of the lattice can be changed such that the physical values can be extrapolated in the continuum limit. This is computationally expensive and requires the use of supercomputers.

## 2.2 Annihilation topology decays

Annihilation topology decays are a class of processes in which the initial state particles annihilate to form a virtual propagator (Fig. 2.1d). This mediator then decays to a (possibly different) set of particles. Perhaps the simplest  $B^+$  meson annihilation process is  $B^+ \rightarrow \ell^+ \nu_\ell$ , shown in Fig. 2.2.

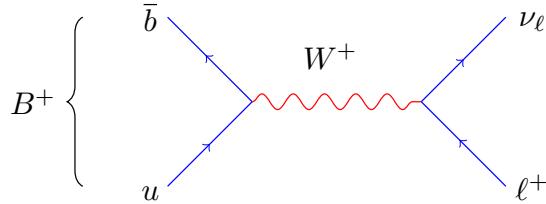


Figure 2.2: Annihilation topology process for  $B^+ \rightarrow \ell^+ \nu_\ell$  decays.

Decay Mode	SM prediction	Measurement
$B^+ \rightarrow e^+ \nu_e$	$(1.3 \pm 0.4) \times 10^{-11}$	$< 9.8 \times 10^{-7}$
$B^+ \rightarrow \mu^+ \nu_\mu$	$(5.6 \pm 0.4) \times 10^{-7}$	$< 1.0 \times 10^{-6}$
$B^+ \rightarrow \tau^+ \nu_\tau$	$(1.6 \pm 0.4) \times 10^{-4}$	$(1.09 \pm 0.24) \times 10^{-4}$

Table 2.1: Theoretical predictions and experimental measurements for  $B^+ \rightarrow \ell^+ \nu_\ell$  processes, from Refs. [65–67].

This purely leptonic  $B^+$  meson decay can proceed to either of the three generations with the branching fractions detailed in Table 2.1. The effect of helicity suppression in the final state is illustrated by the rapid decrease in predicted branching fraction as the mass charged lepton decreases. The neutrino in the final state makes this decay experimentally challenging to reconstruct at hadron colliders such as the LHC. The measurement of  $B^+ \rightarrow \tau^+ \nu_\tau$  was performed at the Belle experiment, a  $B$ -factory colliding  $e^+e^-$  pairs at the  $\Upsilon(4S)$  resonance. Hadronic annihilation decays are necessarily more complicated than these purely leptonic processes as the quark and antiquark must be further hadronised.

Decays that proceed via annihilation amplitudes are important in sectors other than  $B^+$  decays. In the neutral  $B$  meson sector, the decays  $B^0 \rightarrow K^+K^-$  and  $B_s^0 \rightarrow \pi^+\pi^-$  proceed via a combination of penguin annihilation and exchange diagrams [68]. These topologies are affected by QCD hadronic uncertainties therefore precise experimental measurements are required to help understand the QCD dynamics.

In the  $B_c^+$  sector, the recently observed  $B_c^+ \rightarrow D^0K^+$  decay has been shown to proceed via a combination of annihilation and penguin topologies [69]. This motivates the search for pure annihilation topology decays to understand the relative contributions from the annihilation and penguin processes.

### 2.2.1 Pure annihilation topology decays

Pure annihilation topology decays are an interesting subset of processes in which (at lowest order) only annihilation decay diagrams contribute. These are of particular interest because this allows the magnitudes of these processes to be isolated. Additionally, they provide a suitable forum in which to search for new physics; enhanced or decreased branching fractions and non-zero  $CP$  violations are symptomatic of multiple competing processes interfering with one another.

Typically, hadronic *pure* annihilation processes are characterised by a mutually exclusive set of quarks in the initial and final state. This implies that the initial state must have completely annihilated into a mediator for this process to occur.

In the  $B^+$  sector two decays are of particular interest;  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D^+ K^{*0}$ . A search at the LHCb experiment set a stringent limit on the branching fraction of the latter,  $\mathcal{B}(B^+ \rightarrow D^+ K^{*0}) < 6.1 \times 10^{-7}$  at 95% confidence limit [70]. A discussion of the  $B^+ \rightarrow D_s^+ \phi$  decay can be found in Sec. 2.4.

Pure annihilation decays are also of particular interest in the  $B_c^+$  sector. Charmless decays of  $B_c^+$  mesons can proceed exclusively via annihilation topologies [71]. These additionally benefit from the larger size of the CKM matrix element  $V_{cb}$  with respect to  $V_{ub}$ .

## 2.3 Rescattering

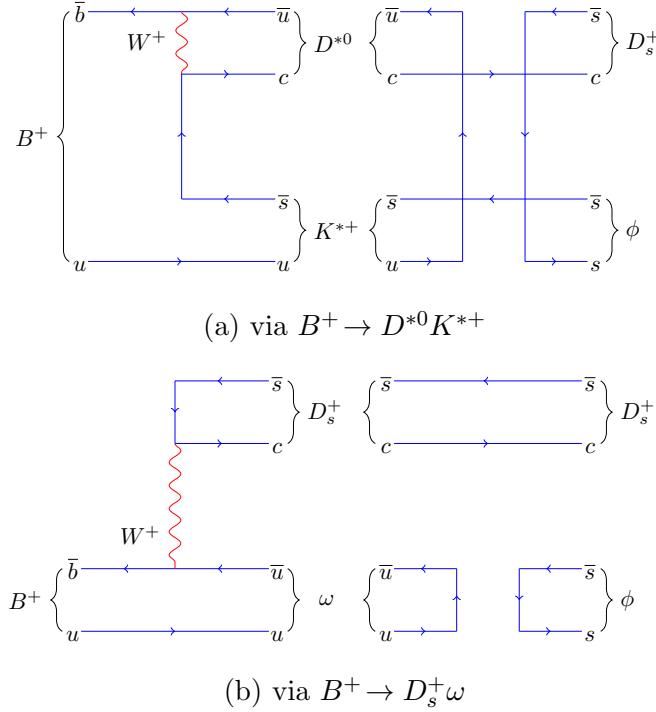
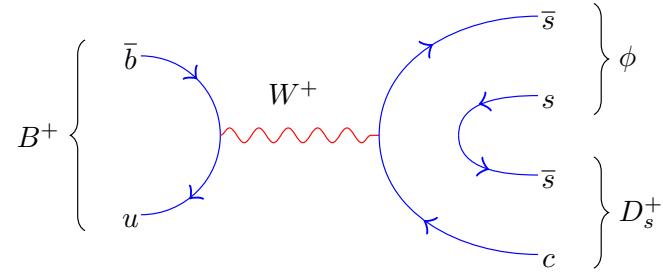
Rescattering allows  $B^+$  mesons to decay to the same final states as pure annihilation topologies via an intermediate process of some other topology. This can potential limit the sensitivity the annihilation decay if these processes happen at a significant rate. Intermediate processes that could contribute to the  $B^+ \rightarrow D_s^+ \phi$  decay are shown in Fig. 2.3 [72]. These rather complicated processes both involve tree-level  $\bar{b} \rightarrow \bar{u}$  transitions in which one of both of the products undergo a final state interactions.

In Fig. 2.3a, the path via the  $B^+ \rightarrow D^{*0} K^{*+}$  intermediate state is shown. The initial process is a weak colour-suppressed tree level decay, followed by a strong final state interaction that interchanges the quark flavours.

An alternative path via the state  $B^+ \rightarrow D_s^+ \omega$  is shown in Fig. 2.3b. This is a colour-favoured tree level process followed by the mixing of a  $\omega$  meson to a  $\phi$  meson. This  $\omega - \phi$  mixing is found to be small as a result of OZI suppression [73, 74].

## 2.4 Theoretical predictions for the $B^+ \rightarrow D_s^+ \phi$ decay

In the SM, the decay  $B^+ \rightarrow D_s^+ \phi$  proceeds dominantly via the pure annihilation diagram shown in Fig. 2.4. This suppressed topology requires the wave functions of the incoming quarks to overlap sufficiently to annihilate into a virtual  $W^+$  boson.

Figure 2.3: Rescattering contributions to  $B^+ \rightarrow D_s^+ \phi$  decays.Figure 2.4: The annihilation topology decay  $B^+ \rightarrow D_s^+ \phi$ .

The decay is further suppressed by the small magnitude of the CKM matrix element  $V_{ub}$  associated with the annihilation vertex.

#### 2.4.1 Standard model predictions

Predictions have been computed for the branching fraction of  $B^+ \rightarrow D_s^+ \phi$  in the SM. These are performed in a range of QCD frameworks, leading to predictions in the range  $(1 - 7) \times 10^{-7}$  as listed in Table 2.2. As detailed in Sec. 2.1.5 these predictions require additional input from Lattice QCD calculations [75–77].

Framework	$\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ Prediction	Ref.
Perturbative QCD	$1.3 \times 10^{-7}$	[78]
QCD Improved Factorization	$6.7 \times 10^{-7}$	[79]
QCD Factorization	$4.2 \times 10^{-7}$	[80]
Perturbative QCD	$3.0 \times 10^{-7}$	[81]

Table 2.2: Predictions for the branching fraction of  $B^+ \rightarrow D_s^+ \phi$  decays computed in different QCD frameworks within the SM.

Unlike many rare hadronic processes, the contribution from the rescattering processes described in Sec. 2.3 are expected to be small. The branching fraction of decays passing through the  $B^+ \rightarrow D_s^+ \omega$  intermediate state are predicted to have a branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)_{\omega-\phi} \sim 0.7 \times 10^{-7}$  [72].

### 2.4.2 BSM models and predictions

Extensions to the SM can enhance the branching fraction and/or produce large  $CP$  asymmetries in  $B^+ \rightarrow D_s^+ \phi$  decays by contributing additional decay amplitudes that interfere with the SM process. Predictions for the branching fraction and  $CP$  asymmetry  $A_{CP} = (N_{B^-} - N_{B^+})/(N_{B^-} + N_{B^+})$  have been computed in a range of BSM scenarios as listed in Table 2.3. In these models the decay is mediated by new parti-

BSM model	$\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ Prediction	$A_{CP}$	Ref.
Two Higgs Doublet Model	$8.0 \times 10^{-6}$	$\leq 59\%$	[79]
Minimal SUSY model with RPV	$3.04 \times 10^{-4}$	$\leq 14\%$	[79]
Scale invariant unparticles [82]	$0 - 20 \times 10^{-6}$	$\leq 50\%$	[80]

Table 2.3: Predictions for the branching fraction of  $B^+ \rightarrow D_s^+ \phi$  decays and  $CP$  asymmetry  $A_{CP}$  computed in various BSM scenarios.

cles, for example a charged Higgs boson, as shown in Fig. 2.5. Overviews of the *Two Higgs Doublet Model* and *Minimal SUSY model with RPV* can be found in Sec. 1.2. The two Higgs doublet model is implemented in the framework of SUSY.

### 2.4.3 Previous measurements

The LHCb experiment previously reported evidence for the decay  $B^+ \rightarrow D_s^+ \phi$  using  $pp$  collision data corresponding to an integrated luminosity of  $1 \text{ fb}^{-1}$  taken during

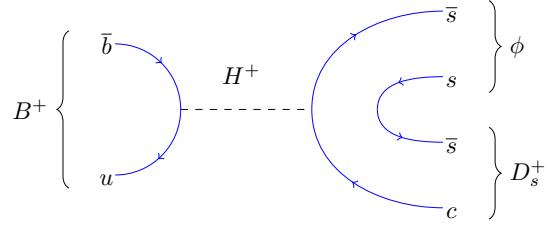


Figure 2.5: The annihilation topology decay  $B^+ \rightarrow D_s^+ \phi$  in a BSM scenario mediated by a charged Higgs boson.

2011, at a centre-of-mass energy of 7 TeV [83]. A total of  $6.7^{+4.5}_{-2.6}$  candidates was observed as shown in Fig. 2.6. The branching fraction was determined to be

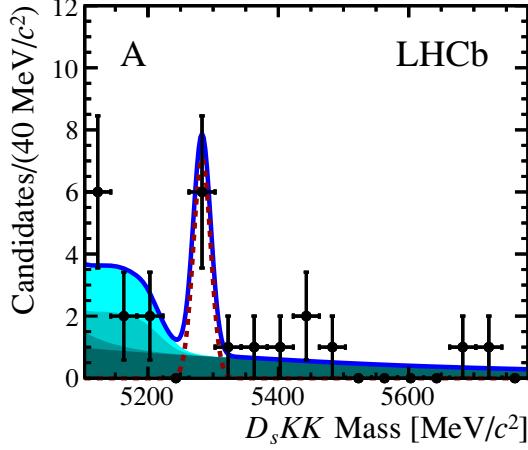


Figure 2.6: The invariant mass of  $B^+ \rightarrow D_s^+ \phi$  candidates previously observed by LHCb, from Ref. [83].

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) = (1.87^{+1.25}_{-0.73} \pm 0.19 \pm 0.32) \times 10^{-6}, \quad (2.9)$$

where the first uncertainty is statistical, the second is systematic and the third is due to the uncertainty on the branching fraction of the decay  $B^+ \rightarrow D_s^+ \bar{D}^0$ , which was used as normalisation. Given the large uncertainties on both the theoretical and experimental values, the previously measured value is consistent with the range of SM values given above.

## 2.5 Theoretical predictions for the $B^+ \rightarrow D_s^+ K^+ K^-$ decay

The decay  $B^+ \rightarrow D_s^+ K^+ K^-$  is an important component to be considered in the search for  $B^+ \rightarrow D_s^+ \phi$  as the  $\phi$  meson is reconstructed decaying to  $K^+ K^-$ . The decay  $B^+ \rightarrow D_s^+ K^+ K^-$  is mediated by a  $\bar{b} \rightarrow \bar{u}$  transition shown in Fig. 2.7 and is therefore suppressed in the SM due to the small size of the CKM matrix element  $V_{ub}$ .

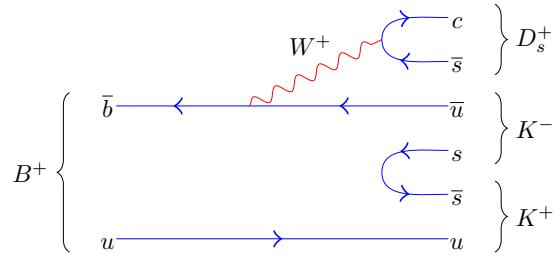


Figure 2.7: The decay process  $B^+ \rightarrow D_s^+ K^+ K^-$ .

This three-body decay differs from the 2-body topologies already considered. In a  $1 \rightarrow 2$  process, the momentum of the decay products in the rest frame of the initial particle is determined exactly. In a  $1 \rightarrow 3$  process this is not the case as different fractions of the momentum can be distributed among the decay products. This leads to two degrees of freedom that characterise the phase-space of three-body decays. This phase space is represented by a Dalitz plot: a two-dimension plot defined by the squared invariant masses of pairs of particles as devised by R. Dalitz [84] (Fig. 2.8). The decay  $B^+ \rightarrow D_s^+ K^+ K^-$  can therefore be characterised by the two-dimensional space defined by  $m^2(D_s^+ K^-)$  and  $m^2(K^+ K^-)$ . Resonances could contribute to  $B^+ \rightarrow D_s^+ K^+ K^-$  decays, for example as  $B^+ \rightarrow D_s^+(X \rightarrow K^+ K^-)$  or  $B^+ \rightarrow K^+(Y \rightarrow D_s^+ K^-)$ . The distribution of decays in the Dalitz plot are indicative of the spin and mass of contributing resonances.

### 2.5.1 Standard model predictions

The branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  is currently not measured. The topologically similar decay  $B^+ \rightarrow D_s^+ \pi^0$  proceeds via a colour-favoured tree level decay of

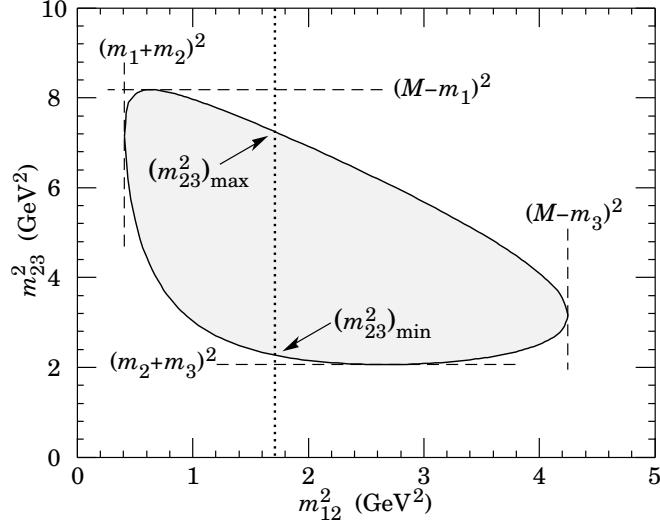


Figure 2.8: An example of a Dalitz plot for a three body final state, from Ref. [28].

the form shown in Fig. 2.7, but without an  $s\bar{s}$  pair produced. The branching fraction for this decay is measured to be  $\mathcal{B}(B^+ \rightarrow D_s^+ \pi^0) = (1.5 \pm 0.5) \times 10^{-5}$  [85].

# Chapter 3

## The LHCb experiment

### Contents

---

<b>3.1 CERN and the LHC . . . . .</b>	<b>30</b>
3.1.1 The accelerator complex . . . . .	31
3.1.2 Beam conditions . . . . .	33
<b>3.2 The LHCb detector . . . . .</b>	<b>34</b>
3.2.1 Magnet . . . . .	36
3.2.2 Vertex Locator . . . . .	37
3.2.3 Silicon Tracker . . . . .	41
3.2.4 Outer Tracker . . . . .	43
3.2.5 Ring imaging Cherenkov detectors . . . . .	45
3.2.6 Calorimeters . . . . .	49
3.2.7 Muon system . . . . .	52
3.2.8 Trigger . . . . .	54
3.2.9 Reconstruction . . . . .	57
<b>3.3 Luminosity determination . . . . .</b>	<b>59</b>
3.3.1 Van der Meer scans . . . . .	60
3.3.2 Beam gas imaging . . . . .	61

---

This chapter describes the LHCb experiment; including both the accelerator complex responsible for providing  $pp$  collisions and the LHCb detector itself. The LHCb experiment is a collaboration of around 800 scientists from 72 different institutions in 16 countries. The primary physics goals are to look for indirect signs of New Physics by searching for sources of  $CP$  violation and searching for rare decays of  $b$ -

and  $c$ -hadrons. The experiment is situated on the Large Hadron Collider at CERN, Geneva.

### 3.1 CERN and the LHC

In the aftermath of the Second World War a number of eminent scientists proposed the creation of a collaborative European laboratory dedicated to the study of atomic physics. With this, the ‘Conseil Européen pour la Recherche Nucléaire’ was born; a provisional council set up in 1952 to oversee the laboratory’s creation. The organisation as it is today was established in 1954, renamed the ‘Organisation Européenne pour la Recherche Nucléaire’, although the acronym CERN remained. The purpose of CERN was clear; the convention dictates that the organisation ‘*shall provide for collaboration among European States in nuclear research of a pure scientific and fundamental character*’. Since then CERN has been a hub of fundamental nuclear and particle physics research, producing a number of Nobel Prize-winning discoveries, alongside technological developments benefiting society as a whole.

The convention stipulates that the organisation ‘*shall have no concern with work for military requirements*’ and requires ‘*the results of its experimental and theoretical work shall be published or otherwise made generally available*’. The choice of the laboratories location followed a similar set of values, picking Geneva, Switzerland owing both to the central European location and neutrality of the host state.

Perhaps currently the most well known accelerator in the complex, CERN is home to the Large Hadron Collider (LHC). Two beams of hadrons circulate in opposite directions around 27 km rings, colliding at four interaction points. The beam pipes and experimental halls are buried deep underground, providing shielding from cosmic radiation and reducing the cost of acquiring large areas of land. The tunnels traverse the Franco-Swiss border at a depth that varies between 50–175 m at the lowest and highest points respectively.

The tunnel and caverns occupied by the LHC pre-date the current accelerators and experiments. There were originally constructed for the Large Electron Positron

Collider (LEP). This machine began operation in 1989 and collided electrons and positrons at a collision energy of  $\sqrt{s} = 209$  GeV.

### 3.1.1 The accelerator complex

The LHC is only one of a vast collection of accelerators at CERN, albeit the largest. The hadrons collided in the LHC travel sequentially through a number of different machines, boosting their energy in each. The full complex is shown in Fig. 3.2 along with a legend detailing the types of particles considered. The protons begin life in a hydrogen gas canister. The gas is ionised and accelerated in a linear accelerator, LINAC2, to an energy of 50 MeV. These then pass into the Proton Synchrotron Booster, raising the energy further from 50 MeV to 1.4 GeV.

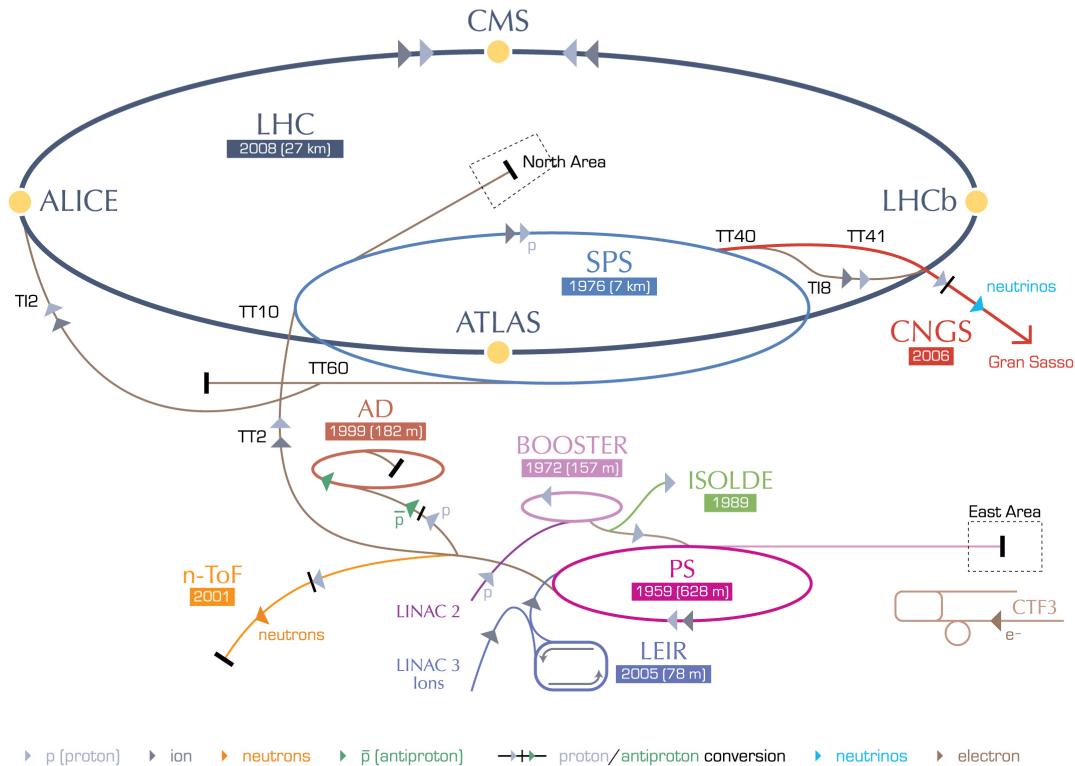


Figure 3.1: The accelerator complex at CERN.

In addition to protons, the complex can accelerate other ions including lead, and more recently, xenon. These start in a dedicated linear accelerator, LINAC3, where the electrons are stripped off the ions to produce bare nuclei before being injected

into the Low Energy Ion Ring (LEIR). This ring raises the ion's energy from 4.2 MeV to 72 MeV.

Either protons or ions can then be injected in the Proton Synchrotron (PS), CERN's first synchrotron. Once the world's highest energy particle accelerator, this now accelerates the particles up to 25 GeV ready to be injected into the Super Proton Synchrotron (SPS). The SPS is the second largest accelerator at CERN, with a circumference of 7 km. It was switched on in 1976 and resulted in the notable discoveries of the  $W$  and  $Z$  bosons. Here, the particles are finally accelerated up to an energy of 450 GeV, ready for injection into the LHC. The particles are transferred from the SPS to the LHC via two lines, one for each of the LHC beam directions. The two beams, referred to as Beam 1 and Beam 2, travel clockwise and anti-clockwise respectively when viewed from above. Beam 2 is injected just before the LHCb detector.

The particles travel through the accelerator complex in groups called bunches. These typically contain around  $10^{11}$  protons and are necessary as a result of the Radio-frequency (RF) cavities that accelerate the particles. The bunches are sequentially transferred between the accelerators in collections called trains. The bunch-trains are transferred from the SPS to the LHC in a series of injections. Although the LHC has a capacity for 3465 bunches, the nominal operating conditions has only 2808 of these spaces filled. These extra gaps allow enough room to safely divert the direction of the beam, without unintentionally damaging the instrumentation around the beam-pipe. The turn-on time for the magnets responsible for dumping the beam in emergencies is longer than the nominal 25 ns bunch spacing.

Once the LHC is at capacity the beams are accelerated from the injection energy of 450 GeV to the nominal energy of up to 7 TeV. This is achieved using eight RF cavities per beam to increase the beam energies and 1232 superconducting dipole magnets that bend the trajectory of the charged particles, ensuring they follow the paths dictated by the circular tunnels. The 15 m long dipole magnets contain superconducting niobium-titanium cables, kept cool at 1.9 K using super-fluid helium. The 11,850 A current flowing through the magnets generates a large magnetic field of 8.33 T. The current

is gradually and synchronously ramped as the particles are accelerated, such that the same curvature trajectory is maintained.

The LHC contains a wealth of other magnets, another 8,361 in addition to the dipole magnets, that optimise the orbit of the beams. These ensure the beams remain tightly packed together, as well as correcting many higher order perturbations.

During the injection and acceleration phase no collisions take place. Once the beams are at the nominal energy and profile distribution have been optimised, the two beams are brought together at four interaction points around the ring. Magnets are used to both redirect the beams such that they crossover and to squeeze the beams to a smaller cross-section at the collision points, increasing the likelihood of collisions. Once the beams overlap and the particles collide the experiments located around the four interaction points can begin to record data. The collisions typically continue for a duration of 10–20 hr (referred to as a fill). The number of particles in the bunches reduce, or *burn-off*, as they collide with one another and the beam-pipes. A fill ends when the beams are intentionally dumped, when the number of protons in each bunch gets too low, or unintentionally dumped, to ensure the safety of the machine. The beam dumps are large graphite blocks surrounded by concrete shielding, designed to absorb the energy of the beams.

A total of seven experiments are located at the LHC, situated in four caverns. The four largest are ATLAS, CMS, LHCb and ALICE. The TOTEM, LHCf and MoDEL experiments are located in the same caverns as CMS, ATLAS and LHCb respectively.

### 3.1.2 Beam conditions

The beam conditions at the LHC have varied between the different years of data taking included in this thesis. A summary of some key parameters are detailed in Table 3.1. This also includes the nominal conditions for the LHC, for which it was designed.

During a normal fill the luminosity delivered to the large experiments slowly decreases as the protons burn off. For ATLAS and CMS the average number of collisions per bunch crossing (known as pile up) is high: around an average of 27 in 2016. The

Condition	2011	2012	2015	2016	Nominal
$\sqrt{s}$ ( TeV)	7	8	13	13	14
Peak luminosity ( $10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ )	3.5	7.6	4.8	14	12
Bunches	1380	1380	1368	2736	2808
Bunch spacing ( ns)	50	50	25	25	25
Average $N_{\text{collisions}}$ , $\mu$	1.7	1.7	1.1	1.1	-
Integrated luminosity ( $\text{fb}^{-1}$ )	1.0	2.0	0.3	1.5	-

Table 3.1: Beam conditions in the different datasets used in this thesis. The last two rows are specific to LHCb.

LHCb experiment chooses to level the luminosity to a lower value than the maximum achievable. This means the average number of visible collisions per bunch crossing,  $\mu$ , is much lower, between 1.1 and 1.7 depending on the year. The same luminosity is maintained for the majority of a fill via a process known as luminosity levelling. At the beginning of the fill the two beams are vertically separated from one another such that the overlap between them is not maximised. As the fill progresses the two beams are brought closer together to compensate for the burn off of the bunches. This way a constant instantaneous luminosity of between  $(2\text{--}4) \times 10^{32} \text{ cm}^2 \text{ s}^{-1}$  can be maintained throughout the given fill.

In addition to the luminosity levelling, the two beams collide at an angle in the horizontal plane. This helps to reduce unwanted collisions between the bunches either side of the intended collision bunches. The angle of collision depends on the polarity of the LHCb spectrometer magnet as the dipole field affects beam-steering fields around the interaction point.

## 3.2 The LHCb detector

This section provides an overview of the experimental apparatus used to obtain the data analysed in this thesis. The LHCb detector is comprised of distinct sub-detectors, each with a dedicated purpose. These help to characterise the sub-atomic particles created in the proton-proton collisions, and enable measurements of their kinematics, trajectories and species. This overview includes a description of the sub-detector's principles, components and performance.

The LHCb detector is found at Point 8 of the LHC ring, in a cavern originally built for the DELPHI detector during the LEP era. A schematic representing the key components of the LHCb detector is shown in Fig. 3.2. This figure displays the axes convention adopted by LHCb, and used henceforth. The horizontal axis is labelled the  $z$ -axis and is parallel to the direction of the beams. The figure's vertical axis is the  $y$ -axis, increasing as one moves from the cavern up to ground level. The  $x$ -axis is in the dimension perpendicular to the plane of the figure and increases as one moves towards the centre of the LHC ring. The counter-rotating beams of protons are collided at the far left of this figure, at the origins of the  $y$ - and  $z$ -axes.

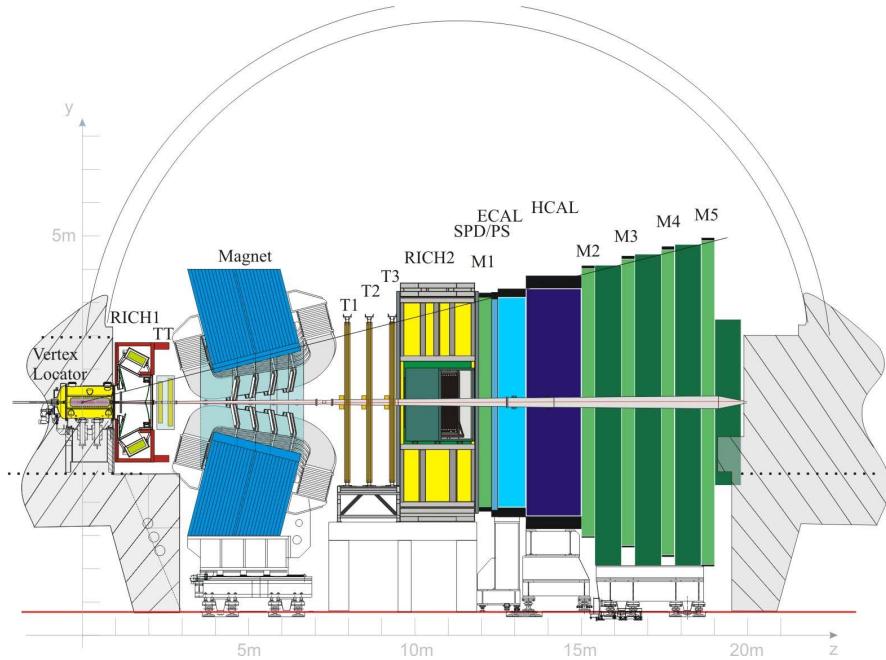


Figure 3.2: Schematic of the LHCb detector, from Ref. [86].

The LHCb detector covers only a small fraction of the area around the collision region; the acceptance covers particles with a pseudo-rapidity in the range  $1.8 < \eta < 4.9$  on one side of the interaction point. However, this configuration is well suited to the reconstruction of particles containing  $b$  or  $\bar{b}$  quarks. The angular distribution of  $b\bar{b}$  pairs is shown in Fig. 3.3, along with the acceptance of LHCb in red. The distribution is highly peaked towards the forward and backward regions as  $b$ -hadrons

tend to receive large boosts at this collision energy. Approximately 25% of decays have both  $b$  and  $\bar{b}$  quarks produced inside the acceptance.

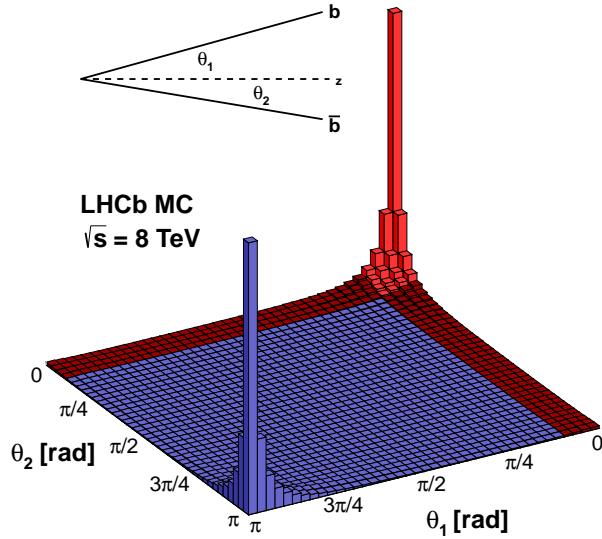


Figure 3.3: The angular distribution of  $b\bar{b}$  pair production in simulated  $pp$  collisions at a centre-of-mass energy of 8 TeV. The acceptance of LHCb is highlighted in red.

Running along the centre of the detector is the LHCb beampipe. The primary role is to separate the inner vacuum chamber from the rest of the cavern, allowing the beams to proceed unimpeded by the air. The majority of the beampipe is made of beryllium, with smaller sections made out of aluminium alloys or stainless steel. Although beryllium is a highly toxic and fragile material it has a long radiation length, allowing the incident particles to traverse the pipe walls with minimal interactions.

### 3.2.1 Magnet

The LHCb detector contains a warm dipole magnet that bends the trajectories of charged particles, allowing measurements of the particles' momentum. The magnet has two saddle-shaped coils inside a square yoke that generate an integrated magnetic field of 4 Tm. The magnetic field is aligned along the  $y$ -axis, bending the charged particles in the horizontal plane. The polarity of the magnet is routinely switched

during data taking. This helps to understand and cancel systematic effects that may affect measurements of  $CP$  asymmetries. The two magnet polarities are referred to *MagDown* and *MagUp*, corresponding to a field in the negative and positive  $y$ -axis direction respectively.

A schematic of the magnet is shown in Fig. 3.4, along with the strength of the magnetic field as a function of the  $z$ -axis position. Both magnet polarities are represented in this figure.

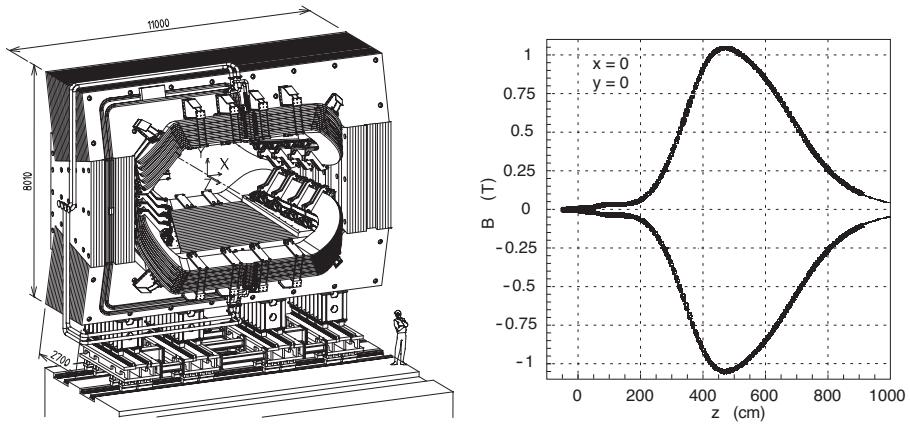


Figure 3.4: Schematic of the LHCb warm dipole magnet (left) and the magnetic field strength along the  $z$ -axis (right) from Ref. [86].

### 3.2.2 Vertex Locator

The first sub-detector to make measurements of the particles produced in proton-proton collisions is the Vertex Locator (VELO) encompassing the collision region. This sub-detector makes precise measurements of the track positions of charged particles as they emanate out of the collisions. A high level of precision is required to identify the secondary vertices characteristic of  $b$ - and  $c$ -hadron decays. These secondary vertices are typically displaced from the interaction position as a result of the long lifetimes associated to these heavy-flavour hadrons. This is achieved by measuring the track coordinates using silicon strip sensors placed as close to the LHC beam as safety allows. The VELO sensors are semicircular devices placed on either side of the beam. To allow the sensors to instrument the innermost region around the

interaction point the two halves of the VELO can move horizontally in and out. During normal data taking this allows the sensors to be 7 mm away from the interaction point yet still be a safe distance away from the beams at injection.

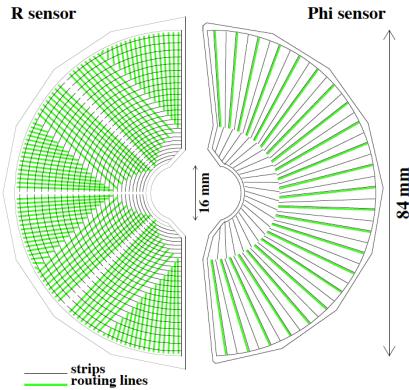


Figure 3.5: Schematic of an  $r$ - and  $\phi$ -sensor in the VELO sub-detector, from Ref. [87].

The sensors are grouped into pairs, called modules. Each of the 21 modules on each side contains sensors for measuring the radial and azimuthal coordinates of the tracks, referred to as  $r$ - and  $\phi$ -sensors respectively. A schematic of the two sensor types is shown in Fig. 3.5, illustrating the silicon strips and readout channels.

The VELO modules are arranged to ensure coverage of particles emerging in forward region at angles of 15–300 mrad with respect to the beam-pipe. This arrangement allows tracks within this acceptance to interact in at least three sensors as shown in Fig. 3.6. The modules extend both forward and backwards of the interaction region. Although momentum measurements are not possible for backward tracks, the vertexing of the primary interaction can benefit from this extra information. In the far backward region there are two additional modules, measuring only the radial coordinate. These help to identify pile-up events in which there is more than one primary vertex.

The VELO sub-detector is constructed to operate in the unique environment close to the LHC beams.

**Radiation resistance:** the VELO modules are subjected to extreme and varying amounts of radiation. The detector is designed to stand three years of

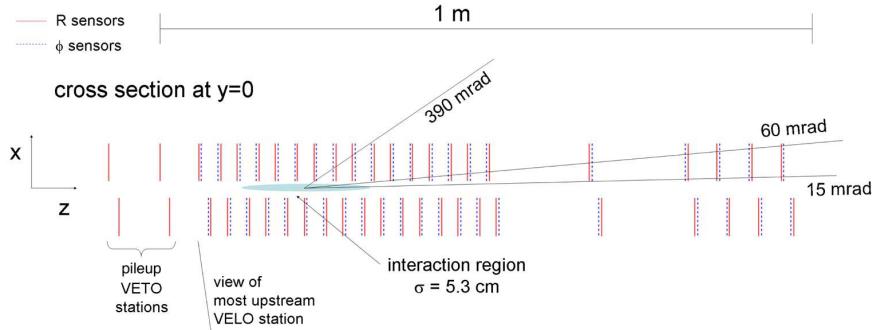


Figure 3.6: Schematic of the sensor layout in the VELO sub-detector, from Ref. [87].

nominal LHC running, using a radiation tolerant semiconductor construction. Additionally, the VELO modules are cooled to remove the heat created from the interactions. The sensors are maintained at a temperature between -10 and 0°C with 24 W of heat removed from each sensor.

**Radio-frequency (RF) pick-up protection:** the electromagnetic fields generated by the LHC beams could cause interference in the VELO detectors electronics. Therefore, a shield is used between the modules and the beam-pipe referred to as the RF-foil. This 0.3 mm thick foil separates the VELO and beam-pipe vacuums, providing additional protection to the conditions of the beams from the VELO detector itself.

The hits arising from particle interactions in the VELO sensors are extracted from the readout channels using custom analogue chips. The signals are digitised and combined into clusters in readout boards called TELL1 boards [88],

The performance of the VELO sub-detector can be quantified using different metrics. Two of particular interest are the track finding efficiency and the resolution of the impact parameter of a track. The track finding efficiency determines the likelihood that a track will be reconstructed and is of particular importance for high multiplicity final states including those presented in this thesis. The average track finding efficiency in data taken during 2011 is found to be around 98% as shown in Fig. 3.7.

The impact parameter (IP) represents the distance between a track and a primary vertex at the point of closest approach. Long-lived particles, such as  $B^+$  mesons, decay at a secondary vertex displaced from the primary collision vertex. Therefore, the decay products of these particles have a larger IP than the tracks originating at the primary vertex. A precise measurement of the IP of a track allows the  $B^+$  mesons and various background processes to be differentiated. The distribution of the  $x$ -component of the IP is shown as a function of the inverse transverse momentum ( $1/p_T$ ) in Fig. 3.7.

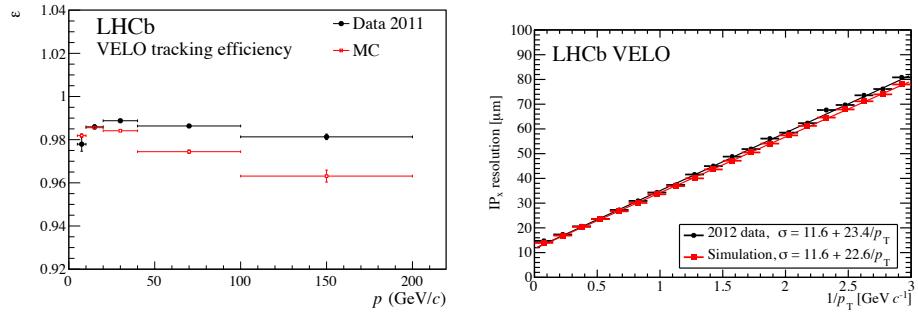


Figure 3.7: The tracking efficiency (left) and  $\text{IP}_x$  resolution (right) in simulation and data, from Ref. [87].

The VELO is situated in a harsh environment, therefore the components gradually degrade as exposed to more and more radiation. The damage to the silicon sensors is monitored over time to ensure the sub-detector is operational and to estimate the remaining lifetime of the sensors. In between LHC fills current-voltage (IV) scans are performed on the silicon semiconductors, measuring the current leaking as a function of the bias voltage. The evolution of this current in Run I is shown in Fig. 3.8, along with the delivered luminosity. These measurements help to predict what bias voltages will be required by the end of Run II.

A comparison of the IP resolution determined using 2012, 2015 and 2016 data is also shown in Fig. 3.8. Although the resolution is broadly similar between the three years, a slight increase in the resolution is present between Run I and Run II.

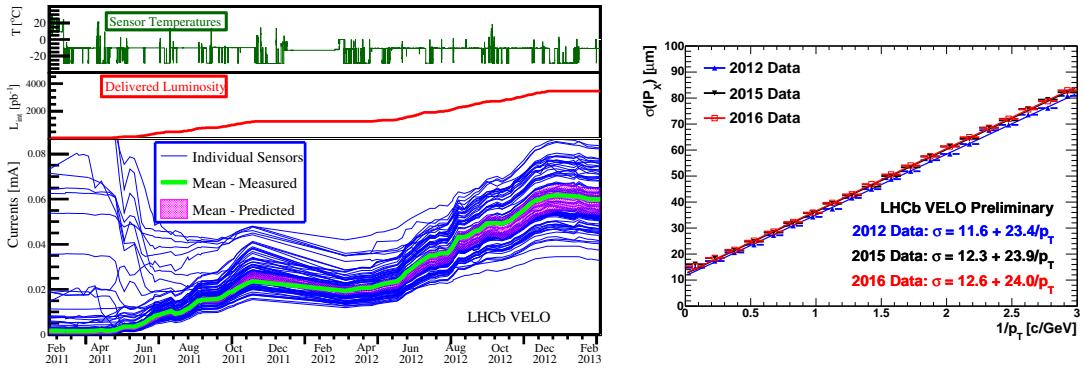


Figure 3.8: The VELO silicon sensor’s leakage current as a function of time through out Run I, from Ref. [89] (left) and the VELO IP resolution in 2012, 2015 and 2016 data (right) . A small decrease in performance is observed between Run I and Run II.

### 3.2.3 Silicon Tracker

In addition to the VELO, there are two more sub-detectors that utilise silicon sensors in order to determine tracking information. These are collectively referred to as the Silicon Tracker (ST), which is made up of two trackers; the Tracker Turicensis (TT) and Inner Tracker (IT). The TT is located before the dipole magnet, whereas the IT is positioned after, as show in Fig. 3.2. Although these two detectors are spatially separated, their common silicon mircostrip sensors and electronics warrant considering them together.

The silicon sensors are made up of single-sided  $p^+$ -on- $n$  sensors. The hits are read out via chips at the end of each module. These chips are the same custom analogue chips used in the VELO sub-detector. The signals pass into digitisers and then through optical fibres into TELL1 boards that perform clustering algorithms. Both sub-detectors are cooled to 5°C and their sealed containers flushed with nitrogen gas to prevent condensation.

#### Tracker Turicensis

The TT is positioned before the dipole magnet and covers the entire LHCb acceptance, standing 130 cm tall and 150 cm wide. It is made up of four layers orientated at angles to one another. The first and fourth layers are parallel, with the second and third at

angles  $-5^\circ$  and  $+5^\circ$  to these respectively. The first and second are separated from the third and fourth by 27 cm along the beam axis. The layout of the silicon modules in the TT is shown in Fig. 3.9. The sensors are grouped into *half-modules* that span half of the vertical height of the sub-detector. These are made up of seven silicon sensors and a readout hybrid at the outermost end. The readout electronics are positioned outside of the LHCb acceptance, limiting the amount of multiple scattering due to interactions with the detector material.

The *half-modules* are arranged to prevent any gaps in the instrumentation. The adjacent *half-modules* are offset by 1 cm along the beam axis, allowing the modules to overlap by a few millimetres in the  $x$ -axis.

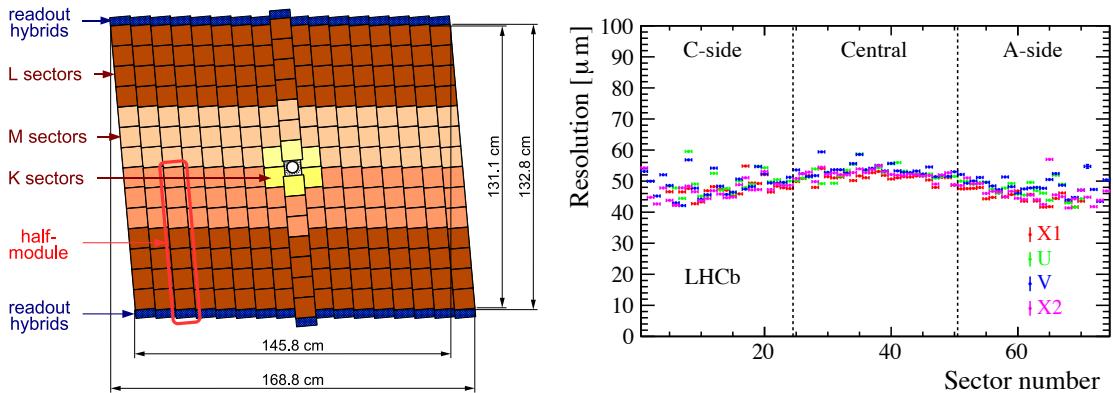


Figure 3.9: Schematic of the TT sub-detector, from Ref. [86] (left) and the TT resolution by sector number, from Ref. [90].

The resolution achieved by the TT is shown in Fig. 3.9, split into different sectors and shown separately for the four layers (labelled X1, U, V and X2). The resolution varies between 40–60  $\mu\text{m}$  and is consistent between the different layers.

### Inner Tracker

The IT is the second silicon detector making up the ST. It is located after the dipole magnet in three tracking stations. As the name implies, it covers only the inner region of the acceptance, measuring 140 cm wide and 40 cm tall, where the particle flux is highest. The rest of the area is covered by the much larger Outer Tracker. Similar to

the TT, the IT is made up of four layers positioned at slight angles to one another. However, as shown in Fig. 3.2, there are three separate IT stations, each containing four layers. These stations are constructed as of four boxes distributed around the beam-pipe in a cross shape, as shown in Fig. 3.10.

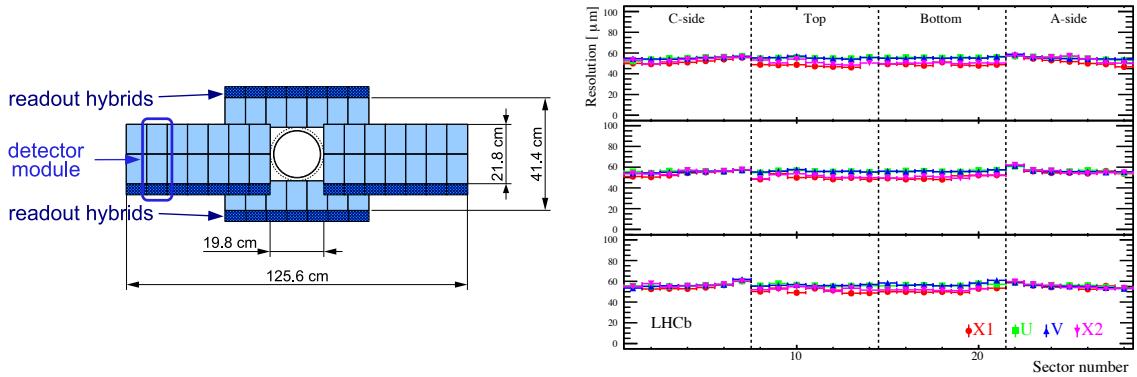


Figure 3.10: Schematic of the IT sub-detector, from Ref. [86] (left) and the IT resolution by sector number, from Ref. [90].

The detector modules consist of either one or two silicon sensors and a readout chip. These are also offset along the beam axis to allow the modules to overlap slightly in the  $x$ -axis.

The resolution of the IT is shown in Fig. 3.10. The resolution is between 40–60  $\mu\text{m}$  and varies slightly between the four different layers.

### 3.2.4 Outer Tracker

In contrast to the the silicon-based sub-detectors already described, the Outer Tracker (OT) is a straw tube tracker filled with a gaseous mixture of argon, carbon dioxide and oxygen. The 4.9 mm diameter straws are 2.4 m in length and arranged in double layers as shown in Fig. 3.11. As with the IT, the OT is split into three stations. Each of these stations similarly has four layers arranged at angles to one another ( $0^\circ, -5^\circ, +5^\circ, 0^\circ$ ). The arrangement of the stations and layers are also shown in Fig. 3.11. The inner region occupied by the IT is visible for scale.

The straws comprising the tracker are composed of an inner wire surrounded by a conducting foil, with a potential difference maintained between the two. When a

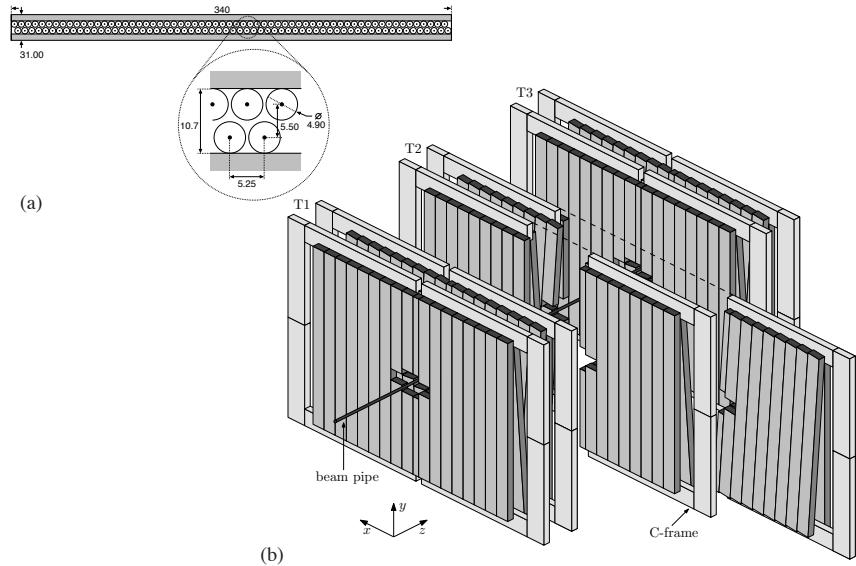


Figure 3.11: Schematic of the OT sub-detector, from Ref. [91].

charged particle passes through the straw the gas becomes ionised. The electrons and gas ions are attracted to the cathode foil and anode wire, resulting in a flow of current. In the OT the inner wire is made of 25  $\mu\text{m}$  thick gold plated tungsten and maintained at a voltage of +1550 V. The surrounding foil is made up of three layers; a 40  $\mu\text{m}$  thick layer of conducting carbon-doped polyimide film; a 12.5  $\mu\text{m}$  layer of insulating polyimide; and 12.5  $\mu\text{m}$  of aluminium.

The double layers of straw tubes as shown in Fig. 3.11 make up a module. Vertically, the modules are split in two, with staggered gaps between the two individual straw layers to avoid an uninstrumented region. Each of the modules are read out from the outermost end. The readout boards contain a number of circuits to process the signals and provide the potential difference to the straw wires. The analogue signals pass through an amplification circuit before being processed and cleaned. These then pass into a digitiser creating a digital value of the drift-time.

The single hit resolution is determined for the OT by determining the width of the hit distance residual distribution. This is shown in Fig. 3.12 giving a resolution of 205  $\mu\text{m}$ .

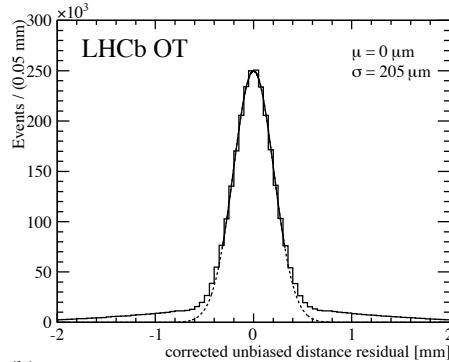


Figure 3.12: The hit distance residuals for the OT, from Ref. [91].

### 3.2.5 Ring imaging Cherenkov detectors

The ring imaging Cherenkov detectors (RICH) provide essential information about the particle identification of tracks, allowing different hadronic species to be distinguished. This allows kaons and protons to be distinguished from one another and the abundant pions tracks in a typical event. It can also help differentiate leptonic tracks in combination with the calorimeters and muon systems.

Two RICH sub-detectors are present, each optimised for particles with different momentums. The first, RICH1, is located between the VELO and TT, before the particles have passed through the magnetic field. This provides discrimination primarily for low momentum tracks between 1–60 GeV/ $c$ . A second sub-detector, RICH2, is located between the IT and OT tracking stations and calorimeters, after the particles have travelled through the magnetic field. This caters for the higher momentum particles in the range 15–100 GeV/ $c$ .

The RICH detectors help to determine the species of the particles passing through it by capturing Cherenkov radiation emitted by the particle as it traverses a transparent medium referred to as a radiator. Cherenkov radiation is produced by particles travelling above the phase velocity of light in that specific medium. The radiation propagates at constant angle  $\theta$  to the particles direction, determined by the speed of the particle  $\beta = v/c$  and refractive index of the material  $n$  as follows

$$\cos \theta = \frac{1}{\beta n}. \quad (3.1)$$

This light is collected by mirrors and focussed onto light sensors.

Although the particles produced in the collisions are all highly energetic and travelling close to the speed of light, the differences in the masses mean the species travel at slightly different velocities. This then results in a different Cherenkov angle. In combination with the momentum measurement from the magnet spectrometer, different mass hypotheses can be compared. The Cherenkov angles for different particle species as a function of momentum are shown in Fig. 3.13.

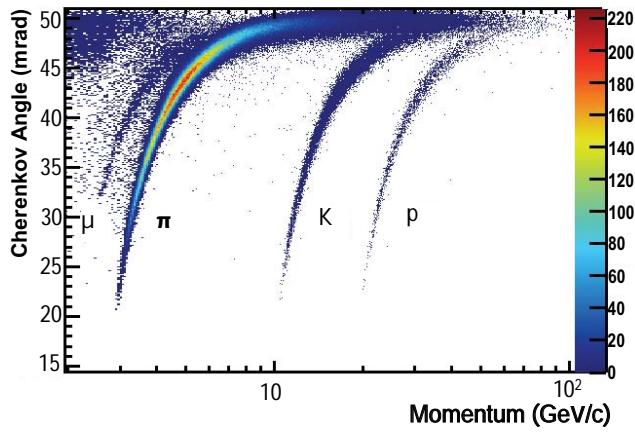


Figure 3.13: The Cherenkov angle for different species of particles as a function of momentum, from Ref. [92].

In both RICH1 and RICH2 the cone of Cherenkov radiation is reflected off spherical and planar mirrors. This not only redirects the radiation onto Hybrid Photon Detectors (HPDs) situated outside of the acceptance but also focusses the radiation into a single ring. The radius of this ring is used to determine the Cherenkov angle  $\theta$ . The identity of particles is determined by comparing likelihood for a given ring of hits to have been created by various species. This is performed globally for all tracks in an event. All tracks are assumed to be pions as these are most abundant. The hypothesis of each is switched to  $e$ ,  $\mu$ ,  $K$  and proton, leaving all other tracks unchanged, and the likelihood recomputed. The species resulting in the minimum value of the global event likelihood is kept. This is repeated until all tracks are set to their preferred hypothesis. The difference in the global event log-likelihood is then

used to indicate the species of each track. Each species is determined relative to the pion hypothesis

$$\text{DLL} = \Delta \log \mathcal{L}(X - \pi) = \log \mathcal{L}(X) - \log \mathcal{L}(\pi), \quad (3.2)$$

where  $\mathcal{L}(X)$  is the global event likelihood when the given track is of the species  $X$ .

## RICH1

The RICH1 sub-detector was designed to aid particle identification for lower momentum tracks in the range 1–60 GeV/ $c$  using two radiators: aerogel and decafluorobutane ( $\text{C}_4\text{F}_{10}$ ). These materials have refractive indices of  $n = 1.030$  and  $n = 1.0014$  respectively for  $\lambda = 400$  nm light. A diagram of the RICH1 detector as seen from the side is shown in Fig. 3.14. To minimise the amount of material that the particles pass through, the RICH1 detector is attached directly onto the VELO exit window.

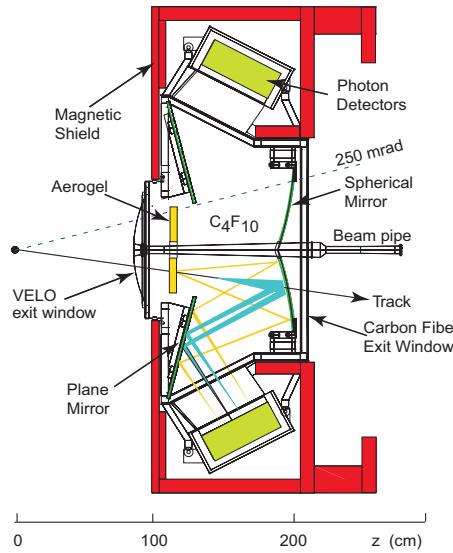


Figure 3.14: Schematic the RICH1 sub-detectors viewed from the side, from Ref. [86].

In Run II the aerogel was removed from the RICH1 detector. Although this catered specifically to the lower momentum tracks, the overall PID efficiency increased [93]. The ability of the aerogel to provide particle identification for kaons below the  $\text{C}_4\text{F}_{10}$  threshold was deteriorated by the number of photons in RICH1 as a result of the high track multiplicity. As well as improving the reconstruction speed,

the removal of the aerogel also allowed full use of the  $\text{C}_4\text{F}_{10}$  gas located between the gel and the VELO entrance window that would otherwise have been blocked.

The single photon resolution for RICH1 is 1.65 mrad, and found to be the same in Run I and Run II.

## RICH2

The RICH2 detector is designed to provide particle identification information for higher momentum tracks in the range 15–100  $\text{GeV}/c$  using a single gas radiator tetrafluoromethane ( $\text{CF}_4$ ). The refractive index for this material is  $n = 1.0005$  for  $\lambda = 400 \text{ nm}$  light.

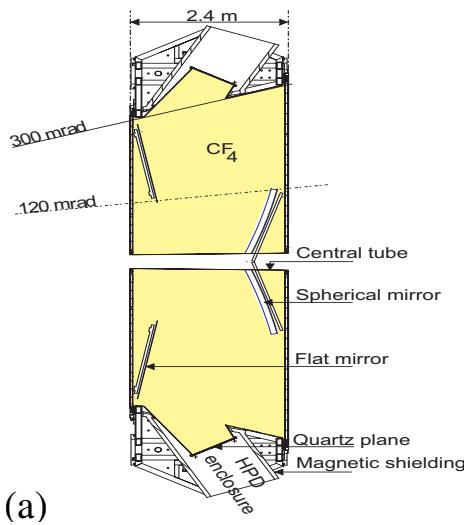


Figure 3.15: Schematic the RICH2 sub-detector viewed from above, from Ref. [86].

The single photon resolution for RICH2 is 0.67 mrad and found to be consistent between Run I and Run II [93].

## Performance

The performance of the RICH detectors can be quantified in terms of the selection efficiency and misidentification rates for different pairs of species. The efficiency of kaon identification and rate of pions being misidentified as kaons is shown in Fig. 3.16 for Run I and Run II. Due to the removal of the aerogel in RICH1, the kaon and pion separation shows a slight improvement in Run II.

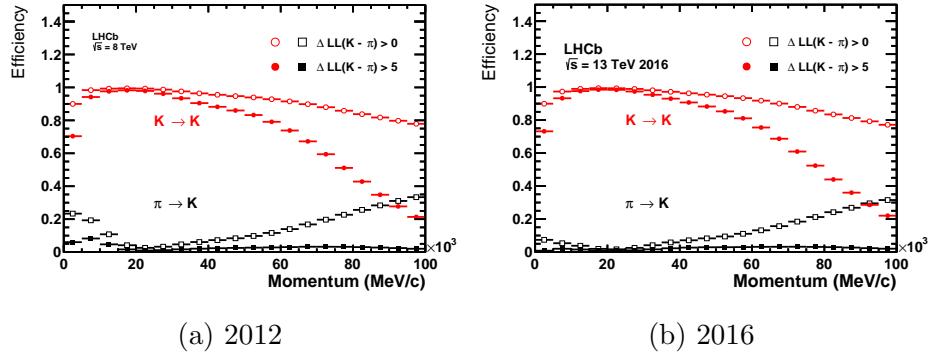


Figure 3.16: The RICH detector performance for kaon and pion separation, from Ref. [92].

### 3.2.6 Calorimeters

The calorimetry system is made of four components that work together to serve a number of purposes. Primarily, the detectors measure energy deposited by particles, vital for neutral pions or photons that don't interact with the tracking stations. Additionally, the calorimeters provide fast measurements of the energy deposited by interacting particles, suitable for use in the trigger. Finally, the four components work together to improve the discrimination of different particle species, in particular between electrons and hadrons.

All of the calorimeter components work on the same basic principle; the particles pass through a transparent materials producing a flash of light, known as scintillation light. This light is collected and measured; the quantity determines the energy deposited by the particle.

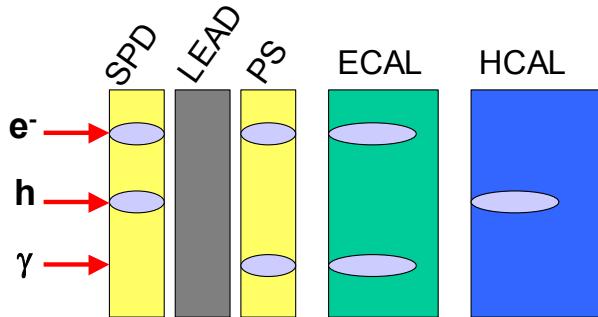


Figure 3.17: A diagram illustrating how the calorimeter system aids particle identification, from Ref. [94].

The first layer of the calorimeter system is the scintillator pad detector (SPD), as shown in Fig. 3.17. This provides information about whether the incident particles are neutral or charged. After this comes the pre-shower detector (PS). This indicates the electromagnetic nature of the particle. For charged particles this separates electrons and charged hadrons; for neutral particle it helps separate photons from neutral hadrons. Next comes the electromagnetic calorimeter (ECAL) tasked with measuring the energy of electromagnetically interacting particles, including electrons and photons. Finally, the hadronic calorimeter (HCAL) determines the energy deposited by hadrons, including pions, kaons, protons and neutrons.

### Scintillator pad detector and Pre-shower detector

The SPD and PS are almost identical scintillating pads detectors constructed from a mix of polystyrene and wavelength-shifting dopants. The pads are coated in light-proof paper and contain loops of wavelength-shifting fibres to collect the scintillation light. The pads are segmented into cells, matching the granularity of the ECAL detector. The fibres are connected to long clear fibres connected to photo-multiplier tubes (PMT) outside the detectors acceptance. Between the SPD and PS detectors is a 15 mm thick layer of lead. This layer initiates electrons and photons to interact electromagnetically, depositing some of their energy in the PS scintillating pads.

### Electronic Calorimeter

The ECAL is made of 66 layers of alternating lead and scintillator, with a total depth of 42 cm. This thickness corresponds to 25 radiation lengths. The layout of the layers are shown in Fig. 3.18. Scintillating light is captured and transported by wavelength-shifting wires that pass through the entire structure. The light is measured by PMTs positioned at the end of the structure. The cells of the ECAL are segmented differently depending on the proximity to the beam-pipe, as shown in Fig. 3.18. The cells are split into inner, middle and outer regions with sizes  $4 \times 4$  cm,  $6 \times 6$  cm and  $12 \times 12$  cm respectively.

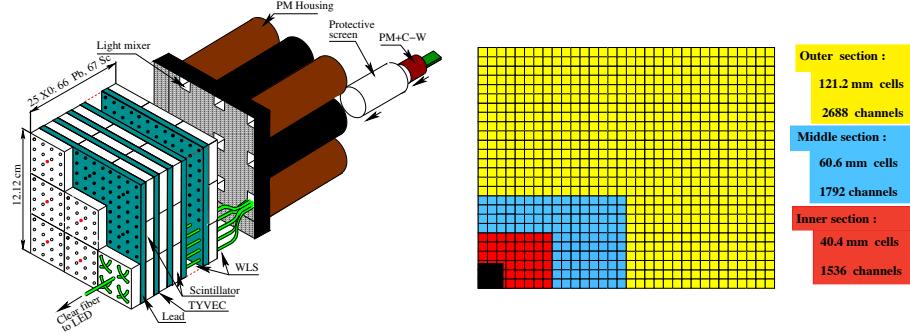


Figure 3.18: A diagram of one of the inner ECAL modules (left) and the layout of one quarter of the ECAL modules (right), from Ref. [86] and Ref. [95].

The resolution of the ECAL can be parametrised as

$$\frac{\sigma_E}{E} = \frac{(8.5 - 9.5)\%}{\sqrt{E}} \oplus 0.8\%, \quad (3.3)$$

where the energy  $E$  is measured in GeV (from Ref. [96]). During data taking the ECAL receives a large dose of radiation. This radiation damages the detector modules as well as the light guides and PMTs. The measurement of irradiation can be quantified by looking at the PMT light yields, as shown in Fig. 3.19. This can be extrapolated to determine the innermost ECAL modules will remain operational for a 13 TeV data set corresponding to an integrated luminosity of  $10\text{--}13 \text{ fb}^{-1}$ .

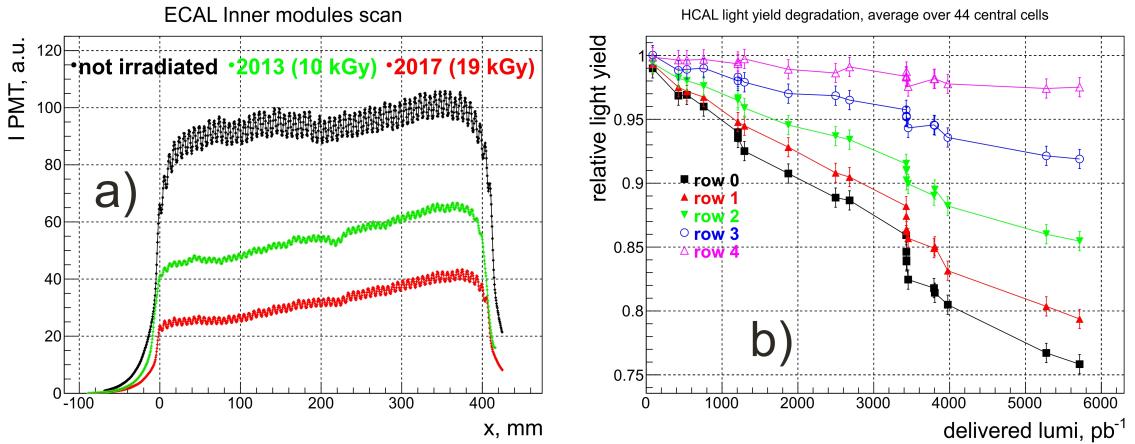


Figure 3.19: The result of radiation damage to the ECAL (a) and HCAL detectors (b), from Ref. [96].

### Hadronic calorimeter

The HCAL is a 500 tonne detector constructed from iron plates and scintillating tiles. Unlike the ECAL, these are arranged parallel to the beam-pipe as shown in Fig. 3.20. The cells vary in size in the different regions of the detector; the inner regions correspond to  $13 \times 13$  cm cells and the outer to  $26 \times 26$  cm. The widths of the iron tiles are chosen to be one radiation length (1 cm), and the lengths are one hadron interaction length. The scintillation light is collected by wavelength-shifting fibres that are read out by PMTs located at the rear of each cell.

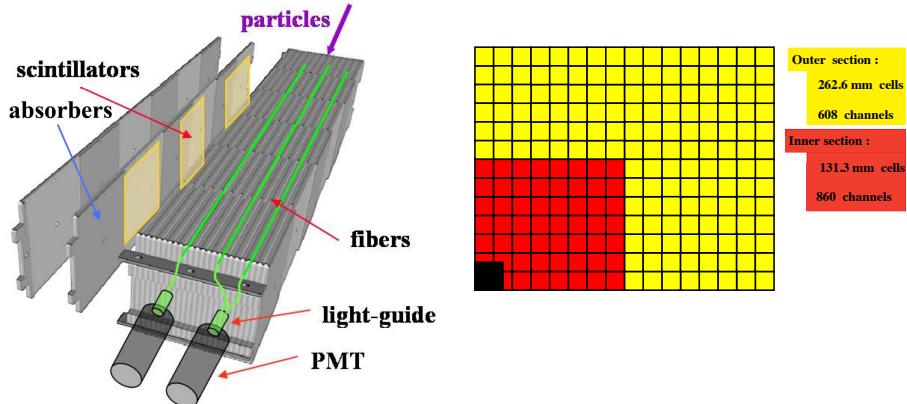


Figure 3.20: A diagram of one of the HCAL modules (left) and the layout of one quarter of the HCAL modules (right), from Ref. [86].

The resolution of the HCAL can be parametrised as

$$\frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%, \quad (3.4)$$

where the energy  $E$  is measured in GeV (from Ref. [96]). The HCAL receives a large dose of radiation during data taking. This mainly affects the detector modules as the PMTs are located further away from the beam. The decrease in the light yield as a function of the accumulated data set is shown in Fig. 3.19.

#### 3.2.7 Muon system

The muon system serves two main purposes: to provide fast muon  $p_T$  measurements for use in the hardware trigger; and to provide muon identification information for

use in the software trigger and offline analyses. The muon system is made of five rectangular stations. The first (M1) is positioned before the calorimeters and the remaining (M2–M4) are positioned after, as shown in Fig. 3.21. The latter stations alternate with 80 cm layers of iron absorbers. The muon system is the last sub-detector as muons are highly penetrating particles, travelling through more material than other species. The interleaving iron absorbers help to isolate muons, as other particles are stopped by these layers.

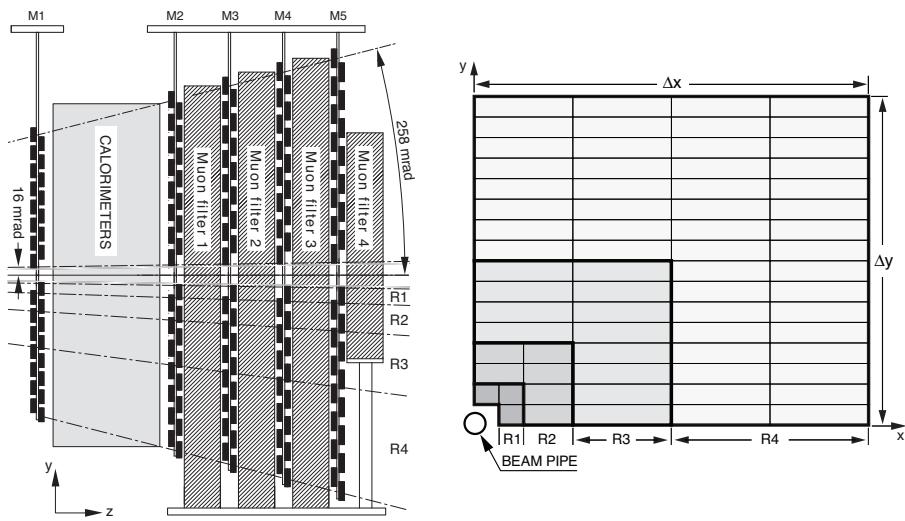


Figure 3.21: Side view of the Muon sub-detector (left) and layout of the muon system regions (right), from Ref. [86].

Each of the stations are divided into four regions R1–R4. These are positioned increasingly further from the beam-pipe with increasingly large areas. Each of the regions are segmented into channels; the inner regions have more segmentation than the outer regions such that the particle flux is roughly constant over the four regions.

The majority of the detector is constructed from Multi Wire Proportional Chambers. These chambers contain a gaseous mixture of carbon dioxide, argon and tetrafluoromethane. When muons pass through this gas they cause ionisation. An array of gold-plated tungsten wires separated by 2 mm are held at a potential difference of around 2.5 kV with respect to the chamber's conductive walls. The ionised gas and electrons are deflected by the electric field and result in a flow of current.

Due to the larger flux of particles, the innermost region (R1) of the first station (M1) is instrumented using triple gas electron multiplier detectors. The muons propagating through these detectors ionise a gas mixture made up of argon, carbon dioxide and tetrafluoromethane. The resulting electrons are accelerated toward an anode through three foil layers. The foil layers multiply the number electrons, creating a cascade. This type of detector was chosen for the inner region as it is required to be especially radiation hard such that the detector can withstand 10 years of operation.

### 3.2.8 Trigger

The primary role of the trigger is to reduce the nominal 40 MHz beam crossing frequency to a more manageable rate that can realistically be recorded. The trigger is composed to two parts: a custom hardware trigger called Level 0 (L0) that operates at the nominal 40 MHz beam crossing rate; and an asynchronous High Level Trigger (HLT) running software algorithms. These two stages reduce the event rate to around 2 kHz, exploiting the characteristics of  $b$ - and  $c$ -hadron decays to preferentially retain signal decays and discard the many uninteresting events.

#### Level 0

The L0 trigger takes input from the muon system, calorimeters and pile-up veto in the VELO. These three systems feed into a single unit to make a global decision about the event. Although the nominal bunch spacing is 25 ns, the L0 decision is made after 4  $\mu$ s. During this time the detector signals are stored in a *pipeline* buffer. Half of this 4  $\mu$ s budget is required for the time-of-flight of the particles and delays in the electronics, leaving the L0 trigger 2  $\mu$ s to reach a decision about each event.

The L0 calorimeter trigger uses input from the SPD, PS, ECAL and HCAL to trigger events. The particles are reconstructed using  $2 \times 2$  cell clusters to calculate the total transverse energy  $E_T$  deposited. The different particle species are inferred by combining the information from the different calorimeter components.

**L0Hadron:** Hadronic deposits are created by finding the highest energy HCAL cluster with a corresponding ECAL cluster in front of it.

**L0Photon:** Photons are created from the highest ECAL deposits with corresponding PS hits.

**L0Electron:** Electrons are created from the highest ECAL deposits with corresponding PS and at least one SPD hit.

If any of these candidates pass corresponding  $E_T$  thresholds, then the trigger fires.

The L0 muon trigger selects the two highest  $p_T$  muons in each of the four quadrant of the muon system.

**L0Muon:** Any of the eight muons can fire this trigger if their  $p_T$  is greater than a threshold.

**L0DiMuon:** Pairs of muons can fire this trigger if the quantity  $p_{T1} \times p_{T2}$  is greater than a threshold.

The third component of the L0 trigger, the pile-up veto, helps to determine how many primary interactions occurred in a given event. Two VELO  $r$ -sensors measure the radii of backward tracks,  $r_a$  and  $r_b$ . For tracks originated at the same vertex the ratio of these is a constant,  $k = r_b/r_a$ . The ratio of hits are compared and events with multiple events can be vetoed.

## High Level Trigger

The HLT is a software trigger comprised of C++ algorithms on a dedicated processing farm. It is further subdivided into two sections: HLT1 and HLT2. The first stage, HLT1, reduces the rate of events by matching L0 deposits to tracks in the VELO and T-stations. The software algorithm performs partial pattern recognition identify track candidates and to ensure that  $\gamma$  or  $\pi^0$  L0 candidates are not associated to tracks. This part of the trigger reduces the L0 output rate (870 kHz) to around 43 kHz. The VELO tracks are used to create vertices with 5 or more tracks. Tracks are accepted if they are consistent with originating at one of these vertices. For events selected with the muon L0 lines, the VELO tracks are extrapolated to the muon stations to

form muon candidates. These VELO or muon candidates are extrapolated to the T-stations (OT and IT) and reconstructed with a Kalman filter [97, 98].

The second stage, HLT2, performs a full event reconstruction, allowing explicit decays to be reconstructed. Many different trigger algorithms are included in HLT2, reconstructing a wide range of decays. Additionally, a set of inclusive  $b$ -hadron decay lines exist, exploiting the common characteristics of these decays. These *topological* trigger lines make decisions based on 2-, 3- or 4-body combinations of tracks that can form a partially or fully reconstructed  $b$ -hadron decay. These tracks are required to pass a minimum distance of closest approach requirement with respect to each other. The invariant mass of these tracks would span a large range in the case of additional non-reconstructed particles, so a quantity called the corrected mass is calculated instead:

$$m_{\text{corr}} = \sqrt{m^2 + |p_{T\text{miss}}'|} + |p_{T\text{miss}}'|, \quad (3.5)$$

where  $p_{T\text{miss}}$  is the missing momentum transverse to the flight direction of the particle, defined by the vector between the PV and 2-, 3-, or 4-body vertex. Requirements are made to  $m_{\text{corr}}$  to ensure the vertex is consistent with a  $b$ -hadron decay. The tracks are required to have a significant impact parameter with respect to the PV. This helps to remove  $c$ -hadrons combined with an additional track both originating at the PV.

The HLT underwent various improvements between Run I and Run II. The number of CPU cores in the Event Filter Farm (EFF) increased from around 29,000 to 52,000. Most notably the HLT1 and HLT2 components were modified to operate asynchronously. This means during Run II the events passing HLT1 are stored in a temporary buffer (around 11 petabytes in size) to be processed later by HLT2. This allows alignment and calibration to be performed between the running of HLT1 and HLT2. Dedicated calibration samples are produced by HLT1 and processed on the EFF. This includes calculating the VELO and tracking station alignments to optimise the momentum resolution. Additionally, the refractive index of the RICH radiators is calibrated using samples of particles with known masses from specific decays. The

difference in the measured and expected Cherenkov angle is used to determine the required calibration.

These improvements in calibration allow the HLT2 reconstruction to be effectively the same as can be achieved offline. An additional trigger system called the Turbo Stream [99] was added in Run II to create candidates ready for analysis straight out of HLT2. However, candidates analysed in this these were processed in the trigger and again offline.

### 3.2.9 Reconstruction

The reconstruction of charged and neutral particles is described in this section.

#### Charged particle reconstruction

Charged tracks are reconstructed using the hits in the VELO, TT, IT and OT. These are classified into different categories depending on their route taken through the detector as shown in Fig. 3.22. The IT and OT are collectively referred to as the T-stations.

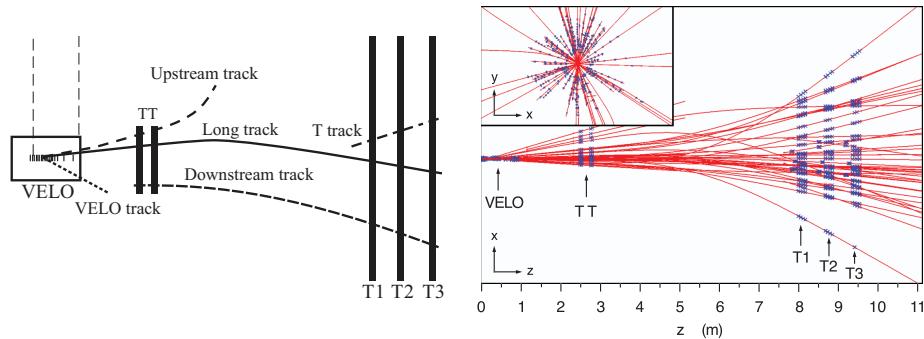


Figure 3.22: Diagram of the different track reconstruction types (left) and an illustration of the reconstructed tracks in a typical event from Ref. [90].

**Long tracks:** these tracks have passed through the VELO, TT and T-stations.

**Upstream tracks:** these only pass through the VELO and TT tracker, not the T-stations. This may be because their momentum is too low so the tracks get bent out of the acceptance in the magnetic field.

**Downstream tracks:** these tracks only contain TT and T-station hits. They provide vital information about long-lived particles that decay after they have traversed the VELO.

**VELO tracks:** these tracks only contain hits in the VELO. They can help improve the primary vertex determination.

**T tracks:** these tracks don't pass through the VELO or TT and are likely to be produced in secondary interaction between the particles and detector material.

The identification of charged tracks used input from the RICH sub-detectors, along side input from the calorimeter and muon systems to identify electrons and muons respectively. As already described in Sec. 3.2.5, the RICH particle identification is performed by comparing the global event likelihood when changing the particle hypothesis. This is combined with information from the calorimeter and muon systems to create an overall particle identification variable

$$\Delta \log \mathcal{L}_{\text{comb}}(X - \pi) = \Delta \log \mathcal{L}_{\text{RICH}}(X - \pi) + \Delta \log \mathcal{L}_{\text{Calo}}(X - \pi) + \Delta \log \mathcal{L}_{\text{Muon}}(X - \pi), \quad (3.6)$$

where RICH, Calo and Muon refer to the difference in the log-likelihood of the  $X$  and  $\pi$  hypothesis as determined by the RICH, calorimeter or muon systems respectively.

The calorimeter log-likelihood is constructed by comparing the electron to hadron hypotheses. This uses input from the ECAL, HCAL and PS sub-systems

$$\Delta \log \mathcal{L}_{\text{Calo}}(e - h) = \Delta \log \mathcal{L}_{\text{ECAL}}(e - h) + \Delta \log \mathcal{L}_{\text{HCAL}}(e - h) + \Delta \log \mathcal{L}_{\text{PS}}(e - h) \quad (3.7)$$

For the HCAL and PS this input is determined using the energy deposited. For the ECAL this is constructed comparing the ratio of energy to momentum,  $E/pc$ , for the candidates, which peaks around one for electrons and at lower values for hadrons. Additionally, the centre of the ECAL deposit and extrapolated track position are compared.

The muon system log-likelihood is determined by extrapolating the track in question to the muon stations. The average squared distance of the extrapolated positions and hits in the stations are used to compare the muon to non-muon hypothesis.

### Neutral particle reconstruction

The energy of neutral particles are determined directly from the calorimeter system. This system is calibrated during running to maintain the optimal resolution.

**Photons:** Isolated ECAL clusters with an corresponding cluster in the PS are used to create photon candidates. The energy is determined from the total energy deposited in the ECAL and PS.

**Neutral pions:** These mesons decay dominantly to two photons, however the reconstruction depends on the transverse momentum of the  $\pi^0$ . For low  $p_T$  decays the two photons are well separated (*resolved*) leading to two photon deposits. For higher  $p_T$  decays the two photon deposits overlap leading to a single ECAL cluster (*merged*).

**Bremsstrahlung correction:** as a result of their small mass, electrons can radiate Bremsstrahlung photons. If this occurred before the magnetic field the photon produced results in a separate ECAL cluster. If this radiation occurs after the magnet the photon cluster can overlaps with the electron cluster. Clusters in the ECAL that are consistent with radiated Bremsstrahlung photons are used to correct the initial energy of the electron.

## 3.3 Luminosity determination

The observed rate  $R$  of a particular physical processes at particle colliders are defined by the luminosity  $\mathcal{L}$  and the cross-section of the process,  $\sigma$ ,

$$R = \mathcal{L}\sigma. \quad (3.8)$$

Experimentally, the rate of specific interactions are measured, so precise measurements of the luminosity required to determine the corresponding cross-section. Additionally, the total integrated luminosity is useful to understand the size of the data set.

The luminosity can be determined *indirectly*, by reversing Eq. 3.8 and inputting a well-known cross section for a calibration process. Alternatively the luminosity can be determined *directly* by calculating it from the machine parameters. This is determined as follows

$$\mathcal{L} = f_{\text{rev}} N_1 N_2 \Omega, \quad (3.9)$$

where  $f_{\text{rev}}$  is the revolution frequency,  $N_1$  and  $N_2$  are the number of protons in the two beams and  $\Omega$  is a geometrical factor that quantifies the overlap of the two bunches [100]. For two beams of ultra-relativistic particles the overlap integral can be approximated as

$$\Omega = 2c \int \rho_1(x, y, z, t) \rho_2(x, y, z, t) dx dy dz dt \quad (3.10)$$

where  $\rho$  describes the particle density distribution.

Two methods are used to *directly* determine the luminosity at LHCb, described here.

### 3.3.1 Van der Meer scans

The luminosity of the collisions can be determined using the van der Meer technique [101]. This involves sweeping two beams across each other in the transverse plane. By measuring the interaction rate as a function of the beam displacement, the overlap of the two beams can be determined, and hence the absolute luminosity. The two beams are moved simultaneously, alternating between the two transverse directions as shown in Fig. 3.23.

The absolute luminosity is calculated for the specific van der Meer scan fills, and extrapolated to the full data set using the interaction rate of standard processes. The effective visible cross-section  $\sigma_{\text{vis}}$  is calibrated for the standard processes during the

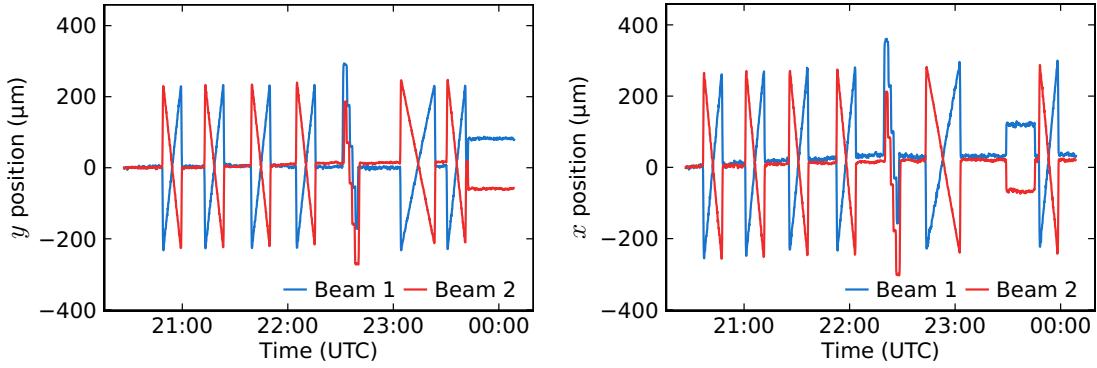


Figure 3.23: Beam positions during a van der Meer scan to determine the absolute luminosity, from Ref. [100].

absolute luminosity determination fill. The total integrated luminosity for a data sample can be determined by the accumulated rate of the standard process.

### 3.3.2 Beam gas imaging

The luminosity can also be calculated by exploiting the interactions between the beams and residual gas molecules in the beam pipe vacuum. The precision of the VELO allows the beam-gas interaction vertices to be reconstructed. These vertices allow the beam position, shape and angle to be determined. The typical distribution of beam-gas vertices in the  $xz$  and  $yz$  planes are shown in Fig. 3.24. The crossing angle in the horizontal plane is clearly visible. The overlap of the two bunches is

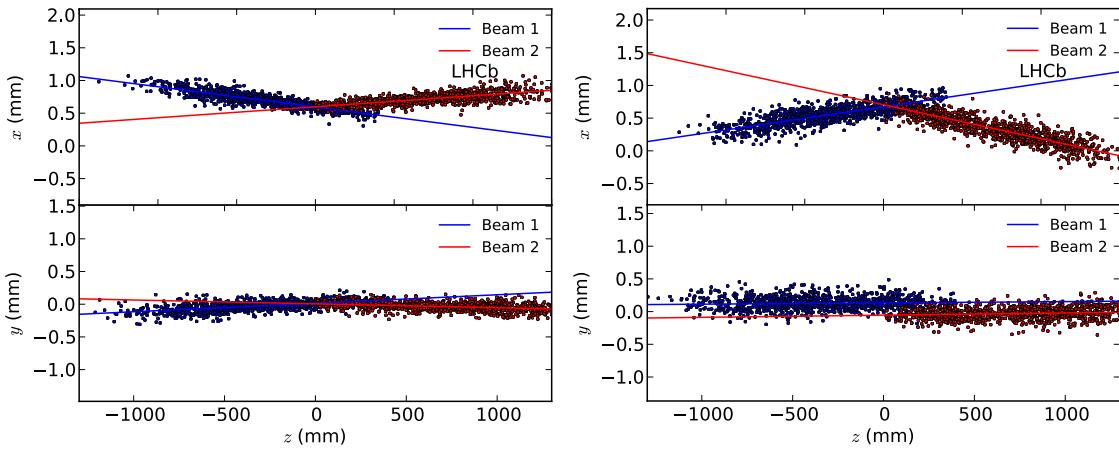


Figure 3.24: The positions of beam-gas interactions reconstructed in the VELO sub-detector, from Ref. [100].

determined by first fitting the beam profiles to obtain the three dimensional density  $\rho(x, y, z)$  for each beam. This is possible due to beam crossings in which only one of the two beams were present. This means all reconstructed beam-gas interactions can be attributed to the single bunch.

### VELO resolution

The distributions of the beam gas interactions as reconstructed by the VELO sub-detector are not simply the beam profiles as vertex position resolution must be accounted for. The resulting distribution is the convolution of the beam profile and resolution distribution. The VELO vertex resolution is strongly dependent on the number of tracks used to create the vertex, as well as the longitudinal position of the vertex.

The resolution of vertices is estimated by splitting the tracks contributing to vertices into two random sub-samples. The vertex position for each of these sub-samples,  $v_{1,2}$  is determined and the difference in the vertex positions  $\Delta v = v_1 - v_2$  is found. The width of the  $\Delta v$  distribution when splitting many vertices is used as a proxy for the resolution. The resolution determined for different values vertex track multiplicity and longitudinal position is shown in Fig. 3.25. These resolution values

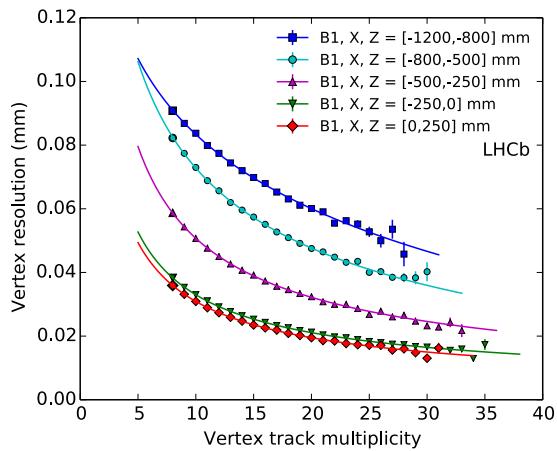


Figure 3.25: The resolution of beam-gas interactions as a function of the vertex track multiplicity and longitudinal position, from Ref. [100].

are used to subtract the smearing from the VELO resolution such that the beam profiles can be accurately determined.

### Real-time beam measurements

The beam gas imaging method additionally allows measurements of the beam profiles to be made in real-time during data taking. As detailed in Sec. 3.1.2 the LHC runs with a smaller number of bunches than the full capacity would allow. The distribution of filled and empty bunch slots are usually chosen to maximise the luminosity for ATLAS and CMS, therefore only a subset of the total bunches collide at LHCb. Those beam crossings in which only one of the two bunches pass through LHCb can be used to make measurements of the beam-gas interactions during normal running. These provide information about the beam angle, positions and shapes. The shape parameters can be used to calculate the emittance of the beams. This information is published back to the LHC, providing important information about the performance of the machine [102].

# Chapter 4

## Event selection

### Contents

---

<b>4.1</b>	<b>Data samples and Simulations</b>	<b>65</b>
4.1.1	Data samples	65
4.1.2	Simulation samples	65
<b>4.2</b>	<b>Online selection</b>	<b>66</b>
<b>4.3</b>	<b>Offline selection</b>	<b>69</b>
4.3.1	Selection requirements	70
4.3.2	Particle identification requirements	79
4.3.3	Charmless and single-charm backgrounds	79
4.3.4	Misidentified $D$ and $A_c^+$ hadrons	81
4.3.5	Invariant mass vetoes	83
4.3.6	Normalisation mode veto	85
4.3.7	Multivariate analysis	87
4.3.8	Impact parameter requirements	103
<b>4.4</b>	<b>Multiple candidates</b>	<b>106</b>
<b>4.5</b>	<b>Refitting the decay chain</b>	<b>107</b>
<b>4.6</b>	<b>Invariant mass distributions</b>	<b>108</b>

---

In this chapter the procedures developed to reconstruct and select  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates are described. In both cases the branching fractions are measured relative to the normalisation channel  $B^+ \rightarrow D_s^+ \bar{D}^0$ . The corresponding selection for the normalisation channel  $B^+ \rightarrow D_s^+ \bar{D}^0$  is also described.

## 4.1 Data samples and Simulations

### 4.1.1 Data samples

The searches for  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are performed using a combined data sample comprising of the entire Run I data set and a part of the Run II data set, namely the years 2015 and 2016. The total integrated luminosity obtained for each year is listed in Table 4.1, along with the corresponding centre-of-mass energies.

Year	Integrated luminosity (fb <sup>-1</sup> )	$\sqrt{s}$ (TeV)
2011	1.0	7
2012	2.0	8
2015	0.3	13
2016	1.5	13

Table 4.1: The integrated luminosities obtained during the different data taking periods used in this analysis and the corresponding centre-of-mass energies ( $\sqrt{s}$ ).

### 4.1.2 Simulation samples

Many aspects of the searches for  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays requires input from simulation samples. This includes the relative selection efficiencies of the signal and normalisation channels and the determination of the invariant mass distributions of the candidates. The samples are generated separately for each of the running periods listed in Table 4.1 and for both polarities of the LHCb dipole magnet. The simulation samples for  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays are generated assuming they proceed via pseudo-two-body decays. In contrast, the phase-space dependence of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays must be considered when generating these simulations. Samples are generated assuming a flat distribution across the phase-space. Additionally, simulation samples are required for many of the background processes considered in the analyses. These are generated using the same framework and processed using the same reconstruction as the signal modes.

In the simulations,  $pp$  collisions are generated using PYTHIA [103, 104] with a specific LHCb configuration [105]. Decays of hadronic particles are described by EVTGEN [106], in which final-state radiation is generated using PHOTOS [107]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [108] as described in Ref. [109].

## 4.2 Online selection

The online event selection is performed by the LHCb trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which reconstructs the full event, as detailed in Sec. 3.2.8. Trigger decisions are made globally for a given event; either the event is saved or it's discarded. However, for events that are retained more detailed information about the specific interaction that initiated the trigger can be used to help isolate events containing signal candidates. At both hardware and software level there are a number of different triggers that can lead to an event being retained. In hardware this is broadly separated according to the sub-detector that contained the triggering deposit. Deposits in the muon chambers or hadronic calorimeter would fire `L0Muon` or `L0Hadron` respectively. Electromagnetic calorimeter deposits could result in `L0Electron` or `L0Photon` firing. The  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s^+ \phi$  decays are both reconstructed in fully hadronic final states, therefore events retained as a result of the `L0Electron`, `L0Photon` or `L0Muon` trigger firing are unlikely to have been caused by the signal decays.

In addition to specifying which triggers fired, it is possible to use more granular information about the specific hits initiating the trigger to improve the selection of signal candidates. The tracks constituting the reconstructed candidate are matched to the deposits in each sub-detector. The fraction of hits required to match in order for the trigger decision to be associated to that object vary depending on the sub-detector in question, as listed in Table 4.2. Reconstructed objects are classified into categories when considering their relationship to the various triggers that fired in a given event.

Sub-detector	Fraction of matching hits
HCAL	1%
ECAL	1%
Muon	0.01%
TT	0%
IT and OT	70%
VELO	70%

Table 4.2: The fraction of matching hits required for a offline reconstructed object to be associated to a online sub-detector deposit.

**Triggered on Signal:** if the deposited matched to a reconstructed object was sufficient to have fired a given trigger then this object is said to be *Triggered on Signal* (**TOS**) with respect to that trigger (Fig. 4.1a). As such, if the hits that this object caused were removed from the events the trigger would no longer fire.

**Triggered Independently of Signal:** conversely, if the specific trigger fired solely due to other interactions within the same event the object is *Triggered Independently of Signal* (**TIS**) with respect to that trigger (Fig. 4.1b). In this case, removal of hits matched to the object would not affect the trigger decision. It is possible for a candidate to be both **TIS** and **TOS** (Fig. 4.1d).

**Triggered on Both:** a third category is possible, *Triggered on Both* (**TOB**), where both the signal object and another object are required to reach the threshold to fire a trigger (Fig. 4.1c). In this situation neither are sufficient to individually fire the trigger, but removing either one of them would prevent the trigger firing. Events in this category are not considered in this analysis.

At the hardware trigger stage, the selected candidates are required to be **TOS** with respect to the hadronic trigger **L0Hadron**. This ensures the selected candidates were retained due to corresponding deposits in the hadronic calorimeter. Alternatively, candidates are selected if they are **TIS** with respect to the global hardware trigger **L0Global**; any of the hardware trigger subsystems can contribute to the **L0Global**

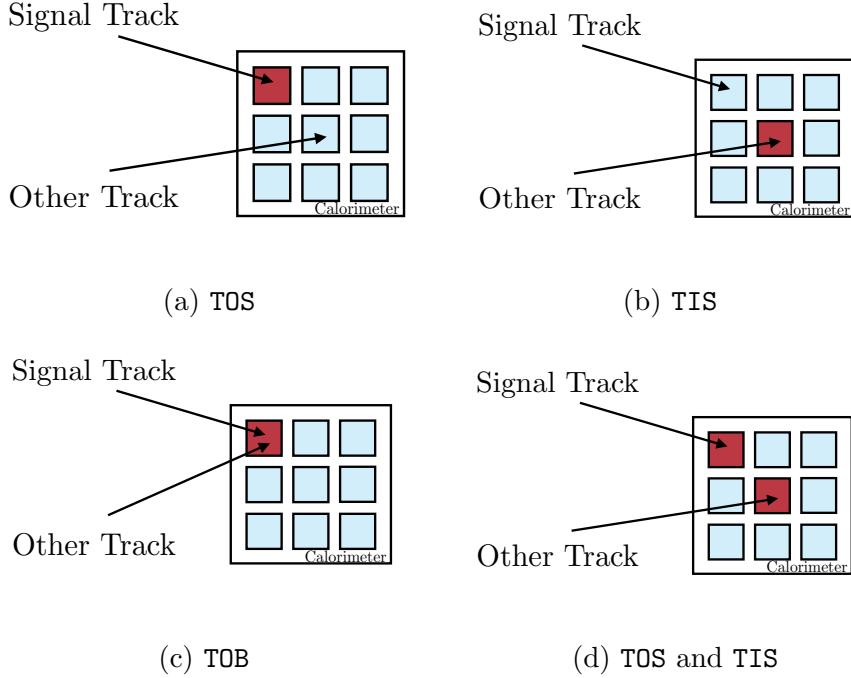


Figure 4.1: Simple schematics of the different trigger matching possibilities, illustrated with one signal track and one non-signal track. Cells with a deposit above the trigger threshold are shown in red.

decision. This allows candidates that have been retained due another highly energetic decay in the same event to contribute. This could be the decay of hadron resulting from the other  $b$  quark in a  $b\bar{b}$  pair production. The relative fractions of simulated decays selected by these **L0Hadron TOS** and **L0Global TIS** for  $B^+ \rightarrow D_s^+\phi$  decays decaying to various  $D_s^+$  final states are shown in Table 4.3.

$D_s^+$ decay mode	TOS & !TIS	TIS&!TOS	TIS&TOS	TIS or TOS
$D_s^+ \rightarrow K^+ K^- \pi^+$	34.0%	41.2%	22.9%	96.9%
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	34.9%	39.5%	23.8%	96.9%
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	35.9%	38.7%	23.5%	96.9%

Table 4.3: The fraction of simulated  $B^+ \rightarrow D_s^+\phi$  decays in each of the hardware trigger categories for the Run I conditions. Here **TOS** refers to candidates found to be **TOS** with respect to the **L0Hadron** trigger, and **TIS** to candidates found to be **TIS** with respect to the **L0Global** requirement. The percentages are given relative to the number of decays passing the reconstruction requirements.

The software trigger stage is split into two parts, HLT1 and HLT2. The first

stage, HLT1, requires that the selected candidates are associated with at least one well reconstructed track. This track must be significantly displaced from the primary interaction and has a large transverse momentum. In Run II, this trigger was reimplementing using a multivariate analysis trained to select a well reconstructed, high transverse momentum track that was significantly displaced from the primary interaction. An additional trigger was also included to select pairs of high transverse momentum tracks that form a vertex that is displaced from the primary interaction. The signal candidate is required to be TOS with respect to these trigger lines such that one of the candidate tracks passes this requirement.

At the second software stage, HLT2, two different algorithms are used to select candidates for this analysis. The first one uses a multivariate algorithm [110] to identify the presence of a secondary vertex that has two, three or four tracks and is displaced from any PV. At least one of these charged particles must have a transverse momentum  $p_T > 1.7 \text{ GeV}/c$  and be inconsistent with originating from a PV. This trigger algorithm is referred to as the topological trigger. The second algorithm selects  $\phi$  candidates decaying to two charged kaons. Each kaon must have a transverse momentum  $p_T > 0.8 \text{ GeV}/c$  and be inconsistent with originating from a PV. The invariant mass of the kaon pair must be within  $20 \text{ MeV}/c^2$  of the known  $\phi$  mass [28]. This inclusive  $\phi$  line is used to maximise the selection efficiencies as  $\phi$  mesons can contribute directly to  $B^+ \rightarrow D_s^+ \phi$  decays, as well as via the large fractions of  $D_s^+ \rightarrow \phi \pi^+$  decays that contribute to the  $D_s^+ \rightarrow K^+ K^- \pi^+$  final state. The relative fractions of simulated  $B^+ \rightarrow D_s^+ \phi$  decays selected by these different trigger algorithms are listed in Table 4.4. It can be seen that the inclusive  $\phi$  trigger algorithm helps additionally retain a larger fraction of decays for the  $D_s^+ \rightarrow K^+ K^- \pi^+$  decay.

### 4.3 Offline selection

Events passing any trigger requirement are saved to tape for processing offline. The stages of the offline reconstruction are detailed in this section.

$D_s^+$ decay mode	Topo & !Phi	Phi&!Topo	Topo&Phi	Topo or Phi
$D_s^+ \rightarrow K^+ K^- \pi^+$	50.9%	4.6%	43.1%	98.7%
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	67.6%	2.7%	28.3%	98.5%
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	67.5%	2.6%	28.5%	98.7%

Table 4.4: The fraction of simulated  $B^+ \rightarrow D_s^+ \phi$  decays in each of the HLT2 software trigger categories for the Run I conditions. Here **Topo** refers to candidates found to be **TOS** with respect to the topological algorithm, and **Phi** to candidates found to be **TOS** with respect to the inclusive  $\phi$  algorithm. The fractions of selected decays are measured relative to the HLT1 requirements.

### 4.3.1 Selection requirements

The large offline data samples passing the online trigger selection are habitually processed in a procedure known within LHCb as *Stripping*. This centrally managed processing builds candidates from tracks and neutral calorimeter objects in each event according to a set of predefined *Stripping Lines*. Each line builds a specific candidate decay, applying a loose set of preselection requirements, including kinematic, geometric and invariant mass selections.

The candidate  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are built using almost identical set of requirements due to the similar topologies of these decays. The  $\phi$  meson has a lifetime of  $(1.55 \pm 0.01) \times 10^{-22}$  s [28], therefore the kaon decay products effectively originate at the  $B^+$  decay vertex in a similar way to  $B^+ \rightarrow D_s^+ K^+ K^-$  decays. The normalisation channel,  $B^+ \rightarrow D_s^+ \bar{D}^0$  is built similarly, however the difference in lifetime between the  $\bar{D}^0$  and  $\phi$  mesons necessitates a slightly different topology. The  $\bar{D}^0$  meson has a lifetime of  $(4.101 \pm 0.015) \times 10^{-13}$  s, allowing the meson to travel away from the  $B^+$  decay vertex. Schematics of the signal and normalisation decay topologies are shown in Fig. 4.2.

The candidate  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s^+ \phi$  signal decays and  $B^+ \rightarrow D_s^+ \bar{D}^0$  normalisation decay are reconstructed in fully hadronic final states. Both the  $\bar{D}^0$  and  $\phi$  mesons are reconstructed using pairs of oppositely charged kaons. The branching fractions are  $\mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-) = (3.97 \pm 0.07) \times 10^{-3}$  and  $\mathcal{B}(\phi \rightarrow K^+ K^-) = (48.9 \pm 0.5)\%$  respectively [28]. Although the similar two-body hadronic decay  $D^0 \rightarrow K^- \pi^+$

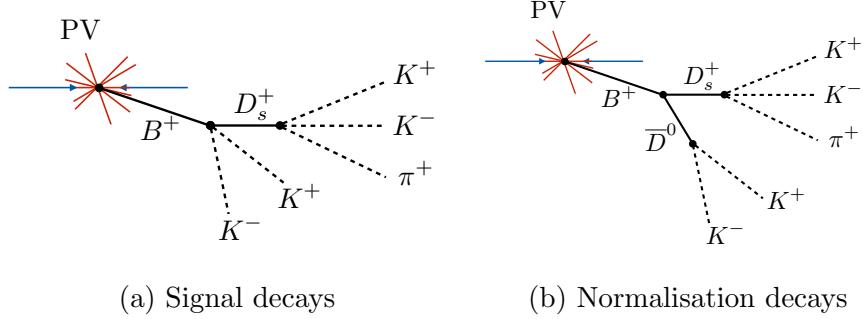


Figure 4.2: Schematics of the signal and normalisation decay topologies. The collision of the two protons (blue) results in the primary collision vertex (PV). Many promptly produced tracks (red) originate at the PV. In both cases, the long-lived  $B^+$  meson decay vertex is displaced from the PV. The long-lived charm mesons are also displaced from the  $B^+$  meson decay vertex.

has a larger branching fraction than mode chosen for the normalisation channel,  $\mathcal{B}(D^0 \rightarrow K^-\pi^+) = (3.89 \pm 0.04)\%$ , sharing the same final state helps to reduced systematic uncertainties in the ratio of selection efficiencies. Many differences in how pions and kaons interact with the detector can be neglected as they would affect the signal and normalisation channel in the same way. The  $D_s^+$  mesons used in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are reconstructed using the  $D_s^+ \rightarrow K^+ K^- \pi^+$  decay. The search for  $B^+ \rightarrow D_s^+ \phi$  additionally includes the modes  $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$  and  $D_s^+ \rightarrow K^+ \pi^- \pi^+$  to increase the sensitivity of the search. The branching fractions for these decays are listed in Table 4.5. In each instance the normalisation channel is reconstructed using the same  $D_s^+$  decay mode as the signal.

Decay	Branching fraction
$D_s^+ \rightarrow K^+ K^- \pi^+$	$5.45 \pm 0.17\%$
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	$1.09 \pm 0.05\%$
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	$0.66 \pm 0.04\%$

Table 4.5: The branching fractions for the different  $D_s^+$  decay modes used to reconstruct the signal and normalisation modes.

The selection of candidates makes requirements on many different distinguishing characteristics. The definitions of the relevant quantities are as follows:

**Mass**,  $m(X)$ : the invariant mass of the indicated particle or combination of particles.

**Momentum**,  $p$ : the total momentum of the particle or combination of particles.

**Transverse momentum**,  $p_T$ : The projection of the total momentum,  $p$ , on the plane perpendicular to the proton beam axis.

**Lifetime**,  $\tau$ : the time taken for the candidate to travel from its production vertex to its decay vertex.

**Products  $p_T$  scalar sum**,  $\sum |p_T|$ : the sum of the magnitudes of the transverse momentum for each of the decay products.

**Vertex quality**,  $\chi^2_{\text{Vtx}}$ : the minimised value of  $\chi^2/N_{\text{DOF}}$  as determined in the fit to the vertex location.

**Track quality**,  $\chi^2_{\text{Trk}}$ : the minimised value of  $\chi^2/N_{\text{DOF}}$  as determined in the fit track hits.

**Impact parameter, IP**: The shortest distance between a given track and a vertex location, as shown in Fig. 4.3.

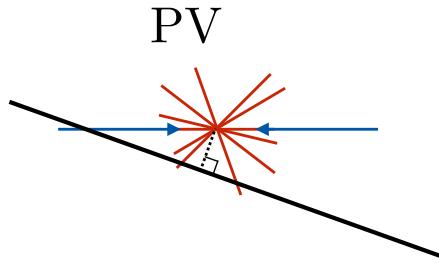


Figure 4.3: Impact parameter (dotted black line) between a track (black) and a vertex (red lines).

**Impact parameter significance**,  $\chi^2_{\text{IP}}$ : The difference of a given vertex's  $\chi^2/N_{\text{DOF}}$  with and without a specific track included in the fitting procedure.

**Flight distance significance,  $\chi^2_{\text{FD}}$ :** A measure of how significant the flight distance of a combination of particle is. Defined as the  $\chi^2$  associated with the difference in position of the two vertices,  $\vec{d} = \vec{v}_2 - \vec{v}_1$ , where  $\vec{v}_1$  and  $\vec{v}_2$  are positions of the first and the second vertices.

**Distance of closest approach, DOCA( $h, h'$ ):** The shortest distance between two tracks as shown in Fig 4.4.

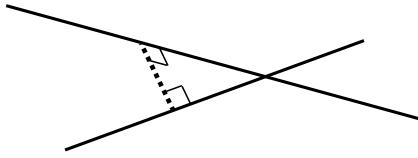


Figure 4.4: Distance of closest approach (dotted black line) between two tracks (black).

**Ghost track probability,  $P_{\text{Ghost}}$ :** This parameter quantifies the probability that a given track is an incorrect combination of tracking stations hits, known as a ghost track. The numerical value is the output of a Neural Network algorithm trained to separate true tracks from ghost tracks using simulated events. Various tracking parameters are inputs to the Neural Network including the number of hits in various tracking stations, the track fit quality and the number of tracks per event.

**Direction angle:** The angle between the particles momentum vector and the vector connecting the primary vertex and decay vertex as shown in Fig. 4.5.

**Particle identification, PIDK:** Something

As the final states are fully hadronic, the candidates are built from the combination of five tracks. Only *long* tracks (those with hits in the Velo and tracking stations) are used to build these mesons. The track  $\chi^2/N_{\text{DOF}}$  is required to be below 4.0 to ensure these are well reconstructed. Additionally, they are required to have a total

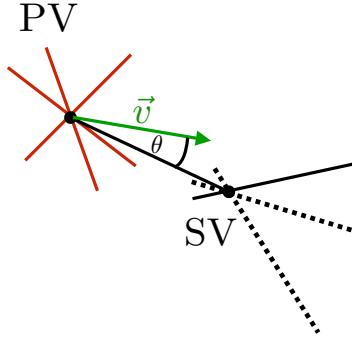


Figure 4.5: Direction angle,  $\theta$ , between the line joining two vertices, primary vertex (PV) and secondary vertex (SV), and the momentum vector shown in green. The momentum vector corresponds to the momentum of the three tracks contributing to the SV.

momentum  $p > 1000 \text{ MeV}/c$  and the transverse momentum is required to be  $p_T > 100 \text{ MeV}/c$ . Due to the relatively long lifetime of the  $B^+$  and  $D$  mesons, the decay products originate from a vertex that is displaced from the proton-proton collision vertex. Therefore it is possible that the trajectory that the decay products followed didn't pass through the collision position as shown in Fig. 4.3. A requirement is placed on the significance of the impact parameter between the track and the proton-proton collision vertex of  $\chi_{\text{IP}}^2 > 4$  to ensure all of the tracks used are inconsistent with originating at the primary interaction.

Loose requirements are placed on *Particle Identification* variables to ensure the tracks are of the required species. These are further tightened as detailed in Section 4.3.2. Incorrect track candidates created by combining unrelated VELO and tracking station hits are suppressed by requiring the ghost track probability,  $P_{\text{Ghost}}$ , to be less than 0.4.

The tracks passing these requirements are combined in pairs or threes to form the  $D_s^+$  and  $\phi$  or  $\bar{D}^0$  meson candidates. To create the  $\phi$  candidates (or  $K^+K^-$  pair), all combinations of two tracks are considered. The two tracks are assigned the mass hypothesis of kaons. Only pairs which pass within 0.5 mm of one another at their closest point are retained. A number of requirements are imposed to ensure the pair are consistent with coming from a  $\phi$  meson (or  $K^+K^-$  pair) originating from

a  $B^+$  meson decay. The vertex is required to be of good quality, the scalar sum of the transverse momentums must be greater than  $1000 \text{ MeV}/c$  and the flight distance significance is required to be  $\chi_{\text{FD}}^2 > 16$ . Additionally the cosine of the direction angle, defined to be the angle between the momentum vector and flight vector, is required to be  $\cos \theta > 0$ , preventing the vertex being backward of the PV when the momentum is in the forward direction. In the search for  $B^+ \rightarrow D_s^+ \phi$  mesons, the invariant mass of the two kaons is required to be within  $150 \text{ MeV}/c^2$  of the known  $\phi$  meson mass.

The  $\bar{D}^0$  candidates are selecting using a similar parameters, however the exact values of the cuts are changed to reflect the differences in properties of the  $\phi$  and  $\bar{D}^0$  mesons. The vertex quality and flight distance significance requirements are tightened to  $\chi^2/N_{\text{DOF}} < 10$  and  $\chi_{\text{FD}}^2 > 36$  respectively. Additionally, the scalar  $p_{\text{T}}$  sum requirement is increased to  $\sum |p_{\text{T}}| > 1800 \text{ MeV}/c$  as the decay products tend to have higher transverse momentum.

The  $D_s^+$  mesons for the signal and normalisation modes are created using combinations of three tracks given the mass hypotheses kaons or pions, depending on the decay mode being constructed. All combinations of three tracks are considered and only those in which all three tracks are within 0.5 mm of one another are retained. Similar to the  $\bar{D}^0$  meson the vertex quality, flight distance significance and scalar  $p_{\text{T}}$  sum requirements are  $\chi^2/N_{\text{DOF}} < 10$ ,  $\chi_{\text{FD}}^2 > 36$  and  $\sum |p_{\text{T}}| > 1800 \text{ MeV}/c$  respectively.

The  $B^+$  meson candidates are constructed by creating all possible combinations of the  $D_s^+$  and  $\phi$  or  $\bar{D}^0$  meson candidates in each event. Requirements are placed on these combination to select just those consistent with  $B^+$  mesons; the combination is required to have a lifetime  $\tau_{B^+} > 0.2 \text{ ps}$  and direction angle  $\cos \theta > 0.999$ . The impact parameter significance is required to be  $\chi_{\text{IP}}^2 < 25$  to ensure the  $B^+$  meson originated at the primary interaction. The vertex is required to have a quality of  $\chi^2/N_{\text{DOF}} < 10$ . More requirements are additionally placed on the decay products that contribute to the  $B^+$  meson. The selection requirements imposed on candidate  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays in their respective *Stripping Lines* are detailed in Table 4.6.

Two slightly different strategies are used for the normalisation channel selection in the search for  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  events. In the former, a dedicated  $B^+ \rightarrow D_s^+ \bar{D}^0$  *Stripping Line* is used to reconstruct the normalisation channel decays. The *Stripping Line* selection for this line is listed in Table 4.7.

The *Stripping Line* used in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays covers the full  $m(K^+ K^-)$  phase-space. This includes the  $\bar{D}^0$  mass such that this line reconstructs both the signal and normalisation channels simultaneously. Both modes are selected using this line to reduce systematic uncertainty in the ratio of selection efficiencies.

Particles	Quantity	Requirement
$B^+$	Mass	$4750 < m(D_s^+ \phi) < 7000 \text{ MeV}/c^2$
	Products $p_T$ scalar sum	$\sum  p_T  > 5000 \text{ MeV}/c$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Lifetime	$\tau_{B^+} > 0.2 \text{ ps}$
	Impact parameter significance	$\chi^2_{\text{IP}} < 25$
	Direction angle	$\cos \theta > 0.999$
	$>0$ decay products with:	
	Momentum	$p > 10000 \text{ MeV}/c$
	Transverse momentum	$p_T > 1700 \text{ MeV}/c$
	Impact parameter significance	$\chi^2_{\text{IP}} > 16$
	Impact parameter	$\text{IP} > 0.1 \text{ mm}$
	$>1$ decay products with:	
	Momentum	$p > 5000 \text{ MeV}/c$
	Transverse momentum	$p_T > 500 \text{ MeV}/c$
$D_s^+$	Mass	$1770 < m(h^+ h^- h^+) < 2068 \text{ MeV}/c^2$
	Products $p_T$ scalar sum	$\sum  p_T  > 1800 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(h^+, h^-) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^-, h'^+) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^+, h'^+) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Flight distance significance	$\chi^2_{\text{FD}} > 36$
$\phi$	Mass (only for $B^+ \rightarrow D_s^+ \phi$ )	$ m(K^+ K^-) - m_\phi  < 150 \text{ MeV}/c^2$
	Products $p_T$ scalar sum	$\sum  p_T  > 1000 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(K^+, K^-) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 16$
	Flight distance significance	$\chi^2_{\text{FD}} > 16$
$K^\pm (\pi^\pm)$	Track quality	$\chi^2/N_{\text{DOF}} < 4.0$
	Transverse momentum	$p_T > 100 \text{ MeV}/c$
	Momentum	$p > 1000 \text{ MeV}/c$
	Impact parameter significance	$\chi^2_{\text{IP}} > 4$
	Ghost track probability	$P_{\text{Ghost}} < 0.4$
	Particle identification	$\text{PIDK} > -10 \text{ (PIDK} < 20)$

Table 4.6: Selection requirements for  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates.

Particles	Quantity	Requirement
$B^+$	Mass	$4750 < m(D_s^+ \bar{D}^0) < 7000 \text{ MeV}/c^2$
	Products $p_T$ scalar sum	$\sum  p_T  > 5000 \text{ MeV}/c$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Lifetime	$\tau_{B^+} > 0.2 \text{ ps}$
	Impact parameter significance	$\chi_{\text{IP}}^2 < 25$
	Direction angle	$\cos \theta > 0.999$
	<i>&gt;0 decay products with:</i>	
	Momentum	$p > 10000 \text{ MeV}/c$
	Transverse momentum	$p_T > 1700 \text{ MeV}/c$
	Impact parameter significance	$\chi_{\text{IP}}^2 > 16$
	Impact parameter	$\text{IP} > 0.1 \text{ mm}$
	<i>&gt;1 decay products with:</i>	
	Momentum	$p > 5000 \text{ MeV}/c$
	Transverse momentum	$p_T > 500 \text{ MeV}/c$
$D_s^+$	Mass	$1770 < m(h^+ h^- h^+) < 2068 \text{ MeV}/c^2$
	Products $p_T$ scalar sum	$\sum  p_T  > 1800 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(h^+, h^-) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^-, h'^+) < 0.5 \text{ mm}$
	Distance of closest approach	$\text{DOCA}(h^+, h'^+) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Flight distance significance	$\chi_{\text{FD}}^2 > 36$
$\bar{D}^0$	Mass	$1765 < m(h^+ h^- h^+) < 1965 \text{ MeV}/c^2$
	Products $p_T$ scalar sum	$\sum  p_T  > 1800 \text{ MeV}/c$
	Distance of closest approach	$\text{DOCA}(K^+, K^-) < 0.5 \text{ mm}$
	Direction angle	$\cos \theta > 0$
	Vertex quality	$\chi^2/N_{\text{DOF}} < 10$
	Flight distance significance	$\chi_{\text{FD}}^2 > 36$
$K^\pm (\pi^\pm)$	Track quality	$\chi^2/N_{\text{DOF}} < 4.0$
	Transverse momentum	$p_T > 100 \text{ MeV}/c$
	Momentum	$p > 1000 \text{ MeV}/c$
	Impact parameter significance	$\chi_{\text{IP}}^2 > 4$
	Ghost track probability	$P_{\text{Ghost}} < 0.4$
	Particle identification	$\text{PIDK} > -10 \text{ (PIDK} < 20)$

Table 4.7: Selection requirements for  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates.

### 4.3.2 Particle identification requirements

Particle identification variables help to determine the species of tracks passing through the LHCb detector. Using information from the RICH sub-detectors, the likelihood of different mass hypotheses are compared to the pion hypothesis. Loose requirements are made on the kaon hypothesis PID variable to reduce the contribution from other types of hadrons and background from other  $b$ -hadron decays with misidentified hadrons. The parameter is defined to be

$$\text{PIDK} = \Delta \log(K - \pi) = \log \mathcal{L}(K) - \log \mathcal{L}(\pi) \quad (4.1)$$

where  $\mathcal{L}(K)$  and  $\mathcal{L}(\pi)$  are the likelihoods of the kaon and pion hypotheses respectively. The requirements placed on each decay product are listed in Table 4.8.

Decay mode	Species	PID requirement
$\phi \rightarrow K^+ K^-$	$K^+$	$\text{PIDK} > 0$
	$K^-$	$\text{PIDK} > 0$
$\bar{D}^0 \rightarrow K^+ K^-$	$K^+$	$\text{PIDK} > 0$
	$K^-$	$\text{PIDK} > 0$
$D_s^+ \rightarrow K^+ K^- \pi^+$	$K^+$	$\text{PIDK} > -5$
	$K^-$	$\text{PIDK} > -5$
	$\pi^+$	$\text{PIDK} < 5$
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	$\pi^+$	$\text{PIDK} < 5$
	$\pi^-$	$\text{PIDK} < 5$
	$\pi^+$	$\text{PIDK} < 5$
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	$K^+$	$\text{PIDK} > -5$
	$\pi^-$	$\text{PIDK} < 5$
	$\pi^+$	$\text{PIDK} < 5$

Table 4.8: Particle identification requirements applied to kaons and pions.

### 4.3.3 Charmless and single-charm backgrounds

Decays of  $B^+$  mesons that didn't proceed via  $D$  mesons could form a peaking background below the signal invariant mass distributions when they decay to the same final state. The signal mode could receive contributions from the decays  $B^+ \rightarrow h^+ h^- h^+ \phi$  or  $B^+ \rightarrow h^+ h^- h^+ h^-$ , referred to as charmless backgrounds. Here  $h^\pm$  is used to

represent  $K^\pm$  or  $\pi^\pm$  in the specific  $D_s^+$ ,  $\bar{D}^0$  or  $\phi$  final state. The normalisation mode is also susceptible, however as it involves two charm mesons it could receive contributions from the decays  $B^+ \rightarrow h^+ h^- h^+ \bar{D}^0$  or  $B^+ \rightarrow D_s^+ h^+ h^-$ , referred to as single-charm backgrounds, and  $B^+ \rightarrow h^+ h^- h^+ h^-$  referred to as a charmless background. These backgrounds can be suppressed by requiring the  $D$  meson decay vertex to be displaced from the  $B^+$  meson decay vertex. Requirements are applied to the significance of the vertex separation ( $\chi_{\text{FD}}^2$ ).

The residual yields of charmless backgrounds in the signal mode are estimated by performing a fit to the  $B^+$  invariant mass for candidates with  $25 < |m(h^+ h^- h^+) - m(D_s^+)| < 50 \text{ MeV}/c^2$ . This background estimation is performed separately for the  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  searches.

For the  $B^+ \rightarrow D_s^+ \bar{D}^0$  normalisation channel, a two-dimensional optimisation is performed to calculate the contribution from decays without a  $D_s^+$  meson,  $\bar{D}^0$  meson or both. The two-dimensional space defined by the  $D_s^+$  and  $\bar{D}^0$  masses is split into four types of area as shown in Fig 4.6.

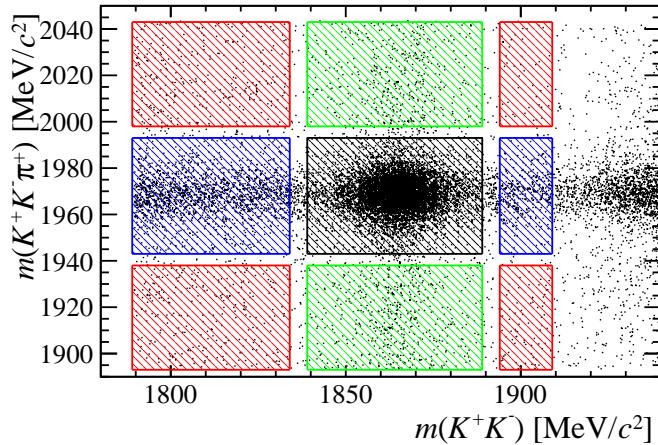


Figure 4.6: Two dimensional normalisation

1. Areas in which only  $B^+ \rightarrow h^+ h^- h^+ h^-$  decays contribute (red).
2. Areas in which either  $B^+ \rightarrow D_s^+ h^+ h^-$  or  $B^+ \rightarrow h^+ h^- h^+ h^-$  decays can contribute (blue).

3. Areas in which either  $B^+ \rightarrow h^+h^-h^+\bar{D}^0$  or  $B^+ \rightarrow h^+h^-h^+h^-$  decays can contribute (green).
4. The signal region in which  $B^+ \rightarrow D_s^+\bar{D}^0$ ,  $B^+ \rightarrow h^+h^-h^+\bar{D}^0$ ,  $B^+ \rightarrow D_s^+h^+h^-$  or  $B^+ \rightarrow h^+h^-h^+h^-$  decays could contribute (black).

Asymmetric  $\bar{D}^0$  sidebands are used to prevent misidentified  $B^+ \rightarrow D_s^+(\bar{D}^0 \rightarrow K^-\pi^+)$  decays from being included in the sideband sample. The optimal selection requirements are chosen such that the maximal signal efficiency is achieved for a residual charmless contribution of 2% of the normalisation yield.

The optimisation of the signal and normalisation cuts is performed separately for each different  $D_s^+$  decay mode, and for the  $B^+ \rightarrow D_s^+K^+K^-$  and  $B^+ \rightarrow D_s^+\phi$  selections. The optimised requirements are listed in Table 4.9, along with the estimated residual yields of charmless and single charm yields in the signal region.

$B^+$ decay mode	$D_s^+$ decay mode	$\chi_{\text{FD}}^2(D_s^+)$	$\chi_{\text{FD}}^2(\bar{D}^0)$	Residual yields
$B^+ \rightarrow D_s^+\phi$	$D_s^+ \rightarrow K^+K^-\pi^+$	0.0	-	0.0
$B^+ \rightarrow D_s^+\phi$	$D_s^+ \rightarrow K^+\pi^-\pi^+$	25.0	-	2.6
$B^+ \rightarrow D_s^+\phi$	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	5.0	-	0.0
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow K^+K^-\pi^+$	8.0	0.0	21.6
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow K^+\pi^-\pi^+$	18.0	0.0	3.3
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	16.0	0.0	3.9
$B^+ \rightarrow D_s^+K^+K^-$	$D_s^+ \rightarrow K^+K^-\pi^+$	5.0	-	0.19
$B^+ \rightarrow D_s^+\bar{D}^0$	$D_s^+ \rightarrow K^+K^-\pi^+$	8.0	0.0	7.95

Table 4.9: Charmless and single charm minimum flight distance significance requirements applied to the  $D_s^+$  and  $\bar{D}^0$  candidates.

#### 4.3.4 Misidentified $D$ and $\Lambda_c^+$ hadrons

It is possible for the samples  $D_s^+$  mesons to be contaminated by other misidentified decays of  $D^+$  mesons or  $\Lambda_c^+$  baryons in which one of the decay products has been incorrectly identified. The invariant mass of the  $D_s^+$  meson is recalculated, swapping the mass hypothesis of the ambiguous track to that of the  $K$ ,  $\pi$  or  $p$ , depending on the decay mode. The particle identification requirements are tightened within a mass

window around the  $D^+$  or  $\Lambda_c^+$  mass, effectively removing this crossfeed. For the mode  $D_s^+ \rightarrow K^+ K^- \pi^+$ , the vetoes are not applied to candidates for which  $|m(K^- K^+) - m_\phi| < 10 \text{ MeV}/c^2$  as there are a high purity of  $D_s^+ \rightarrow K^+ K^- \pi^+$  decays in this region.

The specific vetoes included in this selection are listed in Table 4.10. The invariant

Decay Mode	Misidentified decay
$D_s^+ \rightarrow \textcolor{red}{K}^+ K^- \pi^+$	$D^+ \rightarrow \pi^+ K^- \pi^+$
	$\Lambda_c^+ \rightarrow \textcolor{red}{p} K^- \pi^+$
$D_s^+ \rightarrow \textcolor{red}{K}^+ \pi^- \pi^+$	$D^+ \rightarrow \pi^+ \pi^- \pi^+$

Table 4.10: Misidentified decays targeted by vetoes. The ambiguous track is highlighted in red in each case.

mass distributions for each the misidentified  $D_s^+ \rightarrow K^+ K^- \pi^+$  decays are shown with and without the MVA requirements in Figs. 4.7 and 4.8 for both the signal  $B^+ \rightarrow D_s^+ \phi$  and normalisation  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays.

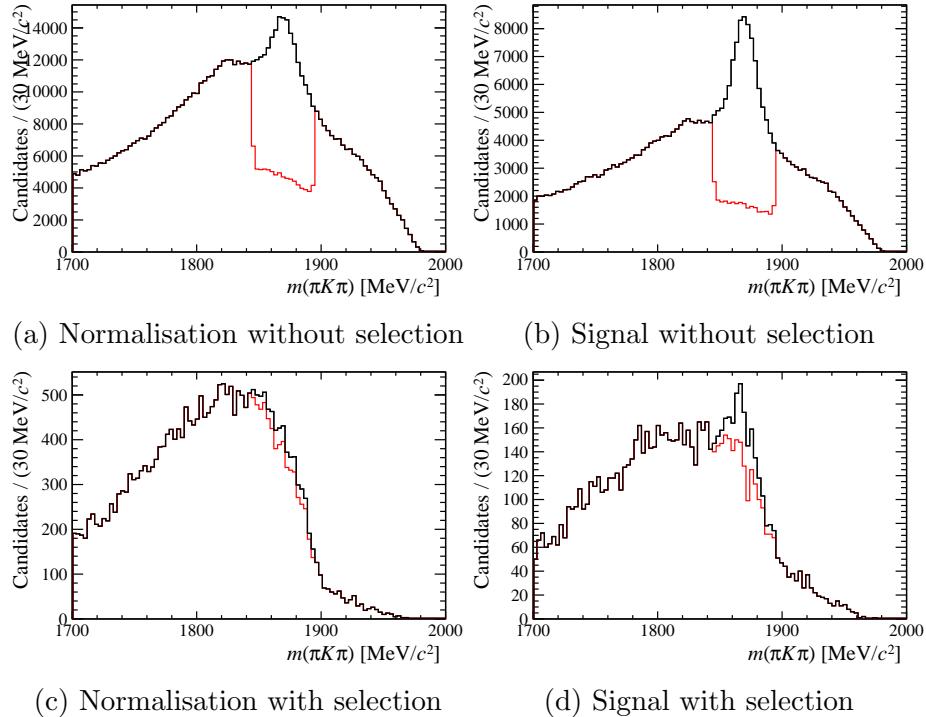


Figure 4.7: Invariant mass distributions of  $D_s^+ \rightarrow K^+ K^- \pi^+$  samples reconstructed as  $D^+ \rightarrow \pi^+ K^- \pi^+$  for the signal and normalisation samples. The samples are shown with (red) and without (black) the veto described in Sec. 4.3.4. The distributions are shown before (top) and after (bottom) the MVA requirements have been applied.

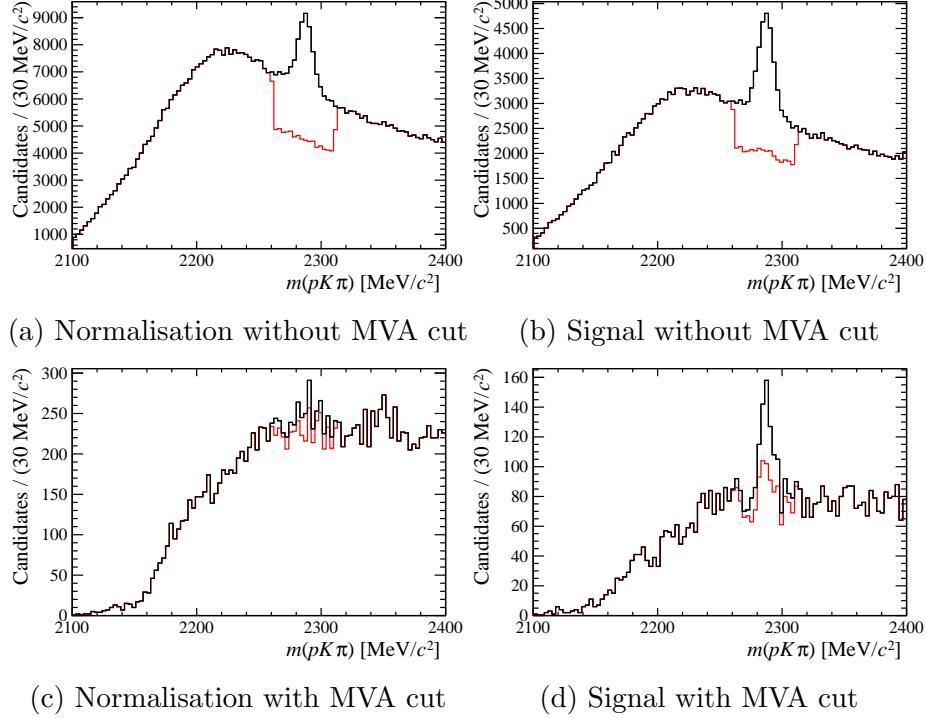


Figure 4.8: Invariant mass distributions of  $D_s^+ \rightarrow K^+ K^- \pi^+$  samples reconstructed as  $\Lambda_c^+ \rightarrow p K^- \pi^+$  for the signal and normalisation samples. The samples are shown with (red) and without (black) the veto described in Sec. 4.3.4. The distributions are shown before (top) and after (bottom) the MVA requirements have been applied.

#### 4.3.5 Invariant mass vetoes

Sharp peaking structures are observed in subsets of the final state particles. These are removed with simple invariant mass cuts to remove combinatorial or partially reconstructed backgrounds that result from these incorrectly reconstructed decays. For simplicity the final state particles for each mode are labelled with a number between 1–5 as described in Table 4.11.

Decay Mode	1	2	3	4	5
$B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) \phi$	$K^+$	$K^-$	$\pi^+$	$K^+$	$K^-$
$B^+ \rightarrow (D_s^+ \rightarrow \pi^+ \pi^- \pi^+) \phi$	$\pi^+$	$\pi^-$	$\pi^+$	$K^+$	$K^-$
$B^+ \rightarrow (D_s^+ \rightarrow K^+ \pi^- \pi^+) \phi$	$K^+$	$\pi^-$	$\pi^+$	$K^+$	$K^-$
$B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) K^+ K^-$	$K^+$	$K^-$	$\pi^+$	$K^+$	$K^-$

Table 4.11: Particle labels used when studying invariant mass vetoes for  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates.

All combinations of the final state particles that create a neutral or singly-charged candidate are investigated. Significant structures are observed for all three  $D_s^+$  decay modes in some combination.

The following vetos are applied to remove these incorrectly reconstructed decays.

- For the mode  $B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) \phi$ 
  - $|m(1245) - m(B_s^0)| > 50 \text{ MeV}/c^2$
  - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$
- For the mode  $B^+ \rightarrow (D_s^+ \rightarrow \pi^+ \pi^- \pi^+) \phi$ 
  - $|m(145) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(145) - m(D^+)| > 25 \text{ MeV}/c^2$
  - $|m(245) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(245) - m(D^+)| > 25 \text{ MeV}/c^2$
  - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$
- For the mode  $B^+ \rightarrow (D_s^+ \rightarrow K^+ \pi^- \pi^+) \phi$ 
  - $|m(245) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(245) - m(D^+)| > 25 \text{ MeV}/c^2$
  - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$

In the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays the increased size of the  $m(K^+ K^-)$  phase-space means more of the combinations of final state particles are susceptible to sharp peaking structure from incorrectly reconstructed backgrounds. Those spectra found to have significant peaking structures are additionally vetoed, as shown in Fig. 4.10.

- Vetoes for the mode  $B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) K^+ K^-$ :
  - $|m(1245) - m(B_s^0)| > 50 \text{ MeV}/c^2$
  - $|m(345) - m(D_s^+)| > 25 \text{ MeV}/c^2$  and  $|m(345) - m(D^+)| > 25 \text{ MeV}/c^2$
  - $|m(135) - m(D_s^+)| > 25 \text{ MeV}/c^2$
  - $|m(234) - m(D_s^+)| > 25 \text{ MeV}/c^2$

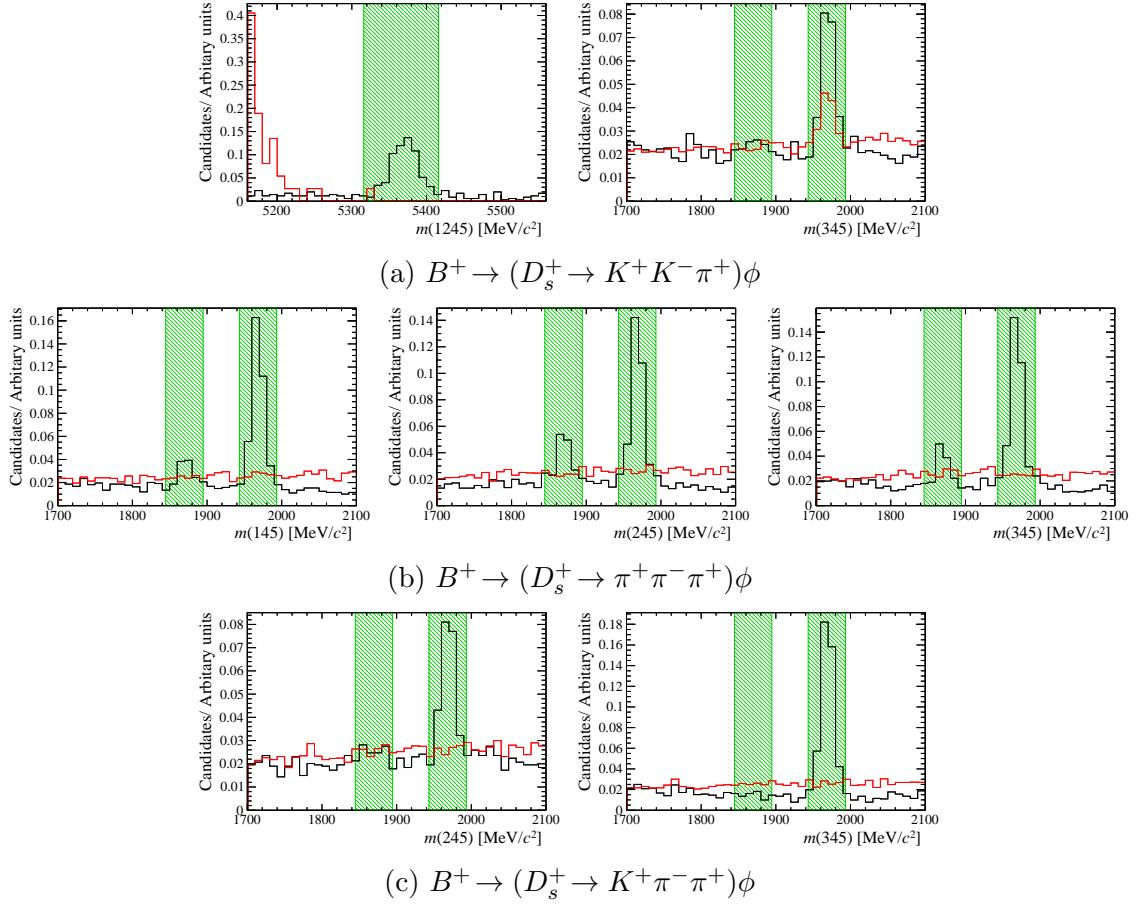


Figure 4.9: Invariant mass distributions for subsets of decay products in data (black) and simulation (red). The green region show the regions removed by the vetoes listed in Sec 4.3.5.

Another set of vetoes rejects decays where the tracks forming the  $D_s^+$  candidate originate from an excited charged charm meson decay, for example  $D^{*+} \rightarrow (D^0 \rightarrow h^+ h'^-) \pi^+$ . By requiring  $\Delta m = m(h^+ h'^-) - m(h^+ h'^-) > 150 \text{ MeV}/c^2$  decays of this type are efficiently removed. These are applied to both the signal and normalisation modes for all  $D_s^+$  decays.

#### 4.3.6 Normalisation mode veto

In the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays, the entire  $m(K^+ K^-)$  phasespace is used. This ranges from the  $K^+ K^-$  mass threshold at around  $990 \text{ MeV}/c^2$  to the kinematic limit at  $m(B^+) - m(D_s^+) = 3300 \text{ MeV}/c^2$ . This range is wide enough to include the mass of the  $\bar{D}^0$  meson,  $m(D^0) = 1864 \text{ MeV}/c^2$ . Consequently, when inspecting the

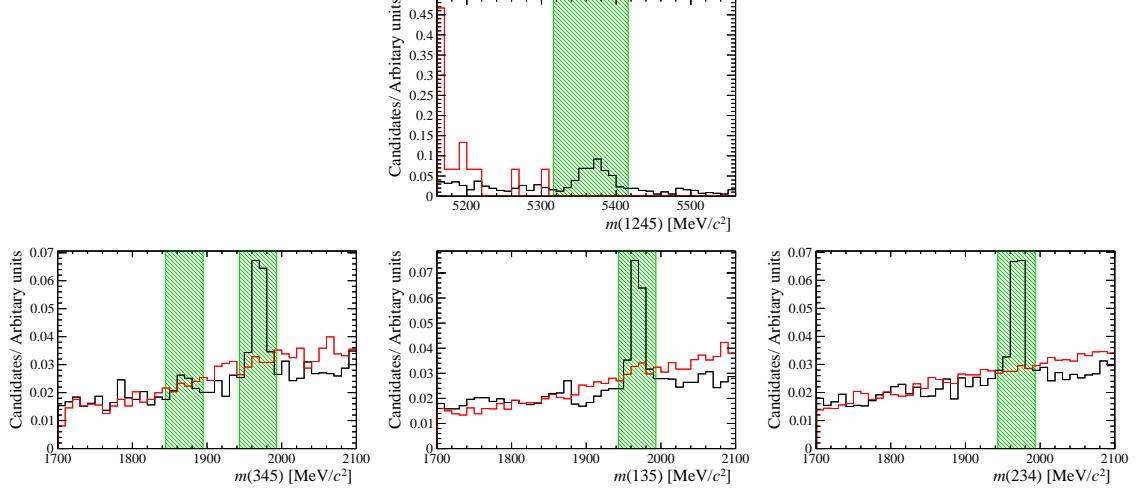


Figure 4.10: Invariant mass distributions for subsets of decay products for  $B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) K^+ K^-$  decays in data (black) and simulation (red). The green region show the regions removed by the vetoes listed in Sec 4.3.5.

$m(K^+ K^-)$  spectrum for selected signal candidates there is an excess of events at the  $D^0$  mass. It is necessary to remove these from the signal samples as, unsurprisingly, they result in a peak at the  $B^+$  mass in the  $m(D_s^+ K^+ K^-)$  spectrum and lead to an incorrect signal yield. As described in Section 4.3.1, these removed events are used as the normalisation channel for the measurement of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays. The region affected by the veto  $|m(K^+ K^-) - m(\bar{D}^0)| > 25 \text{ MeV}/c^2$  is shown in Fig. 4.11a.

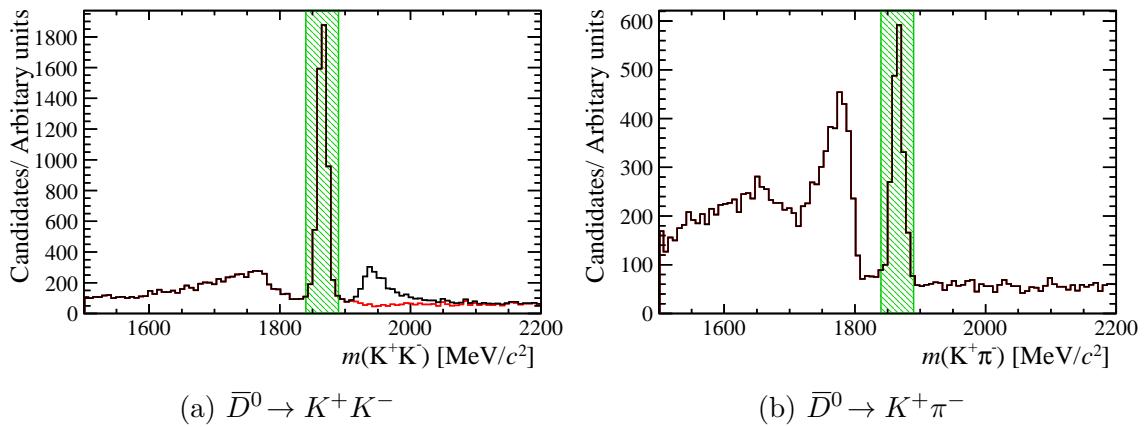


Figure 4.11: Normalisation channel vetoes.

In addition to the correctly reconstructed normalisation channel, the presence of the incorrectly reconstructed  $B^+ \rightarrow D_s^+(\bar{D}^0 \rightarrow K^+ \pi^-)$  decay is observable in

Fig. 4.11a. This appears as smeared out peak to the right of the  $\bar{D}^0$  peak. Although the probability of the  $\pi^-$  meson being misidentified as a  $K^-$  meson is low, the branching fraction for  $\bar{D}^0 \rightarrow K^+ \pi^-$  is larger, leading to the observed excess. It is possible, and necessary, to remove this contribution. The  $K^+ K^-$  candidates are reconstructed again, swapping the mass hypothesis of the  $K^-$  to  $\pi^-$ . The distribution of these candidates in the vicinity of the  $\bar{D}^0$  mass is shown in Fig. 4.11b. A distinct peak is observed at the  $\bar{D}^0$  mass. This contribution is removed by the requirement  $|m(K^+ \pi^-) - m(\bar{D}^0)| > 25 \text{ MeV}/c^2$ . The effect of this requirement on the  $m(K^+ K^-)$  spectrum is shown by the red line in Fig. 4.11a, which represents the sample with this veto applied. This structure is effectively removed.

Another structure appears to be present to the left of the  $\bar{D}^0$  peak in Fig. 4.11a. This is likely to be due to  $\bar{D}^0 \rightarrow K^+ K^- X$  decays in which one or more particles have not been reconstructed. This partially reconstructed background would not peak at the  $B^+$  mass in the  $m(D_s^+ K^+ K^-)$  spectrum due to the missing particles, therefore no attempt is made to remove this contribution.

### 4.3.7 Multivariate analysis

Multivariate Analyses (MVAs) are used to help discriminate between genuine  $D_s^+$  and  $\phi$  meson decays and combinations of unrelated tracks. These MVAs are trained using large samples of candidates from other  $B$  mesons decays with similar topologies reconstructed in data. This data-driven approach can benefit from an expanded set of variables that are not perfectly represented in simulation, including track quality and particle identification information, in addition to the widely used kinematic and geometric properties. The method is based on the approach used in Ref. [111], however the choice of input variables has been optimised and training samples expanded to include Run II data.

The sample of  $D_s^+$  mesons is obtained from the relatively abundant  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decay. Similarly, the sample of  $\phi$  mesons is obtained from  $B_s^0 \rightarrow J/\psi \phi$  decays. Large, high purity samples are reconstructed using similar requirements to those applied in the selection of signal  $D_s^+$  and  $\phi$  mesons. A sample is selected for each of the  $D_s^+$  and

$\phi$  meson decays uses in this analysis, as listed in Table 4.12. The training of separate MVAs for the different  $D_s^+$  modes allows the use of particle identification variables that separate kaons and pions to be fully exploited. The MVA for  $\phi \rightarrow K^+K^-$  decays is used to select both  $\phi$  and  $\bar{D}^0$  mesons. For the normalisation mode this may be suboptimal, however it ensures the selection of the signal and normalisation channels are almost identical such that the systematic uncertainty in the ratio of efficiencies is minimised. The samples of  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  and  $B_s^0 \rightarrow J/\psi \phi$  decays are randomly split into two subsamples. The first is used to train the MVAs and the second is used to determine the efficiency of the selection. To prevent the difference between  $\phi$  and  $\bar{D}^0$  decays from affecting the ratio of efficiencies, the normalisation channel MVA efficiency is instead determined from a dedicated sample of  $\bar{D}^0 \rightarrow K^+K^-$  decays as detailed in Table 4.12. The  $\phi \rightarrow K^+K^-$  MVA is also used to select the  $K^+K^-$  pair in  $B^+ \rightarrow D_s^+ K^+ K^-$  decays. A total of eight MVAs are trained, one for each of the  $D_s^+$  and  $\phi$  modes, in both Run I and Run II. Changes to the particle identification variables between the two running periods necessitates separate trainings.

Sample	Mode	Use
$\bar{B}_s^0 \rightarrow D_s^+ \pi^-$	$D_s^+ \rightarrow K^+ K^- \pi^+$	Training, Efficiency
$\bar{B}_s^0 \rightarrow D_s^+ \pi^-$	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	Training, Efficiency
$\bar{B}_s^0 \rightarrow D_s^+ \pi^-$	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	Training, Efficiency
$B_s^0 \rightarrow J/\psi \phi$	$\phi \rightarrow K^+ K^-$	Training, Efficiency
$B^+ \rightarrow \bar{D}^0 \pi^+$	$\bar{D}^0 \rightarrow K^+ K^-$	Efficiency

Table 4.12: The decay modes used to train and determine the efficiency of the various MVAs used in this analysis.

## Preselection

Before the samples of  $D_s^+$  and  $\phi$  mesons are used to train the MVAs, precautions are taken to ensure the samples are representative of the  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  signal decays. The  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  and  $B_s^0 \rightarrow J/\psi \phi$  decays are selected using a similar procedure to the signal modes. Firstly, *Stripping Lines* reconstruct the candidates. The  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decays are built using the same software module as the signal and

normalisation channel. Therefore, the selection requirements for  $D_s^+$  candidates in  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decays are identical to those listed for the signal in Tables 4.6 and 4.7. The  $B_s^0 \rightarrow J/\psi \phi$  decays are reconstructed using  $J/\psi \rightarrow \mu^+ \mu^-$  decays, therefore they are built using a different software module as the final state is not fully hadronic. As such there are some differences in the selection requirements for the  $\phi$  mesons from the two sources. In general the selection requirements for the  $B_s^0 \rightarrow J/\psi \phi$  channel are looser as the presence of the muons allows an efficient triggering without the need to make tight selections on the  $K^+ K^-$  pair. A direct comparison of the relevant quantities are listed in Table 4.13.

Particle	Quantity	Signal	Control
$\phi$	Mass minimum	$870 \text{ MeV}/c^2$	$980 \text{ MeV}/c^2$
	Mass maximum	$1170 \text{ MeV}/c^2$	$1050 \text{ MeV}/c^2$
	Transverse Momentum	-	$p_T > 500 \text{ MeV}/c$
	Products $p_T$ scalar sum	$\sum  p_T  > 1000 \text{ MeV}/c$	-
	DOCA( $K^+, K^-$ )	$< 0.5 \text{ mm}$	-
	Direction angle	$\cos \theta > 0$	-
	Vertex quality	$\chi^2/N_{\text{DOF}} < 16$	$\chi^2/N_{\text{DOF}} < 25$
	Flight distance significance	$\chi^2_{\text{FD}} > 16$	-
$K^\pm$	Track quality	$\chi^2/N_{\text{DOF}} < 4.0$	$\chi^2/N_{\text{DOF}} < 5.0$
	Transverse momentum	$p_T > 100 \text{ MeV}/c$	-
	Momentum	$p > 1000 \text{ MeV}/c$	-
	Impact parameter significance	$\chi^2_{\text{IP}} > 4$	-
	Ghost track probability	$P_{\text{Ghost}} < 0.4$	-
	Particle identification	$\text{PIDK} > -10$	$\text{PIDK} > 0$

Table 4.13: The *Stripping Line* requirements for  $\phi \rightarrow K^+ K^-$  candidates in the signal and MVA training (control) mode selection. All requirements are looser for the control channel with the exception of the particle identification requirements.

The difference in the two selections could be potentially biasing when calculating the efficiency of the MVA, therefore the requirements are tightened on the half of the  $B_s^0 \rightarrow J/\psi \phi$  sample used to calculate the efficiency. The looser set of requirements are still used when training the MVA methods to maximise the sample sizes to help prevent overtraining. As detailed in Section 4.3.7, when training MVA the  $\phi$  meson invariant mass sidebands in the  $B_s^0 \rightarrow J/\psi \phi$  sample are used as the background

sample. It was found that using the tighter set of requirements led to a negligible amount of background candidates, resulting in overtraining.

The second step of preselection aims to apply the same sequence of requirements to the modes used to train the MVAs as are applied to the signal and normalisation before the MVAs are applied. These are made up of the same trigger, background veto and PID requirements. For  $D_s^+$  candidates from  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decays, the  $\bar{B}_s^0$  meson is required to either be **TOS** with respect to the **L0Hadron** trigger or **TIS** with respect to **L0Global** as with the signal. The  $\bar{B}_s^0$  candidates are then required to be **TOS** with respect to the same HLT1 and HLT2 triggers as used with the signal modes.

For the  $B_s^0 \rightarrow J/\psi \phi$  candidates, a large fraction are triggered due to the muons from the  $J/\psi$  decay. To ensure the sample is reflective of the signal decays the  $B_s^0$  meson is required to be **TOS** with respect to the **L0Hadron** trigger or **TIS** with respect to **L0Global**. This effectively excludes events that have been selected solely due to the muons firing **L0Muon**. This consequently results in a large decrease in the available statistics.

As shown in Table 4.13, the MVA training mode has tighter particle identification requirements than the signal mode. Particle identification plays a crucial role in the efficacy of the MVA methods, therefore the signal PID requirements for the  $\phi$  meson decay products are tightened accordingly as discussed in Section 4.3.2.

Finally, the same misidentified  $D$  and  $A_c^+$  hadron vetoes as detailed in Section 4.3.4 are applied to the  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  sample. These help to ensure the samples of  $D_s^+$  candidates are free from contamination.

## Background subtraction

In order to use the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  samples to train MVAs, background subtracted distributions must be obtained for the variables used to discriminate the signal from backgrounds. Firstly, unbinned extended maximum likelihood fits are performed to the  $D_s^+$  and  $\phi$  invariant mass distribution in order to determine the yield of  $D_s^+$  and  $\phi$  candidates respectively. Maximum likelihood fits determine the values of the signal and background yields for which the given data set was most likely.

The likelihood,  $\mathcal{L}$ , is constructed from the probability density functions (PDFs) for the fit model,  $F(m, \vec{p})$ , for each entry  $i$  in the data set. The PDF for each entry is evaluated at the corresponding value of the invariant mass,  $m = m_i$ , such that  $\mathcal{L}$  is just a function of the PDF parameters,  $\vec{p}$ ,

$$\mathcal{L}(\vec{p}) = \prod_i^N F(m = m_i, \vec{p}). \quad (4.2)$$

The maximum value of  $\mathcal{L}(\vec{p})$  is achieved for the set of PDF parameters for which the data was most likely. It is computationally beneficial to instead compute the negative log-likelihood (NLL) rather than the likelihood directly, as addition is less intensive than multiplication. The NLL,

$$-\log \mathcal{L}(\vec{p}) = -\sum_i^N \log F(m = m_i, \vec{p}), \quad (4.3)$$

is then minimised with respect to the PDF parameters.

This likelihood can be *extended* to allow the total yield attributed to the PDF to be determined as a parameter as well. This is necessary to determine the yields of both signal and background contributions ( $n_S$  and  $n_B$ ). The likelihood is multiplied by the Poisson probability density for measuring  $n = n_S + n_B$  events,

$$\mathcal{L}(n, \vec{p}) = \frac{n^N e^{-n}}{N!} \prod_i^N F(m = m_i, \vec{p}) = \frac{e^{-n}}{N!} \prod_i^N nF(m = m_i, \vec{p}). \quad (4.4)$$

The extended NLL becomes

$$-\log \mathcal{L}(n, \vec{p}) = -\sum_i^N \log nF(m = m_i, \vec{p}) + n + \log N!. \quad (4.5)$$

The  $\log N!$  term can be ignored as it is a constant. The fit model is constructed from signal and background contributions

$$nF(m|n_S, n_B, \vec{p}', \vec{p}'') = n_S f(m|\vec{p}') + n_B g(m|\vec{p}''), \quad (4.6)$$

where  $n_S$  and  $n_B$  are the signal and background yields,  $f$  and  $g$  are the PDFs for the signal and background respectively, and  $\vec{p}'$  and  $\vec{p}''$  represent the parameters controlling the PDF shapes.

After the extended maximum likelihood fits have been performed and the yields established, the subtraction of candidates attributed to the background component is performed using the *sPlot* technique [112]. The *sPlot* method assigns weights for each species in the fit model, namely signal and background. These are assigned to each event in the data set and are only dependent on the value of the fitted variable, in this case  $\phi$  or  $D_s^+$  mass. When plotting other, uncorrelated variables and weighting each event accordingly, the contribution from each component in the fit can be extracted.

For both  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decays, simple fit models are found to be sufficient when performing the *sPlot* technique. For each mode the signal components are modelled with a single Gaussian probability density function

$$f(m|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \times e^{-\frac{(m-\mu)^2}{2\sigma^2}}, \quad (4.7)$$

where  $\mu$  and  $\sigma$  are the mean and width of the Gaussian distribution and  $m$  is the observable invariant mass. The parameters  $\mu$  and  $\sigma$  are allowed to vary freely. The background contributions to the  $\phi \rightarrow K^+ K^-$  decays are modelled using a Chebychev polynomial with two degrees of freedom

$$g(m|a, b) = a \times (2m^2 - 1) + b \times m + 1, \quad (4.8)$$

where  $a$  and  $b$  are parameters that can vary freely and  $m$  is the observable invariant mass. The  $\phi$  meson mass is close to the threshold for  $K^+ K^-$  pair production so this parametrisation has sufficient freedom to successfully model the increasing background shape. The background contribution to all three  $D_s^+$  decays are modelled using exponential functions with a single degree of freedom controlling the effective slope

$$g(m|c) = e^{-m \times c}, \quad (4.9)$$

where  $c$  is a freely varying parameter and  $m$  is the observable invariant mass. The fits to the  $\phi$  and  $D_s^+$  meson masses in the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  samples are performed separately for each year of data taking and polarity of the LHCb dipole magnet. These distributions are shown in Fig. 4.12 for each of the  $\phi$  and  $D_s^+$  decay

modes. The yields of candidates in each of the subsamples are tabulated in Table 4.14, along with the totals for each mode.

Mode	Year	MagDown	MagUp
$\phi \rightarrow K^+ K^-$	2011	$3190 \pm 170$	$2120 \pm 140$
	2012	$5800 \pm 200$	$5540 \pm 200$
	2015	$2240 \pm 110$	$1640 \pm 100$
	2016	$11200 \pm 300$	$11100 \pm 300$
	Total	$42830 \pm 600$	
$D_s^+ \rightarrow K^+ K^- \pi^+$	2011	$97100 \pm 500$	$68600 \pm 400$
	2012	$216600 \pm 800$	$212800 \pm 800$
	2015	$68600 \pm 400$	$43400 \pm 300$
	2016	$321400 \pm 900$	$310800 \pm 900$
	Total	$1339300 \pm 1900$	
$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	2011	$24300 \pm 500$	$17100 \pm 400$
	2012	$57700 \pm 800$	$56500 \pm 900$
	2015	$20400 \pm 500$	$13100 \pm 400$
	2016	$99200 \pm 1200$	$92100 \pm 1200$
	Total	$380400 \pm 2300$	
$D_s^+ \rightarrow K^+ \pi^- \pi^+$	2011	$13200 \pm 600$	$9500 \pm 400$
	2012	$29300 \pm 900$	$29300 \pm 900$
	2015	$10000 \pm 500$	$6200 \pm 400$
	2016	$47800 \pm 1200$	$44200 \pm 1100$
	Total	$189500 \pm 2300$	

Table 4.14: Yields of the samples used to train data-driven MVAs. The total yields are summed over all years and both magnet polarities.

## Input variables

The MVA method is trained using a large set of variables chosen to help discriminate between the signals of interest and combinations of unrelated tracks. These variables include the properties of the  $K^\pm$  or  $\pi^\pm$  decay products as well as those of the  $\phi$  or  $D_s^+$  candidate itself. Many of the quantities are the same as those previously defined in Section 4.3.1. The additional parameters relate to the track quality and particle identification information and are defined as follows:

**Track matching quality,  $\chi^2_{\text{TRKMATCH}}$ :** this parameter is a key component of the pattern recognition software that matches the VELO and tracking station

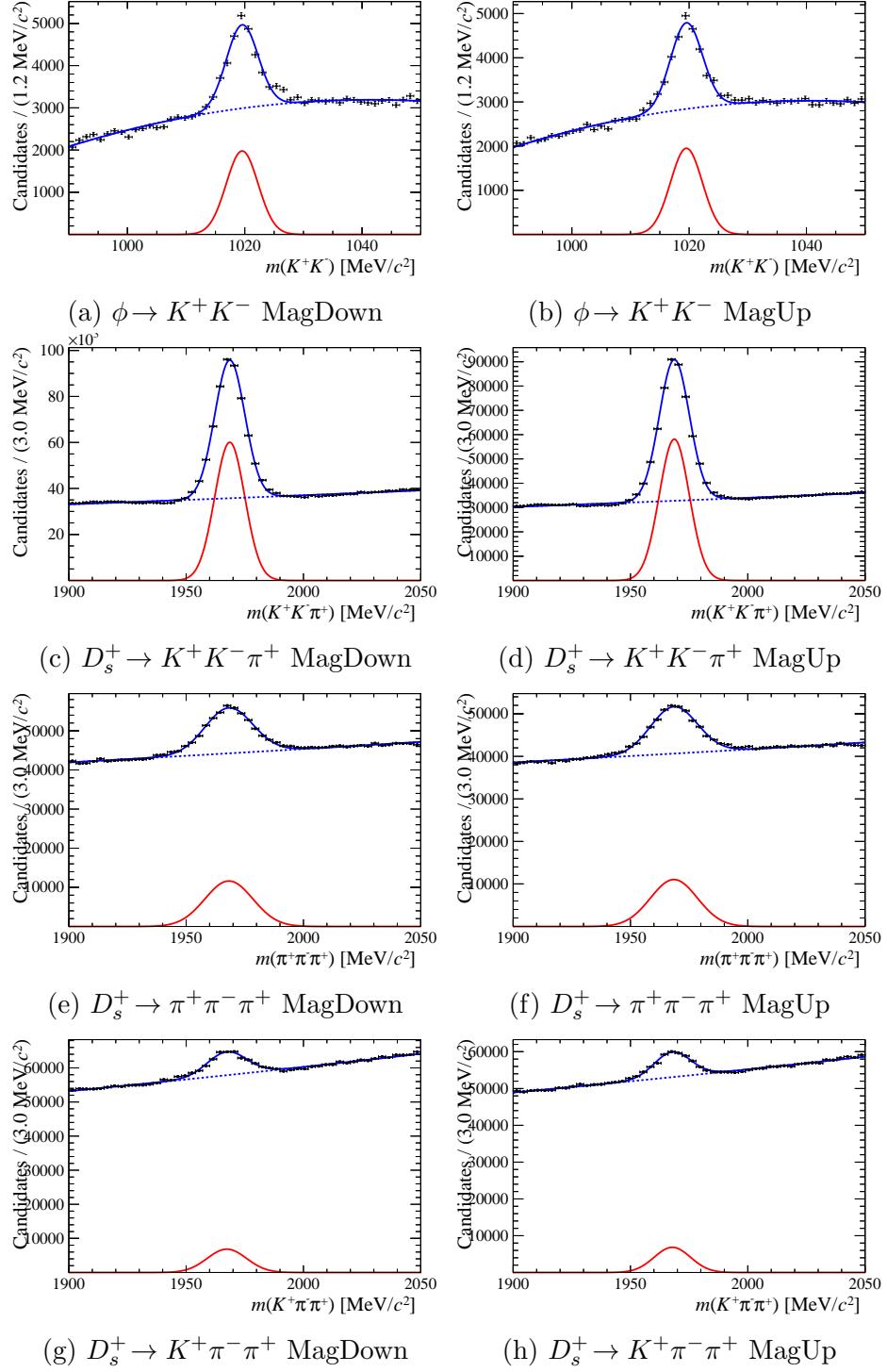


Figure 4.12: Unbinned maximum likelihood fit to  $B_s^0 \rightarrow J/\psi \phi$  (top) and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  (bottom three rows) candidates. These distributions correspond to just the 2016 data samples.

track sections together to form *long* tracks. It quantifies the quality of the track matching.

**VELO track quality,  $\chi^2_{\text{TRKVELO}}$ :** this quantifies the quality of the track fit using just the hits in the VELO.

**Kaon, proton, pion and ghost probabilities, ProbNNx:** these variables are an alternative form of *Particle Identification* variables in addition to those already discussed in Section 4.3.2. Each ProbNNx variable acts as the probability for the reconstructed track to be of the species x. These variables are the responses of Artificial Neural Networks trained to differentiate between the various particle types using a large number of inputs from different sub-detectors. The trainings use simulated samples of signal and background decays and have been retuned between Run I and Run II to reflect the difference in the detector configuration. Four of the available ProbNNx tuning are used in this analysis. Three of these identify kaons, protons and pions. The last aims to identify ghost tracks, similar to the  $P_{\text{Ghost}}$  variable already discussed.

The variables used in the  $\phi$  meson MVA are listed in Table 4.15. Some of the variables are input into the MVA training algorithm as their logarithm, rather than the variable directly. These variables tend to be rapidly increasing in at certain values, so using the logarithm aids visualisation. A total of 24 variables are used in this MVA.

When training the  $\phi$  meson MVA, the flight distance significance is not included as an input variable. This MVA is used to select both  $\phi$  and  $\bar{D}^0$  mesons which are likely to have different flight distance significance distributions as a result of their different lifetimes.

The variables used in the  $D_s^+$  MVA are listed in Table 4.16. These are largely very similar to those used for the  $\phi$  meson MVA, however it also includes the  $D_s^+$  meson's flight distance significance as, unlike the  $\phi$  meson, the  $D_s^+$  meson decays displaced from its production vertex. A total of 35 variables are used in these MVAs.

Particle	Description	Quantity
$\phi$	Momentum	$p$
	Transverse Momentum	$p_T$
	Vertex quality	$\log_{10}(\chi^2_{\text{VXT}})$
	Impact parameter significance	$\log_{10}(\chi^2_{\text{IP}})$
$K^\pm$	Momentum	$p$
	Transverse momentum	$p_T$
	Longitudinal momentum	$p_Z$
	Impact parameter significance	$\log_{10}(\chi^2_{\text{IP}})$
	Track quality	$\chi^2_{\text{TRK}}$
	Track matching quality	$\chi^2_{\text{TRKMATCH}}$
	Kaon probability	ProbNNk
	Proton probability	ProbNNp
	Pion probability	ProbNNpi
	Ghost probability	ProbNNghost

Table 4.15: Discriminating variables used to train the  $\phi$  MVA.

When training the MVA methods, the variables are ranked according to their discrimination power. Examples of these ranking are shown for the Run II  $\phi \rightarrow K^+ K^-$  MVA and Run II  $D_s^+ \rightarrow K^+ K^- \pi^+$  MVA in Table 4.17. It can be seen that for both the  $D_s^+$  and  $\phi$  meson MVAs that the kaon particle identification variable and various  $\log_{10}(\chi^2_{\text{IP}})$  variables are the most discriminating.

### MVA method

The MVA training and validation is implemented using the TMVA package, part of the ROOT framework [113]. During training, each method uses samples that have been defined to be ‘signal’ and ‘background’ in order to characterise the differences between the two categories. The multidimensional space defined by the input variables is effectively condensed onto a single axis given by the classifier’s response. Typically, high values of the response correspond to ‘signal’-like events and low values to ‘background’-like events. This training produces a *weights file* allowing the response to be calculated for other samples.

A second set of ‘signal’-like and ‘background’-like samples are used to blindly validate the training. The classifier response is compared in the training and validation

Particle	Description	Quantity
$D_s^+$	Momentum	$p$
	Transverse Momentum	$p_T$
	Vertex quality	$\log_{10}(\chi_{\text{VXT}}^2)$
	Impact parameter significance	$\log_{10}(\chi_{\text{IP}}^2)$
	Flight distance significance	$\log_{10}(\chi_{\text{FD}}^2)$
$K^\pm, \pi^\pm$	Momentum	$p$
	Transverse momentum	$p_T$
	Impact parameter significance	$\log_{10}(\chi_{\text{IP}}^2)$
	Track quality	$\chi_{\text{TRK}}^2$
	Velo track quality	$\chi_{\text{TRKVELO}}^2$
	Track matching quality	$\chi_{\text{TRKMATCH}}^2$
	Kaon probability	ProbNNk
	Proton probability	ProbNNp
	Pion probability	ProbNNpi
	Ghost probability	ProbNNghost

Table 4.16: Discriminating variables used to train the  $D_s^+$  MVAs.

samples to demonstrate if the expected separation is reproducible. This is important to determine if the trained method suffers from *overtraining*. This can happen when too few events are passed to the method, such that the statistical fluctuations in the distributions are significant. This can result in differences being identified that are actually artefacts of the low statistics and not representative of the sample as a whole. This results in the classifier being more discriminatory on the training sample than the validation sample.

The ‘signal’ and ‘background’ samples are both taken from the relevant  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  or  $B_s^0 \rightarrow J/\psi \phi$  decay samples. The signal samples include all events within the fit ranges shown in Fig. 4.12. Each candidate is weighted with the appropriate Gaussian fit component weight as determined by the *sPlot* technique. This weight is passed to the TMVA framework when training the methods. The background samples are comprised of the sideband regions of the relevant ranges in Fig. 4.12. These ranges are defined by  $|m(K^+ K^-) - m(\phi)| > 10 \text{ MeV}/c^2$  and  $|m(h^+ h^- h^+) - m(D_s^+)| > 30 \text{ MeV}/c^2$  for the  $\phi$  and  $D_s^+$  MVAs respectively. No weights are used for the background samples as the regions are known to be signal-deficient.

Rank	Variable	Importance	Rank	Variable	Importance
1	$K^+$ ProbNNk	$7.639 \times 10^{-2}$	1	$K^+$ ProbNNk	$6.001 \times 10^{-2}$
2	$K^-$ ProbNNk	$6.043 \times 10^{-2}$	2	$\phi \log_{10}(\chi^2_{\text{IP}})$	$6.001 \times 10^{-2}$
3	$K^+ \log_{10}(\chi^2_{\text{IP}})$	$5.469 \times 10^{-2}$	3	$K^- \log_{10}(\chi^2_{\text{IP}})$	$5.997 \times 10^{-2}$
4	$\pi^+ \log_{10}(\chi^2_{\text{IP}})$	$4.984 \times 10^{-2}$	4	$\phi \log_{10}(\chi^2_{\text{VTX}})$	$5.827 \times 10^{-2}$
5	$K^+ p$	$4.348 \times 10^{-2}$	5	$K^- \text{ProbNNpi}$	$5.826 \times 10^{-2}$
6	$D_s^+ \log_{10}(\chi^2_{\text{IP}})$	$4.100 \times 10^{-2}$	6	$K^+ \log_{10}(\chi^2_{\text{IP}})$	$5.761 \times 10^{-2}$
7	$\pi^+ p_T$	$3.953 \times 10^{-2}$	7	$K^+ \text{ProbNNghost}$	$5.427 \times 10^{-2}$
8	$K^+ p_T$	$3.937 \times 10^{-2}$	8	$K^- \text{ProbNNk}$	$5.404 \times 10^{-2}$
9	$\pi^+ \text{ProbNNpi}$	$3.881 \times 10^{-2}$	9	$K^+ \chi^2_{\text{TRK}}$	$5.359 \times 10^{-2}$
10	$D_s^+ \log_{10}(\chi^2_{\text{FD}})$	$3.788 \times 10^{-2}$	10	$K^- \chi^2_{\text{TRK}}$	$5.201 \times 10^{-2}$
11	$K^- \log_{10}(\chi^2_{\text{IP}})$	$3.735 \times 10^{-2}$	11	$K^+ \text{ProbNNpi}$	$5.162 \times 10^{-2}$
12	$D_s^+ \log_{10}(\chi^2_{\text{VTX}})$	$3.016 \times 10^{-2}$	12	$K^- \text{ProbNNghost}$	$4.985 \times 10^{-2}$
13	$K^- p$	$2.770 \times 10^{-2}$	13	$K^+ \text{ProbNNp}$	$4.486 \times 10^{-2}$
14	$K^+ \text{ProbNNp}$	$2.758 \times 10^{-2}$	14	$K^- \text{ProbNNp}$	$4.410 \times 10^{-2}$
15	$K^- \text{ProbNNp}$	$2.692 \times 10^{-2}$	15	$K^+ \chi^2_{\text{TRKMATCH}}$	$4.291 \times 10^{-2}$

Table 4.17: Ranking of variables for the 15 highest ranked variables used to train the  $D_s^+ \rightarrow K^+ K^- \pi^+$  (left) and  $\phi \rightarrow K^+ K^-$  (right) Run II MVAs.

The MVAs are trained using Gradient Boosted Decision Trees (BDTGs) [114]. This method was compared to a number of different Boosted Decision Tree (BDT) derivatives and found to be most effective.

The BDTG method categorises ‘signal’ and ‘background’ candidates by creating consecutive sets of questions called Decision Trees. Each question can have two possible answers and depends on the previous responses. Eventually the answer to a question results in a candidate being classes as ‘signal’ or ‘background’ rather than asking more questions. A definable number of decision trees are used to collectively categorise each candidate; the overall response combines the decisions of all trees. The trees are trained using *boosting*. This means candidates that are most likely to be misclassified are weighted higher, such that the later trees concentrate on these trickier areas. The boosting method utilised in the BDTG method is designed to be more robust and reproducible in noisy situations, for example with limited statistics. The BDTG method is used in training all MVAs.

The MVA classifier distributions for each of the eight MVAs trained are shown in Fig. 4.15. This figure shows both the signal and background samples, as well as the

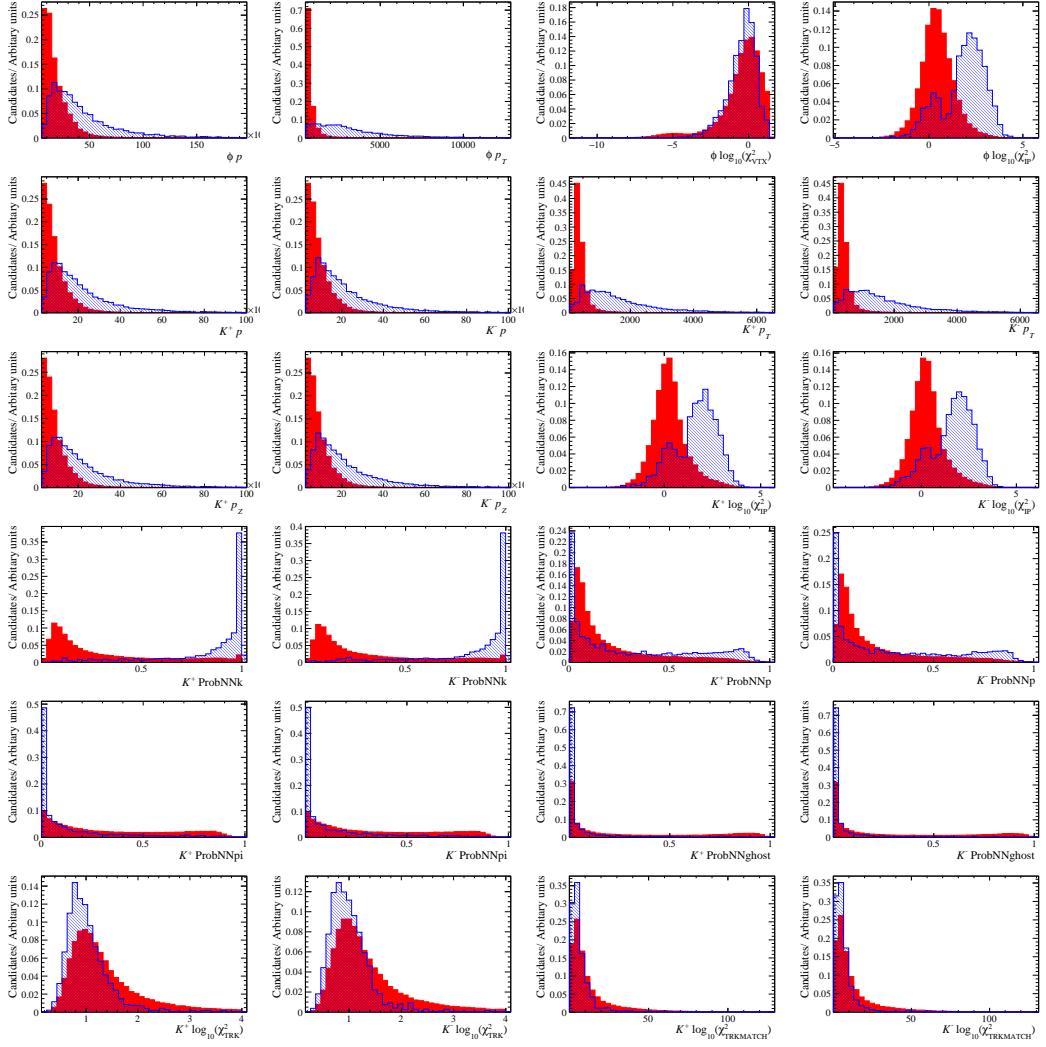


Figure 4.13: Distributions of MVA training variables in the signal (blue) and background (red) samples for the Run II  $\phi \rightarrow K^+K^-$  MVA.

training and validation subsamples. Each mode shows a good separation between the signal and background samples. For the majority of the distributions the training and testing samples have almost identical shapes, implying response is reproducible and the method has not overtrained. There are noticeable exceptions, however. The signal training and validation samples tend to show different distributions in the regions where the background distributions are maximal. This effect is likely to be a result of the use of weights in the MVA training. Ideally, the MVA classifier should remain positive across the whole range. The negative values implies that in that range the MVA classifier is correlated to the parameter used to generate the weights:

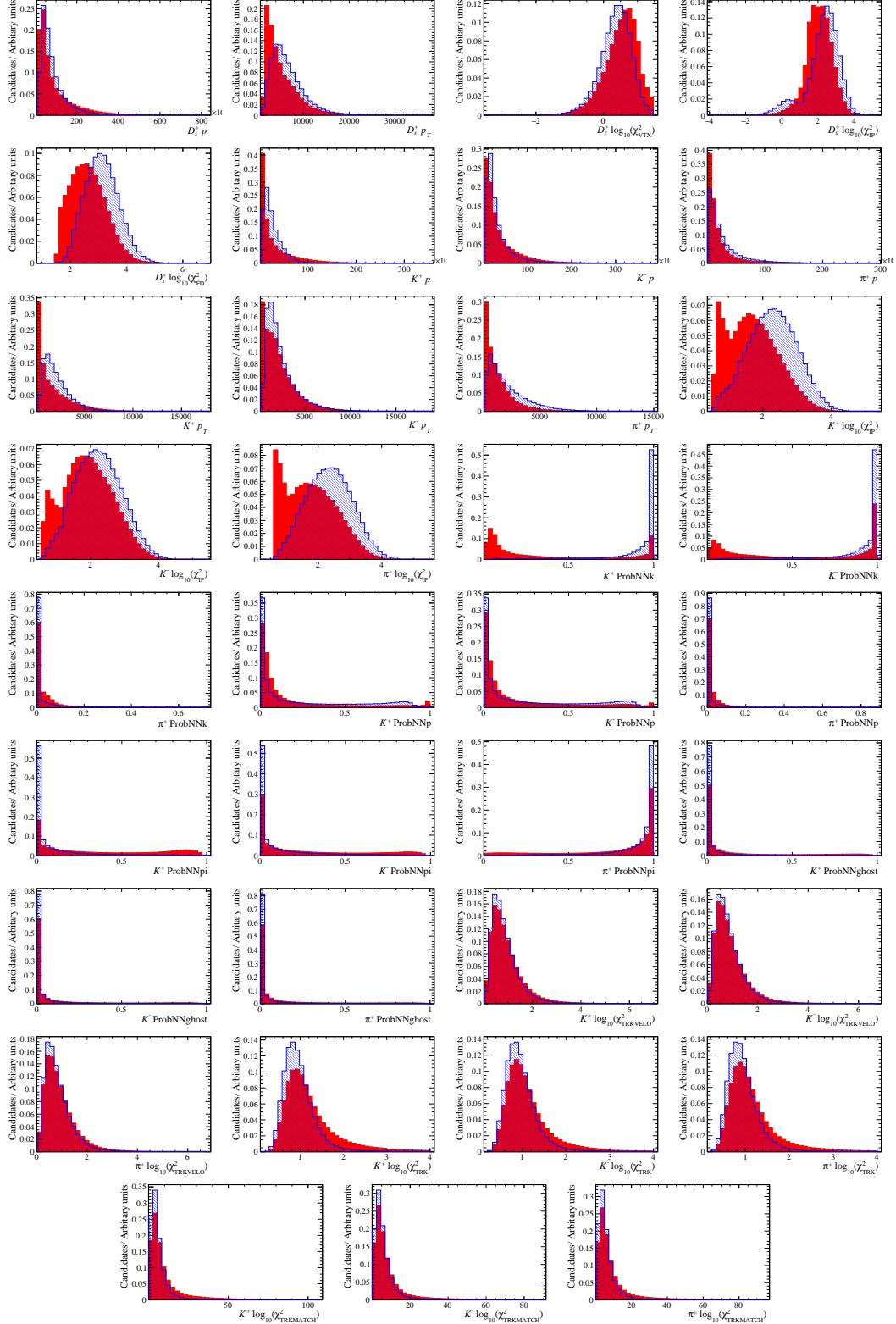


Figure 4.14: Distributions of MVA training variables in the signal (blue) and background (red) samples for the Run II  $D_s^+ \rightarrow K^+ K^- \pi^+$  MVA.

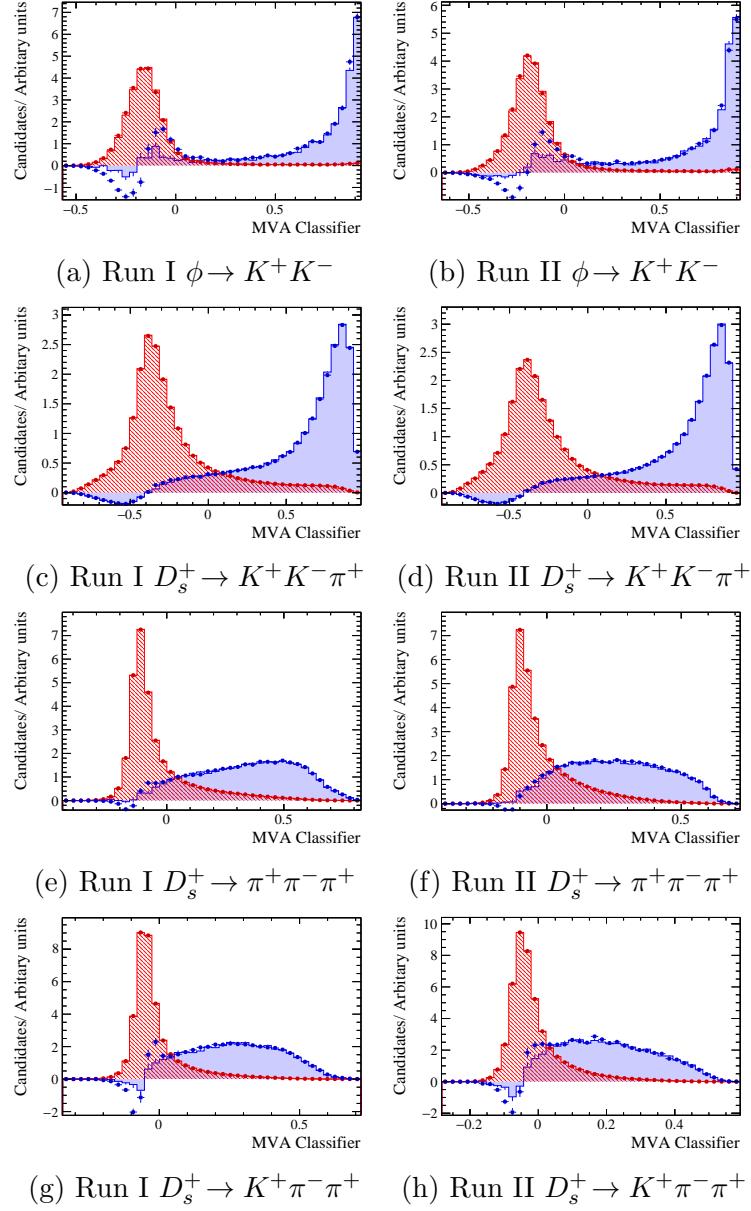


Figure 4.15: MVA classifier response for the eight MVAs trained for both signal (blue) and background (red). The training samples are represented by markers and the validation samples by filled histograms.

the  $\phi$  or  $D_s^+$  mass. As this only affects the range of MVA classifier values that are dominantly ‘background’-like, it likely this discrepancy is due specifically to presence of weighted background events in the signal sample passed to TMVA. This effect is further studied to determine if it results in any systematic uncertainty in the MVA efficiencies as discussed in Sections 5.7 and 6.7. In all cases the discrepancies are far

below the range of values that selection requirements are placed.

### MVA efficiency

The efficiencies of the MVAs on  $D_s^+$  and  $\phi$  meson in the signal decays can be obtained from the validation samples of  $B_s^0 \rightarrow J/\psi\phi$  and  $\bar{B}_s^0 \rightarrow D_s^+\pi^-$  decays. Additionally, a sample of  $B^+ \rightarrow \bar{D}^0\pi^+$  decays is used to calculate the efficiency of  $\bar{D}^0 \rightarrow K^+K^-$  decays in the normalisation channel. The efficiency calculation takes into account the kinematic differences between the training and signal samples, as well as any possible correlations between the  $D_s^+$  and  $\phi$  kinematics, by using input from simulation samples. The signal MVA response is extracted from the validation samples in four bins of both  $p_T$  and  $\chi_{FD}^2$ . Simulation samples for  $B^+ \rightarrow D_s^+\phi$  or  $B^+ \rightarrow D_s^+\bar{D}^0$  decays are iterated through in turn, selecting the corresponding MVA response from the appropriate kinematic bins. The per-candidate efficiency is determined by integrating the validation sample  $D_s^+$  and  $\phi$  MVA responses above the required cut values

$$\varepsilon_{\text{event}} = \int_{\text{cut}_{\phi}}^1 f(x)dx \times \int_{\text{cut}_{D_s^+}}^1 g(y)dy \quad (4.10)$$

where  $f(x)$  and  $g(y)$  represent the weighted MVA classifier output for the  $\phi$  and  $D_s^+$  MVAs respectively, and  $\text{cut}_{\phi}$  and  $\text{cut}_{D_s^+}$  represent the chosen MVA cut value. The total efficiency for each mode is given by the sum of the per-candidate efficiencies within the relevant simulation sample.

The  $\phi \rightarrow K^+K^-$  MVA is also used to select the  $K^+K^-$  pair in  $B^+ \rightarrow D_s^+K^+K^-$  decays. Due to the similar topologies of  $B^+ \rightarrow D_s^+K^+K^-$  and  $B^+ \rightarrow D_s^+\phi$  decays, the  $K^+K^-$  pair share lots of similarities to the  $\phi \rightarrow K^+K^-$  counterparts. One obvious difference is the invariant mass of the pair, as the  $B^+ \rightarrow D_s^+K^+K^-$  decays includes a much larger phase-space. In training the MVA this could lead to the selection being non-optimal for  $B^+ \rightarrow D_s^+K^+K^-$  decays at high  $m(K^+K^-)$ . However, when determining the efficiency of the selection using these same  $B_s^0 \rightarrow J/\psi\phi$  decays these differences could lead to bias. Instead, calibration samples are used to correct for the imperfect modelling of the particle identification in the  $B^+ \rightarrow D_s^+K^+K^-$

simulation samples. These corrected simulations are then used to obtain the variations in the MVA efficiencies as a function of the phase-space position, in particular of the  $m(K^+K^-)$  invariant mass as discussed in Section 5.6.

### MVA optimisation

The selection criteria for each of the BDTG classifiers are determined by optimising the Punzi figure of merit [115],

$$\frac{\epsilon_s}{\left(\frac{a}{2} + \sqrt{N_{\text{BKG}}}\right)} \quad (4.11)$$

with  $a = 5$ , where  $\epsilon_s$  is the signal efficiency and  $N_{\text{BKG}}$  is the number of background candidates determined from fits to data, calculated in the signal region. The signal efficiency is calculated for the relevant mode and MVA requirement values using the data-driven technique previously described. The optimisation is performed in two dimensions, one for each of the  $\phi$  and  $D_s^+$  MVAs. This is performed separately for the three different  $D_s^+$  decay modes. The results of the optimisation for each  $D_s^+$  decay mode and running period are shown in Fig. 4.16.

The optimal values are selected separately for each of the modes and tabulated in Table 4.18.

Period	$D_s^+$ decay mode	$D_s^+$ requirement	$\phi/\bar{D}^0$ requirement
Run I	$D_s^+ \rightarrow K^+K^-\pi^+$	0.5	0.8
	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	0.3	0.8
	$D_s^+ \rightarrow K^+\pi^-\pi^+$	0.2	0.8
Run II	$D_s^+ \rightarrow K^+K^-\pi^+$	0.5	0.8
	$D_s^+ \rightarrow \pi^+\pi^-\pi^+$	0.2	0.8
	$D_s^+ \rightarrow K^+\pi^-\pi^+$	0.2	0.8

Table 4.18: Optimised MVA cuts for the different  $D_s^+$  decay modes and datasets.

#### 4.3.8 Impact parameter requirements

The data-driven MVAs target the selection of  $\phi$  and  $D_s^+$  candidates, removing a significant background contribution. However it is possible to further purify the candidates

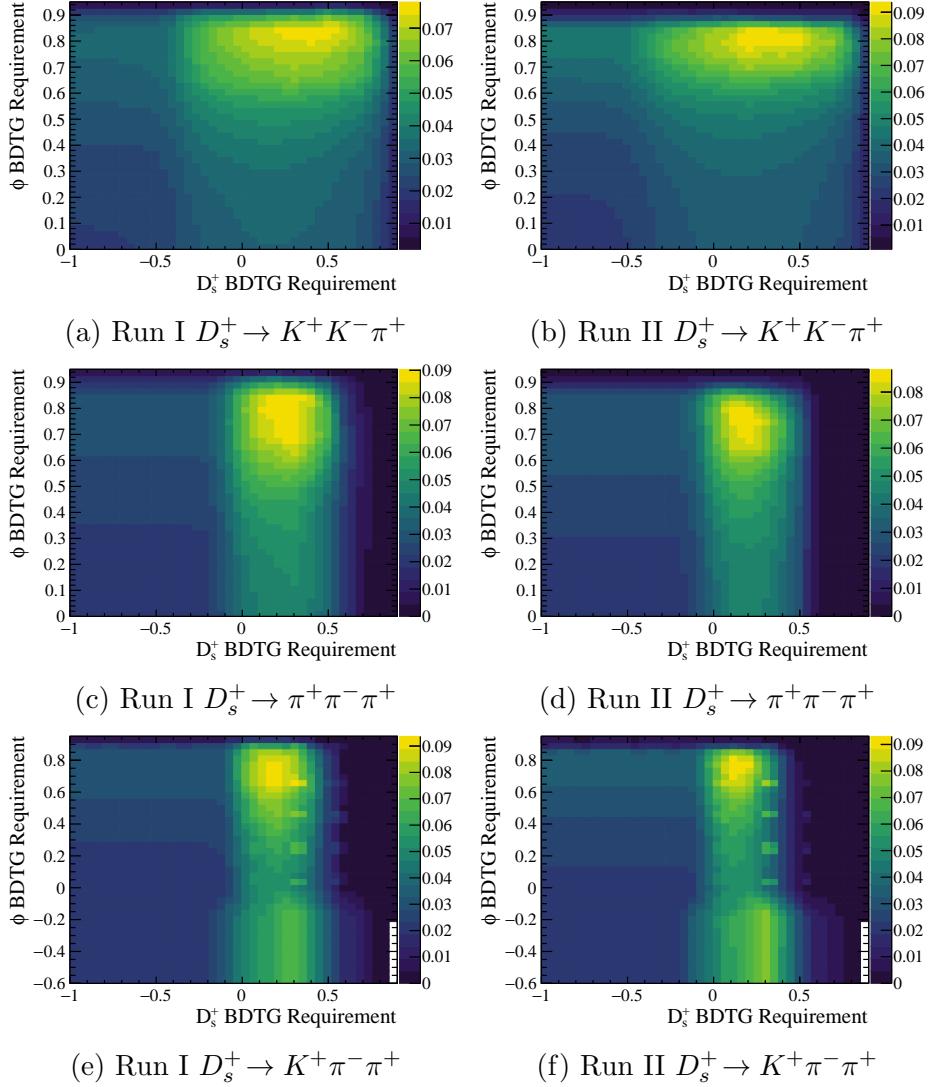


Figure 4.16: MVA optimisation

by making requirements on the  $B^+$  candidates formed from the combination of the two. This helps to remove possible backgrounds coming from the incorrect combination of two unrelated mesons. Requirements are placed on the impact parameter significance  $\chi_{\text{IP}}^2$  to help ensure that the different components of the decay chain have a topology consistent with those desired.

The  $B^+$  meson candidate is required to be consistent with originating at the PV by requiring  $\chi_{\text{IP}}^2 < 10$ . This helps to remove background as unrelated combinations of  $D_s^+$  and  $\phi$  mesons will not necessarily point back towards the PV. Additionally, the opposite requirement is made of the  $D_s^+$  meson,  $\chi_{\text{IP}}^2 > 10$ . It is possible for  $D_s^+$

mesons to be produced promptly at the PV rather than in a  $B^+$  meson decay and this requirement helps to remove those candidates.

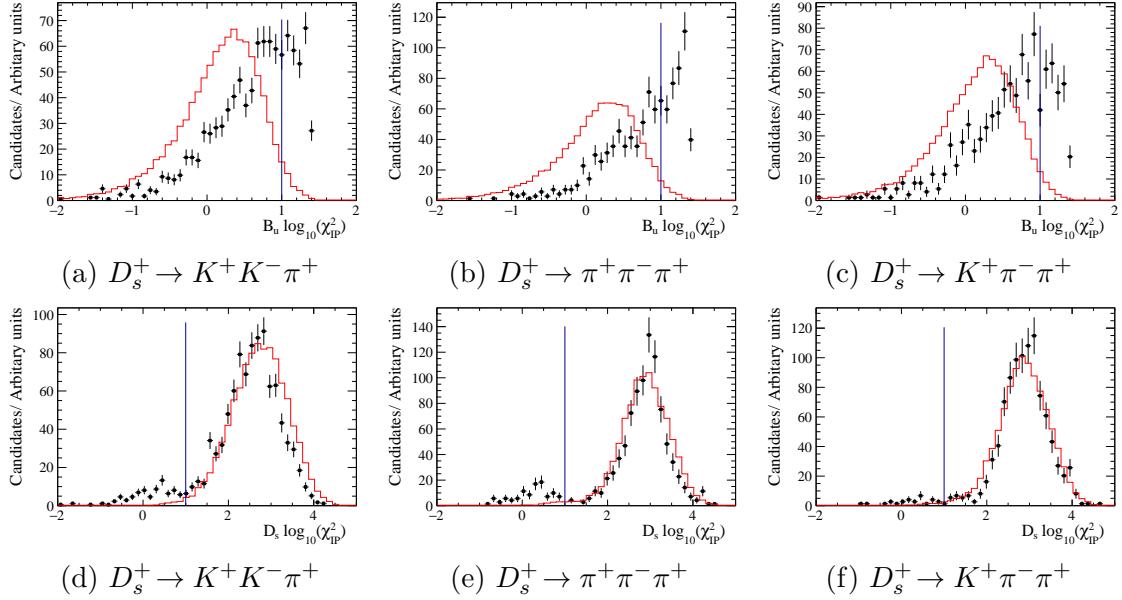


Figure 4.17: The  $B^+$  and  $D_s^+$  meson  $\chi_{\text{IP}}^2$  distributions for  $B^+ \rightarrow D_s^+ \phi$  candidates in data (black) and simulation (red). Candidates within the range  $4900 < m(D_s^+ \phi) < 5900 \text{ MeV}/c^2$  are included in these figures.

The distributions for the  $B^+$  and  $D_s^+$  meson impact parameter significance is shown in Fig. 4.17 for  $B^+ \rightarrow D_s^+ \phi$  candidates. The requirement values are represented by vertical lines. Small peaks can be observed at low  $D_s^+$  meson impact parameter significance, corresponding to  $D_s^+$  mesons that originated at the PV. As a cross-check, the same distributions are produced for the normalisation channel,  $B^+ \rightarrow D_s^+ \bar{D}^0$ . However, instead of including all candidates in a wide  $m(D_s^+ \phi)$  mass range, only those within  $20 \text{ MeV}/c^2$  of the  $B^+$  mass are included. As the normalisation sample has a very high purity this effectively isolates these decays in data. This comparison of data and simulation distributions is shown in Fig. 4.18.

Small shifts are observed in the  $\chi_{\text{IP}}^2$  distributions. These differences are studied as a source of systematic uncertainty as discussed in Sections 5.7 and 6.7.

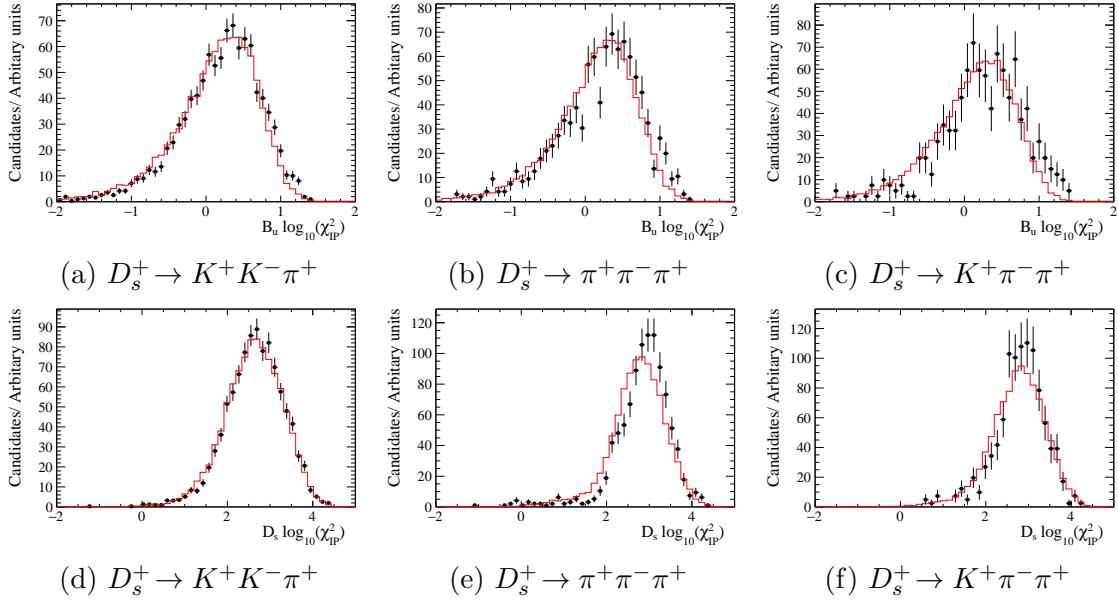


Figure 4.18: The  $B^+$  and  $D_s^+$  meson  $\chi_{\text{IP}}^2$  distributions for  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates in data (black) and simulation (red). Candidates within the range  $|m(D_s^+\phi) - m(B^+)| < 20 \text{ MeV}/c^2$  are included in these figures, effectively isolating the normalisation events in data.

## 4.4 Multiple candidates

Signal and normalisation candidates are constructed by combining five of the large number of tracks in each event in different combinations. It is possible that after the many steps of selection have been performed, that there are still multiple candidates reconstructed in a single event. It is, of course, possible that there are more than one  $B$  meson decays in a single event. Indeed production via  $b\bar{b}$  pairs would lead to two  $B$  hadrons in the same event. However both the singal and normalisation channels searched for here have relatively small branching fractions, therefore the probability for both  $B^+$  mesons to decay the same final state is low. The rate of multiple candidates is studied for each of the samples after passing through the full selection. The run and event number of each candidate is compared, and the number of events that contain more than one candidate is divided by the total number of candidates to produce a rate. These are tabulated in Table 4.19. The rates of multiple candidates are sufficiently low that no attempt is made to remove them.

$B^+$ decay mode	$D_s^+$ decay mode	Multiple candidate rate
$B^+ \rightarrow D_s^+ \phi$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.33%
	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	0.46%
	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	0.00%
$B^+ \rightarrow D_s^+ \bar{D}^0$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.27%
	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	0.11%
	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	0.17%
$B^+ \rightarrow D_s^+ K^+ K^-$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.25%
$B^+ \rightarrow D_s^+ \bar{D}^0$	$D_s^+ \rightarrow K^+ K^- \pi^+$	0.21%

Table 4.19: Multiple candidate rates for signal and normalisation decays.

## 4.5 Refitting the decay chain

The yields of  $B^+$  meson candidates for the signal and normalisation channels are determined from fits to the  $B^+$  meson invariant mass, the exact details of which will be in Chapters 5 and 6. The *raw*  $B^+$  meson invariant mass is calculated in the vertex fits used to constructed the  $B^+$  candidates from combinations of  $D_s^+$  and  $\phi$  or  $\bar{D}^0$  candidates. The 4-momentum vectors for the decay products trajectories are extrapolated to the vertex location and the invariant mass calculated from the sum of the decay products 4-vectors. The width of the  $B^+$  meson distribution is dominated by the resolution of the LHCb detector, as the natural width of the meson itself is  $0.4 \times 10^{-3}$  eV.

It is possible to reduce the effect of the resolution by adding more constraints in the fit to the decay chain. The mass of the  $D_s^+$  meson (and  $\bar{D}^0$  meson for the normalisation channel) are constrained to their known values. Additionally, the momentum vector of the  $B^+$  meson is constrained to be parallel to the line between the primary interaction and the  $B^+$  meson decay vertex. This is implemented with the Decay Tree Fitter algorithm [116]. With these constraints applied, the whole decay chain is refitted, and the  $B^+$  meson mass recalculated. For genuine signal and normalisation decays the resolution is improved; the smearing effects that might lead the track momentum or position to be incorrectly measured can be circumvented by including the

extra information about intermediate meson masses, or  $B^+$  flight direction. Background candidates that have different topologies to the constrained decay chain will be randomly shifted up or down in invariant mass, leading to no overall difference in the combinatorial distribution.

It is possible to calculate all kinematic distributions for the components of the decay chain with these constraints applied, for example the momentum of the decay products. However, these variables are not used when training the MVA methods even though they may result in a greater discrimination between signal and background. The modes used to train the MVA methods have slightly different topologies, therefore the improvements in the resolution achieved in these modes may not be the same as that seen for the signal and normalisation channel.

The search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays involves determining the phase-space distribution of candidates in both simulations and data using the two-body masses  $m^2(D_s^+ K^-)$  and  $m^2(K^+ K^-)$ . These are determined using the decay product 4-vectors calculated with the  $B^+$  meson mass constrained in addition to the  $D_s^+$  mass and  $B^+$  flight direction. This ensures the candidates are found within the kinematically allowed region of the two dimensional  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  space.

In the search for  $B^+ \rightarrow D_s^+ \phi$  decays an angle,  $\theta_K$ , is used to split candidates into two categories. This angle is calculated using the decay products 4-vectors calculated using the  $D_s^+$  mass and  $B^+$  flight direction constraints. This improves the separation power as opposed to using the *raw* 4-vectors.

## 4.6 Invariant mass distributions

The  $D_s^+$  and  $\phi$  meson invariant mass distributions for signal  $B^+ \rightarrow D_s^+ \phi$  candidates passing the full selection requirements are shown in Fig. 4.19. The same distributions for the normalisation channel  $B^+ \rightarrow D_s^+ \bar{D}^0$  are shown in Fig. 4.20. The purity of the  $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$  and  $D_s^+ \rightarrow K^+ \pi^- \pi^+$  decays are significantly lower than the  $D_s^+ \rightarrow K^+ K^- \pi^+$  mode as a result of the smaller branching fractions for these modes.

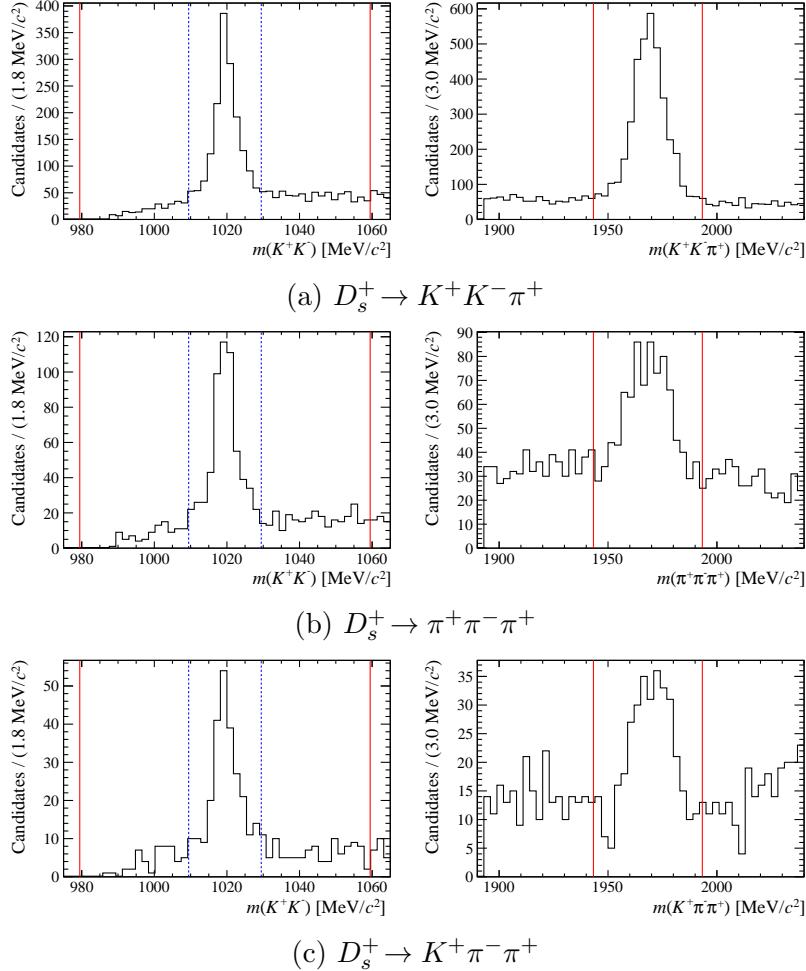


Figure 4.19: The  $D_s^+$  and  $\phi$  invariant mass distributions for signal  $B^+ \rightarrow D_s^+ \phi$  candidates passing the selection requirements. The inner and outer  $m(K^+ K^-)$  ranges, as defined in Chapter 6, are represented by the blue and red vertical lines respectively.

Requirements are placed in the invariant masses of the  $D_s^+$ ,  $\phi$  and  $\bar{D}^0$  candidates to isolate the signal decays. The requirements are listed in Table 4.20 and represented in Figs. 4.19 and 4.20 by red vertical lines. The  $\phi$  meson requirement is wider than those for the  $D_s^+$  and  $\bar{D}^0$  mesons even though the width is clearly narrower. The  $B^+ \rightarrow D_s^+ \phi$  candidates are further split into two categories as detailed in Chapter 6, therefore the sidebands either side of the  $\phi$  meson are retained.

The distribution of  $D_s^+$  and  $K^+ K^-$  invariant masses for  $B^+ \rightarrow D_s^+ K^+ K^-$  and the corresponding  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates are shown in Fig. 4.21.

The  $B^+$  invariant mass distribution for the normalisation mode is shown in Fig. 4.22.

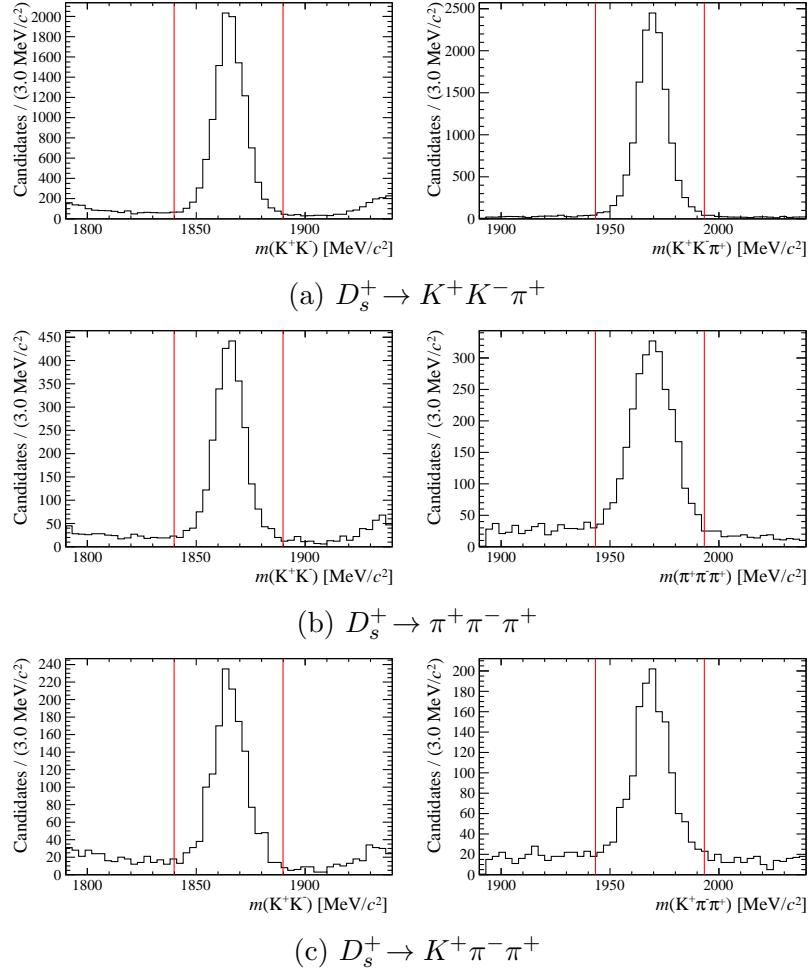


Figure 4.20: The  $D_s^+$  and  $\bar{D}^0$  invariant mass distributions for normalisation  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates passing the selection requirements.

Particle	Requirement
$D_s^+$	$ m(h^+ h^- h^+) - m(D_s^+)  < 25 \text{ MeV}/c^2$
$\bar{D}^0$	$ m(K^+ K^-) - m(\bar{D}^0)  < 25 \text{ MeV}/c^2$
$\phi$	$ m(K^+ K^-) - m(\phi)  < 40 \text{ MeV}/c^2$

Table 4.20: Invariant mass requirements applied to all candidates.

The distribution of candidates is shown with the different aspects of the selection applied. It can be seen that the level of combinatorial background that is present across the whole mass range dramatically decreases as the different selection criteria are applied. Additionally, it is clear the normalisation channel efficiency decreases too, but at a much slower rate. The final distribution has a negligible amount of combinatorial

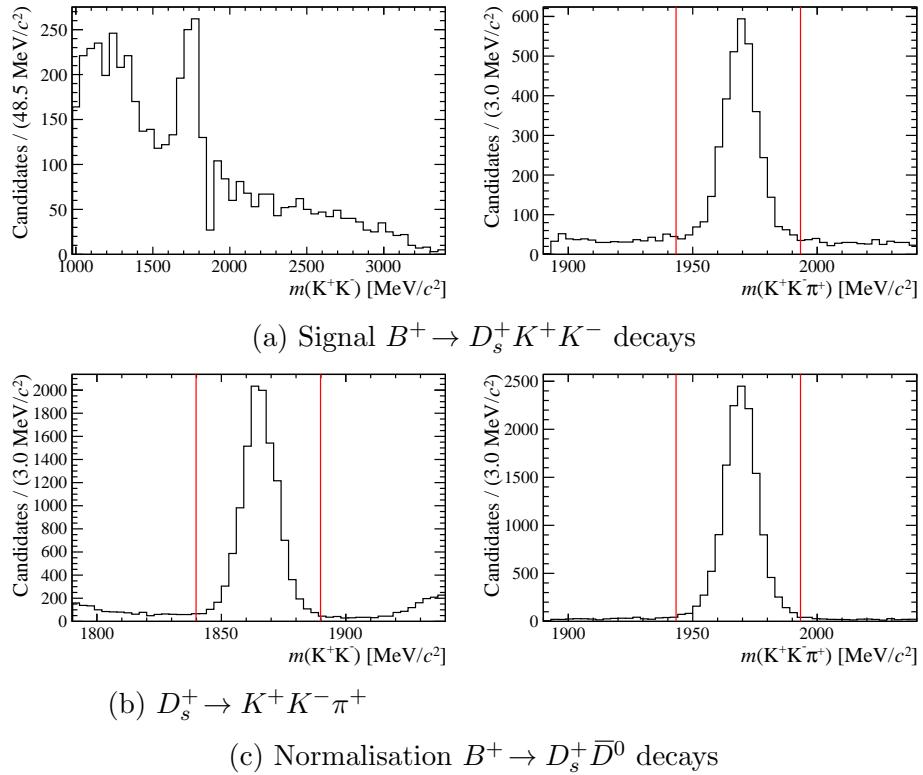


Figure 4.21: The  $D_s^+$  and  $K^+ K^-$  invariant mass distributions for  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates passing the selection requirements.

background with an isolated normalisation peak.

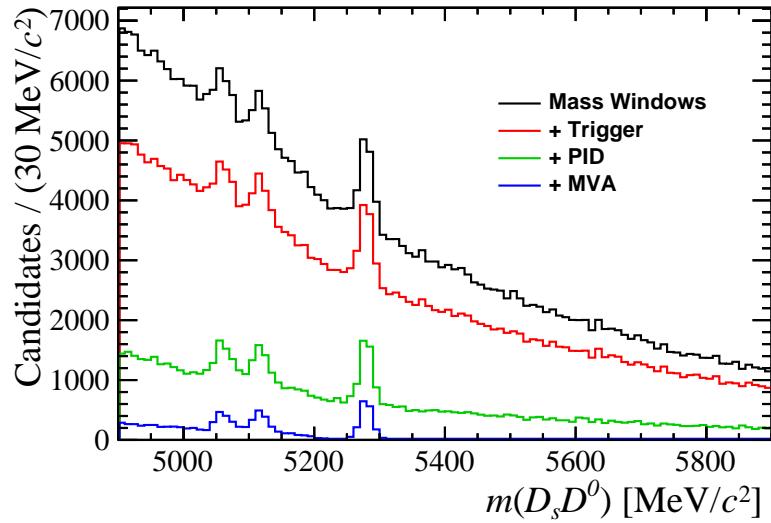


Figure 4.22: Various selection steps applied to the normalisation channel as defined in the legend.

# Chapter 5

## Mass fit to $B^+ \rightarrow D_s^+ K^+ K^-$ candidates

### Contents

---

<b>5.1</b>	<b>Fit strategy</b>	<b>113</b>
<b>5.2</b>	<b>Fit components</b>	<b>114</b>
5.2.1	Signal and normalisation decays	114
5.2.2	Partially reconstructed backgrounds	115
5.2.3	Combinatorial background	121
<b>5.3</b>	<b>Free and constrained parameters</b>	<b>122</b>
5.3.1	Parameter of interest	122
5.3.2	Shape parameters	122
5.3.3	Yields	122
<b>5.4</b>	<b>Fit validation</b>	<b>123</b>
<b>5.5</b>	<b>Normalisation and signal fits</b>	<b>125</b>
<b>5.6</b>	<b>Efficiency corrections</b>	<b>128</b>
5.6.1	Efficiencies from simulations	128
5.6.2	Efficiencies requiring calibration samples	131
5.6.3	Total efficiency	134
<b>5.7</b>	<b>Systematic uncertainties</b>	<b>135</b>
5.7.1	Relative efficiencies	135
5.7.2	Signal and normalisation PDFs	137
5.7.3	Background PDFs	138
5.7.4	Charmless contribution	138
5.7.5	Total systematic uncertainty	139
<b>5.8</b>	<b>Results</b>	<b>139</b>
<b>5.9</b>	<b>Outlook</b>	<b>142</b>

In this chapter the methodology used to search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is described. The branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)$  is determined by measuring the ratio of  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  yields. This ratio is corrected to account for the corresponding selection efficiencies of the two modes. Finally, the corrected yield ratio is multiplied by the externally measured branching fractions for the normalisation channel  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $\bar{D}^0 \rightarrow K^+ K^-$  decays to determine  $\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)$ .

The parametrisations used to model the signal and backgrounds components and extract the candidate yields are described in Section 5.2, the efficiency corrections are described in Section 5.6 and the resulting calculation of the branching fraction is in Section 5.8.

## 5.1 Fit strategy

The search for  $B^+ \rightarrow D_s^+ K^+ K^-$  involves two independent unbinned extended maximum likelihood fits for the signal and normalisation channels implemented with the RooFIT package [117] within ROOT. The extended likelihoods for the two fits are constructed in a similar manner to those already detailed in Sec. 4.3.7. However, as a larger number of components are included there are correspondingly more contributions included in the likelihood

$$-\log \mathcal{L}(n_0 \dots n_j, \vec{p}) = -\sum_i^N \log \left( \sum_j n_j f_j(m = m_i, \vec{p}) \right) + \sum_j n_j, \quad (5.1)$$

where the index  $j$  represents each component of the model with yield  $n_j$  and PDF  $f_j$ . The constant  $\log N!$  term has been ignored. Separate likelihoods are created for the signal and normalisation fits.

The raw  $B^+ \rightarrow D_s^+ K^+ K^-$  yield is corrected on a per-candidate basis to account for the phase-space dependence of the signal efficiencies in this three-body decay. This is implemented using the *sPlot* technique [112] to determine a signal weight  $W_{s,i}$  for

each event  $i$  in the fitted data set. The weights for each component are constructed such that they sum to the fitted value of that components yield

$$n_{\text{sig}} = \sum_i^N W_{\text{sig},i}, \quad (5.2)$$

where  $n_{\text{sig}}$  is the fitted signal yield and  $N$  is the total number of entries in the data set. The efficiencies for the  $B^+ \rightarrow D_s^+ K^+ K^-$  signal decays are determined as a function of the kinematic properties of the decay. The weight for each entry  $i$  in the data set is corrected with the appropriate efficiency for its given kinematics

$$n_{\text{sig,corr}} = \sum_i^N \frac{W_{\text{sig},i}}{\epsilon_i}. \quad (5.3)$$

The propagation of the uncertainty in this corrected yields is described in Sec. 5.8.

The normalisation channel decay proceeds via a pseudo two-body process, therefore no kinematic dependent efficiency correction is required.

## 5.2 Fit components

In order to extract the yields of  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays the invariant mass distributions for the processes contributing within the invariant mass range are parametrised with probability density functions (PDFs). Both the signal and normalisation channels are considered within the same  $B^+$  meson invariant mass range 5100–5900 MeV/ $c^2$ . This is sufficiently wide to allow the contributions from different background components to be distinguished and accurately extrapolated into the signal region. The entire  $m(K^+ K^-)$  phase-space is included in the search for the signal decays, including the range in the vicinity of the  $\phi$  meson used later in Chapter 6.

### 5.2.1 Signal and normalisation decays

The invariant mass distributions of  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are parametrised as the sum of two Crystal Ball (CB) functions. The CB function consists

of a Gaussian function with a power-law tail and is typically used to parametrise losses due to radiative processes. This is defined as

$$\text{CB}(m|\mu, \sigma, n, \alpha) = \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2}, & \text{if } \left(\frac{m-\mu}{\sigma}\right) < -|\alpha| \\ \frac{\left(\frac{n}{|\alpha|}\right)^n \times e^{-\frac{1}{2}|\alpha|^2}}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma}\right)^n}, & \text{otherwise} \end{cases} \quad (5.4)$$

where  $\mu$ ,  $\sigma$ ,  $n$  and  $\alpha$  are adjustable parameters and  $m$  is the  $B$  meson invariant mass observable. The sum of two CB functions is constructed with a variable fraction  $f_\sigma$  assigned to the CB function with the narrower width,

$$\text{DCB}(m|\mu, \sigma_1, \sigma_2, n, \alpha) = f_\sigma \times \text{CB}(m|\mu, \sigma_1, n, \alpha) + (1 - f_\sigma) \times \text{CB}(m|\mu, \sigma_2, n, \alpha), \quad (5.5)$$

where the same tail parameters,  $n$  and  $\alpha$  are used for both functions, but the widths,  $\sigma_1$  and  $\sigma_2$ , are allowed to be different (with  $\sigma_1 < \sigma_2$ ). As both CB shapes have the same parameter  $\alpha$ , the tails are constrained to be on the same side. Values for the adjustable parameters are determined from fits to simulated decays passing the selection requirements applied to the data. However, a number of parameters are not completely constrained from the simulations. The mean position  $\mu$  is allowed vary freely in the fit to data, as is the narrowest CB width of the normalisation and signal decays. The tail parameters  $n$  and  $\alpha$  are highly correlated, therefore the value of  $n$  is fixed to unity in both the fits to simulations and data. The values determined from simulations for  $B^+ \rightarrow D_s^+ \bar{D}^0$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are tabulated in Table 5.1 and the results of the corresponding fits are shown in Fig. 5.1.

### 5.2.2 Partially reconstructed backgrounds

Partially reconstructed decays are those in which the five final state particles combined in the signal mode are only a subset of a background mode's final state. Decays of  $b$ -hadrons can contribute at lower invariant masses below the signal peak when one or more the decay products have not been reconstructed. For decays to contribute within the fitted  $B^+$  invariant mass window, the particle or particles that have not been reconstructed must be fairly low-momentum (soft) such that the invariant mass

Parameter	Value
$B^+ \rightarrow D_s^+ K^+ K^-$	
$\sigma_1/\sigma_2$	$0.53 \pm 0.02$
$f_\sigma$	$0.87 \pm 0.02$
$\alpha$	$2.60 \pm 0.03$
$n$	$1 \pm 0$
$B^+ \rightarrow D_s^+ \bar{D}^0$	
$\sigma_1/\sigma_2$	$0.60 \pm 0.03$
$f_\sigma$	$0.66 \pm 0.12$
$\alpha$	$2.67 \pm 0.12$
$n$	$1 \pm 0$

Table 5.1: Fixed values obtained in fits to simulations used in the model for the signal and normalisation PDFs.

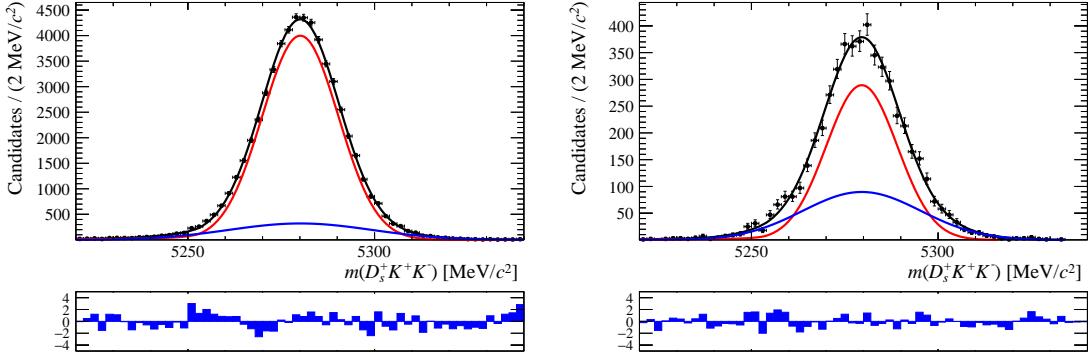


Figure 5.1: Invariant mass fits to signal (left) and normalisation (right) channel simulation samples.

of the remaining particles is large. Accurate parametrisation of these contributions is vital as many of the distributions extend close to or within the range of the signal distribution. Incorrectly attributing these decays to the signal component could lead to the incorrect branching fraction being measured.

### Backgrounds to the normalisation channel

The low invariant mass region of the  $D_s^+ \bar{D}^0$  spectrum is populated by decays of  $B^+$  mesons to combinations of  $D$  and excited  $D$  mesons. These  $\bar{D}^{*0}$  and  $D_s^{*+}$  mesons decay strongly to a ground state  $\bar{D}^0$  or  $D_s^+$  meson and a soft pion or photon. The branching fractions for these decays are listed in Table 5.2.

The excited charm mesons  $\bar{D}^{*0}$  and  $D_s^{*+}$  are vector ( $J^P = 1^-$ ) mesons. The

Decay	Branching fraction
$\bar{D}^{*0} \rightarrow \bar{D}^0 \gamma$	$(64.7 \pm 0.9)\%$
$\bar{D}^{*0} \rightarrow \bar{D}^0 \pi^0$	$(35.3 \pm 0.9)\%$
$D_s^{*+} \rightarrow D_s^+ \gamma$	$(93.5 \pm 0.7)\%$
$D_s^{*+} \rightarrow D_s^+ \pi^0$	$(5.8 \pm 0.7)\%$

Table 5.2: Branching fractions for excited charm mesons [28].

partially reconstructed invariant mass of the  $D_s^+$  and  $\bar{D}^0$  mesons,  $m(D_s^+ \bar{D}^0)$ , varies depending on the spin of the missed particle. Analytical PDFs are used to account for the spin and mass of the missing particle. These PDFs have been used to describe other partially reconstructed  $B \rightarrow DX$  decays investigated by the LHCb collaboration [118].

**$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ [\pi^0]) \bar{D}^0$  and  $B^+ \rightarrow D_s^+ (\bar{D}^{*0} \rightarrow \bar{D}^0 [\pi^0])$ :** the  $\pi^0$  meson is a pseudo-scalar ( $J^P = 0^-$ ) particle with a mass  $m(\pi^0) = 134.9766 \pm 0.0006 \text{ MeV}/c^2$ . The partially reconstructed invariant mass distribution  $m(D_s^+ \bar{D}^0)$  can be described by a parabola convolved with a Gaussian resolution function. The parabola does not extend beyond kinematic endpoints defined by the parameters  $a$  and  $b$  and has a minimum in the centre

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b \left( \mu - \frac{a+b}{2} \right)^2 \left( \frac{1-\xi}{b-a} \mu + \frac{b\xi - a}{b-a} \right) e^{-\frac{-(\mu-(m-\delta))^2}{2\sigma^2}} d\mu. \quad (5.6)$$

Here,  $\sigma$  is the width of the resolution Gaussian and  $\delta$  allows the function to be offset in invariant mass. The parameter  $\xi$  introduces the freedom for the two sides of the parabola to have difference heights, achieved by multiplying the parabola by a line whose slope is depends on  $\xi$ . The resulting function is a double peaked structure shown in Fig. 5.2. The parameters  $a$  and  $b$  are calculated from the kinematics of the  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ \pi^0) \bar{D}^0$  and  $B^+ \rightarrow D_s^+ (\bar{D}^{*0} \rightarrow \bar{D}^0 \pi^0)$  decays respectively, and are listed in Table 5.3.

**$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ [\gamma]) \bar{D}^0$  and  $B^+ \rightarrow D_s^+ (\bar{D}^{*0} \rightarrow \bar{D}^0 [\gamma])$ :** the  $\gamma$  boson is a massless vector ( $J^P = 1^-$ ) particle. The partially reconstructed  $m(D_s^+ \bar{D}^0)$  invariant

Mode	$a$ (MeV/c <sup>2</sup> )	$b$ (MeV/c <sup>2</sup> )
$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\pi^0])\bar{D}^0$	5051.4	5132.9
$B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\pi^0])$	5051.5	5128.6

Table 5.3: Kinematic endpoints for the partially reconstructed  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ \pi^0)\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0 \pi^0)$  decays.

mass is also described by a parabola convolved with a Gaussian resolution function. The parabola does not extend beyond the endpoints  $a$  and  $b$ , and has a maximum in the centre

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b -(\mu - a)(\mu - b) \left( \frac{1 - \xi}{b - a} \mu + \frac{b\xi - a}{b - a} \right) e^{-\frac{-(\mu - (m - \delta))^2}{2\sigma^2}} d\mu. \quad (5.7)$$

Again  $\sigma$ ,  $\delta$  and  $\xi$  control the width, offset and relative heights of two sides of the parabola. The resulting function is a broad single peak shown in Fig. 5.2.

Mode	$a$ (MeV/c <sup>2</sup> )	$b$ (MeV/c <sup>2</sup> )
$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\gamma])\bar{D}^0$	4976.7	5213.1
$B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\gamma])$	4970.1	5216.1

Table 5.4: Kinematic endpoints for the partially reconstructed  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+ \gamma)\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0 \gamma)$  decays.

## Backgrounds to the signal channel

The signal channel receives contributions at low invariant mass from a number of different decays. All modes considered involve a  $B_s^0$  or  $B^0$  meson decay in which one or more soft decay products have not been reconstructed.

$\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0}$ : this decay can form a background to the  $B^+ \rightarrow D_s^+ K^+ K^-$  signal when a soft pion from the  $K^{*0} \rightarrow K^+ \pi^-$  decay is not reconstructed. The  $K^- K^{*0}$  is modelled as originating from the  $a_1(1260)$  resonance. This resonance has a width of 250–600 MeV [28], allowing it to decay to  $K^- K^{*0}$  even though its pole mass is below the  $K^- K^{*0}$  threshold. The PDF for this background is determined from a sample of fully reconstructed simulated events that have been processed using the same procedure as the signal decays. The PDF is created using a

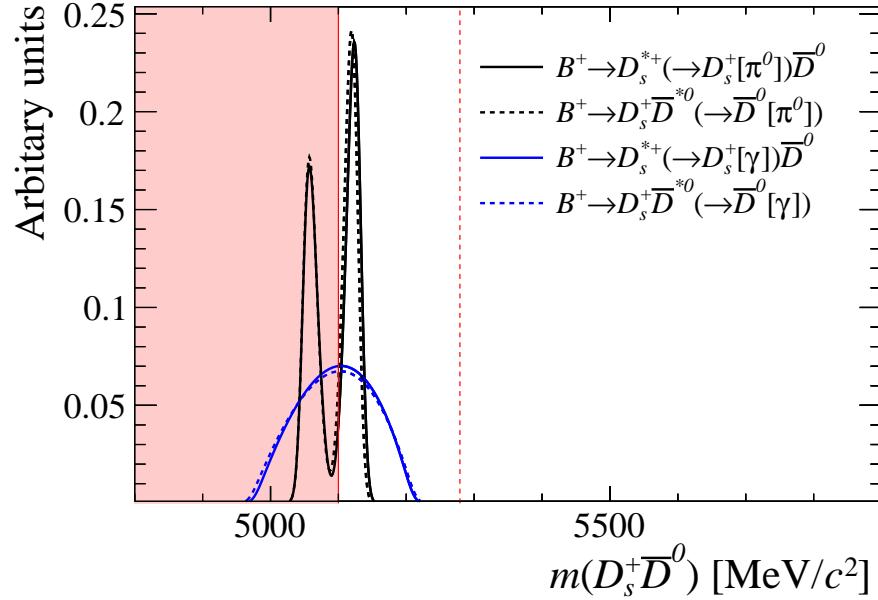


Figure 5.2: Partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  PDFs as described in Sec. 5.2.2. The range below  $5100 \text{ MeV}/c^2$  (highlighted in red) is not included in the fit range, but included to show the full PDF distributions. The  $B^+$  meson mass is represented by a vertical line.

kernel estimation technique [119] implemented in the `RooKeysPDF` class within `ROOFIT`. The resulting PDF is shown in Fig. 5.3.

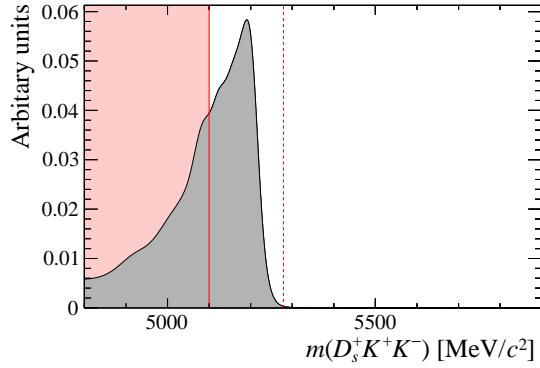


Figure 5.3: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$  mass shape. The  $B^+$  meson mass is represented by a vertical dashed line. The range below  $5100 \text{ MeV}/c^2$  (highlighted in red) is not included in the fit range.

$\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$ : similarly this decay can form a background at low invariant mass when a soft neutral particle is not reconstructed in the  $D_s^{*+} \rightarrow D_s^+ X$  decay

in addition to the pion from the  $K^{*0}$ . This PDF is also determined using the `RooKeysPDF` class and shown in Fig. 5.4.

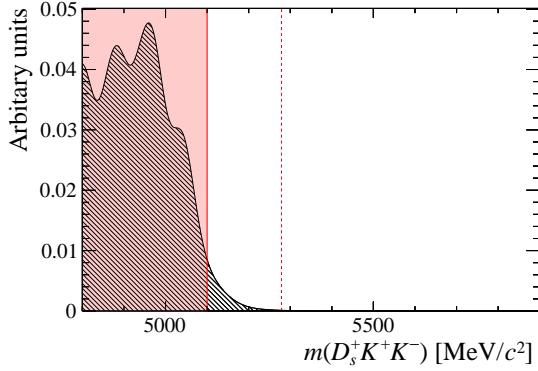


Figure 5.4: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$  mass shape. The  $B^+$  meson mass is represented by a vertical dashed line. The range below  $5100 \text{ MeV}/c^2$  (highlighted in red) is not included in the fit range.

$\bar{B}_s^0 \rightarrow D_s^+ D_s^-$ : this decay can form a background to the signal when a pion is missed from either of the  $D_s^+$  decays. This requires both  $D_s^+$  mesons to decay to the  $D_s^+ \rightarrow K^+ K^- \pi^+$  final state. This background becomes significant at close to the  $B^+$  meson mass as shown in Fig. 5.5. This PDF is similarly determined by creating a kernel estimation of fully-reconstructed simulated events passing the signal selection.

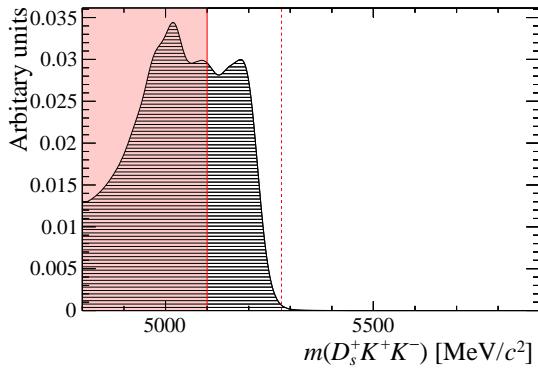


Figure 5.5: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  mass shape. The  $B^+$  meson mass is represented by a vertical dashed line. The range below  $5100 \text{ MeV}/c^2$  (highlighted in red) is not included in the fit range.

$\bar{B}^0 \rightarrow D_s^+ D^-$ : this decay can form a background when both the  $D_s^+$  and  $D^-$  mesons decay to the  $K^\pm K^\mp \pi^\pm$  final state. Due to the similar topology to the  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  decay, the same PDF determined from simulated events is used, however it is shifted down in mass by  $40 \text{ MeV}/c^2$  to account for the difference in kinematics.

$\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$ : similarly this decay can cause a background when both a soft neutral particle is missed from the  $D_s^{*+} \rightarrow D_s^+ X$  decay as well as a pion from either of the  $D_s^+$  mesons. As such this only has a small contribution within the fit range as shown in Fig. 5.6.

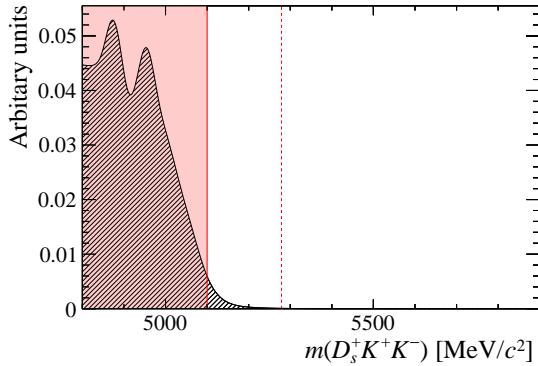


Figure 5.6: Partially reconstructed  $\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$  mass shape.

All PDFs determined using kernel estimations from simulated events are convolved with a Gaussian distribution to account for difference between the simulations and data. The mean position of the Gaussian is given by a single free parameter  $\delta$ , allowing these PDF to move slightly high or lower in mass. The width of the Gaussian is increased to account for the difference in resolution between simulation and data.

### 5.2.3 Combinatorial background

The dominant source of background under the signal peak is due to combinations of unrelated tracks. This combinatorial background is modelled using an exponential function

$$f(m|c) = e^{-m \times c}, \quad (5.8)$$

where  $c$  is the single degree of freedom controlling the effective slope of the function and  $m$  is the observable  $B^+$  meson invariant mass. The separate fits to the signal and normalisation modes have the freedom to have different combinatorial slopes, motivated by the difference in background levels for the two decays.

### 5.3 Free and constrained parameters

The fit to signal and normalisation decays have eleven and nine free parameters respectively. These control the yields and shapes of the fit components.

#### 5.3.1 Parameter of interest

**Signal:** the parameter of interest in the signal decay is  $N(B^+ \rightarrow D_s^+ K^+ K^-)$ , the yield attributed to the signal PDF.

**Normalisation:** the parameter of interest in the signal decay is  $N(B^+ \rightarrow D_s^+ \bar{D}^0)$ , the yield attributed to the normalisation PDF.

#### 5.3.2 Shape parameters

**Signal:** The signal fit contains four free parameters that determine the shape of various PDFs. This corresponds to the mass offset  $\delta$ , signal mean value  $\mu$ , signal width  $\sigma_1$  and combinatorial slope  $c$ .

**Normalisation:** The normalisation fit has five free parameters that control the distributions of the signal and background PDFs. This includes the mean  $B^+$  mass  $\mu$ , signal width  $\sigma_1$ , the combinational slope  $c$ , the mass offset  $\delta$  and the relative heights of the double peaked partially reconstructed shapes  $\xi$ .

#### 5.3.3 Yields

**Signal:** the yields of each background component are left free in the signal fit. This corresponds to six free parameters for  $N_{\text{comb}}$ ,  $N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})$ ,  $N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0})$ ,  $N(B_s^0 \rightarrow D_s^+ D_s^-)$ ,  $N(B^0 \rightarrow D_s^+ D^-)$ , and  $N(B_s^0 \rightarrow D_s^{*+} D_s^-)$ .

**Normalisation:** the yields of the combinatorial and partially reconstructed backgrounds are left free in the fit to the normalisation channel. This corresponds to three parameters;  $N_{\text{comb}}$ ,  $N(B^+ \rightarrow D_s^{*+} \bar{D}^0)$  and  $N(B^+ \rightarrow D_s^+ \bar{D}^{*0})$ . The relatives contributions of the  $D_s^{*+} \rightarrow D_s^+ \pi^0$  and  $D_s^{*+} \rightarrow D_s^+ \gamma$  decays are fixed to their ratio of branching fractions (similarly for  $\bar{D}^{*0}$ ). A factor is included to account for the fraction of each PDF that is within the fitted  $B^+$  invariant mass range.

## 5.4 Fit validation

The fitting framework is validated using large quantities of pseudo-experiments randomly generated using the same fit model. The free parameters are treated using the plug-in method [120]; the generated values of all free parameters are *plugged in* using the fitted values from the fit to data.

The fitted values and uncertainties are determined for each pseudo-experiment and a corresponding pull determined, defined as

$$g_{\text{pull}} = \frac{x_{\text{fit}} - x_{\text{gen}}}{\sigma} \quad (5.9)$$

where  $x_{\text{gen}}$  and  $x_{\text{fit}}$  are the generated and fitted values of the variable, and  $\sigma$  is the parameter's uncertainty. For an ideal unbiased fit model, the pull of each parameter of interest would be normally distributed with unit width and mean of zero.

The distributions of the yields, errors and pulls for the signal and normalisation pseudo-experiments are shown in Fig. 5.7. A fit is performed for each pull distributions using a Gaussian to determine the mean and width. For the signal yield the mean and width are within  $3\sigma$  of zero and one respectively. The normalisation yield shows a significant bias in the width. The bias implies the fit model is overestimating the uncertainty  $\sigma$  of the yield. If the normalisation yield uncertainty dominates the uncertainty in the branching fraction this could lead to an overestimation of the uncertainty on the final measured branching fraction.

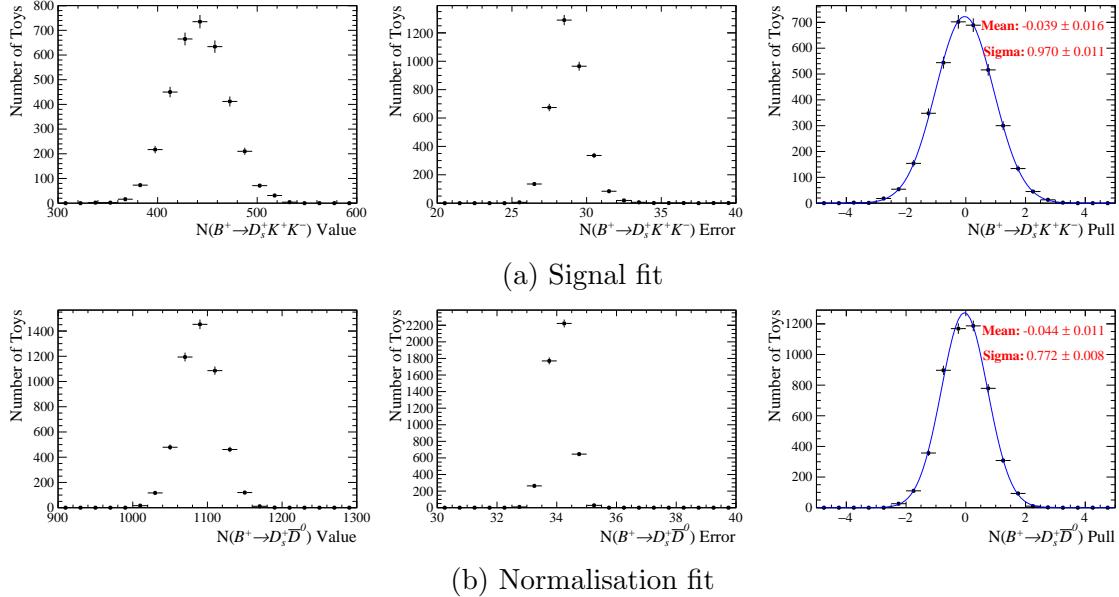


Figure 5.7: The distribution of the signal and normalisation yields, errors and pulls as determined from pseudo-experiments. The result of fit performed to the pull distributions is overlaid in blue, along with the numerical results in red.

To determine how the normalisation uncertainty propagates to the branching fraction another set of pseudo-experiments are produced. This set includes both signal and normalisation decays. To calculate the branching fraction the yield of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is corrected according to the signal efficiency as a function of the kinematics of a given candidate, as detailed previously in Eq. 5.3. The candidates are assumed to have a flat distribution in the two dimensional  $m^2(K^+ K^-)$  vs.  $m^2(D_s^+ K^-)$  space used to parametrise the efficiency. For each pseudo-experiment the branching fraction and uncertainty are produced. As no branching fraction is explicitly used to generate the pseudo-experiment (rather independent signal and normalisation yields), the pulls are redefined to be measured relative to the mean branching fraction

$$g_{\text{pull}} = \frac{x_{\text{fit}} - \bar{x}_{\text{fit}}}{\sigma} \quad (5.10)$$

where  $x_{\text{fit}}$  is the fitted value of the parameter,  $\bar{x}_{\text{fit}}$  is the average of all fitted values and  $\sigma$  is the uncertainty. The distribution of the branching fraction, uncertainty and pull are shown in Fig 5.8.

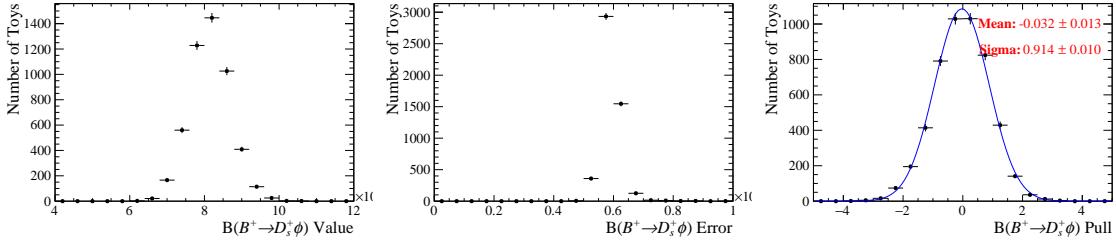


Figure 5.8: The distribution of the branching fraction, error and pull as determined from pseudo-experiments. The result of fit performed to the pull distributions is overlaid in blue, along with the numerical results in red.

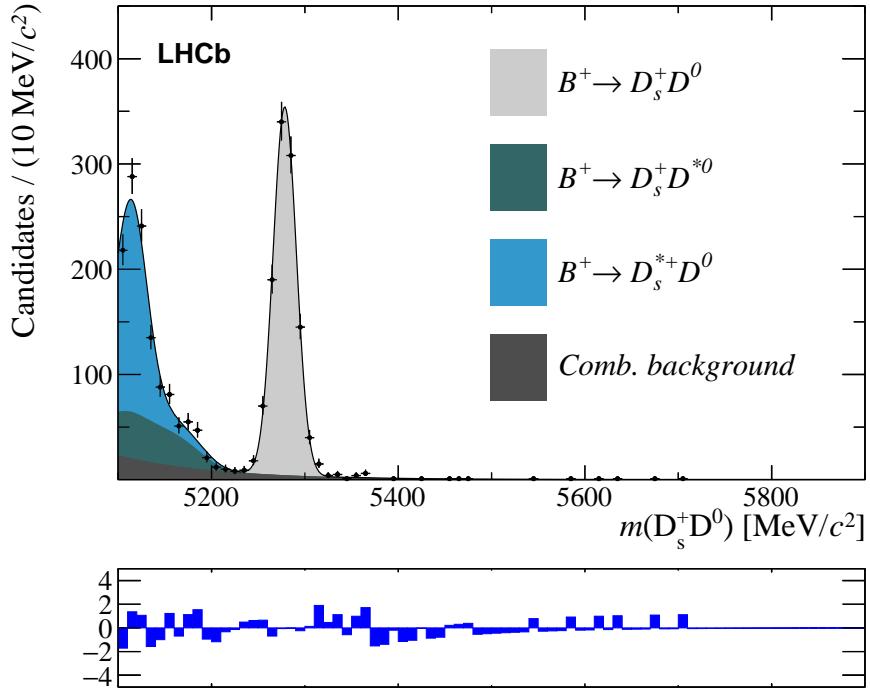
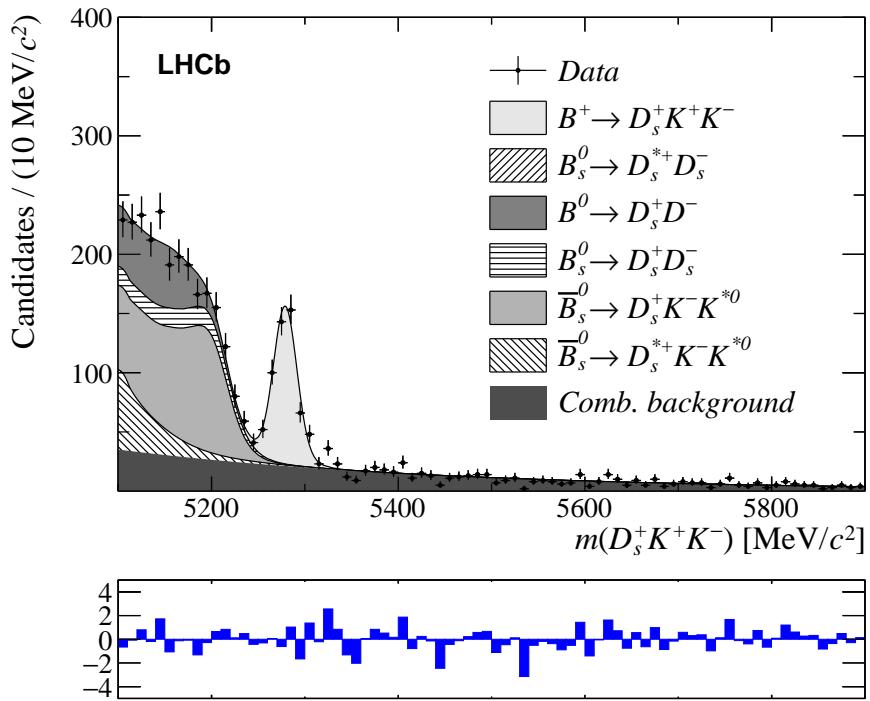
The pull in the branching fraction has a much smaller bias in the width than the normalisation pull. It would be possible to scale the uncertainty in the branching fraction as determined by the fit to account for the residual bias in the uncertainty. However, this would introduce a source of systematic uncertainty associated to scaling factor so the choice is made to not correct the slightly overestimated uncertainty.

## 5.5 Normalisation and signal fits

The signal and normalisation fits are performed using the model and dataset previously described. The results of the two fits are shown in Figs. 5.9 and 5.10. These figures show the distribution of  $B^+$  candidates along with the total model PDF constructed with the values of the free parameters determined in the NLL minimisation process. The contributions from each different component in the model are superimposed, stacked upon one another, and detailed in the legends.

The distribution of  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates in Fig. 5.9 demonstrates the relatively high purity of the selection. Additionally, there is a good separation between the normalisation peak and the partially reconstructed background. Only the combinatorial background is found to have a contribution at the same invariant mass as the normalisation peak.

The distribution of  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates shown in Fig. 5.10 has a significant contribution from the signal decay. The background contribution under the signal PDF is larger than was observed for the normalisation channel. Although the

Figure 5.9: Invariant mass fit to  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates.Figure 5.10: Invariant mass fit to  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates.

dominant background below the signal is still the combinatorial background, the contribution from the partially reconstructed backgrounds is also larger; several of the background PDFs extend underneath the leftmost half of the signal distribution.

The numerical values of the free parameters determined in the fit to the signal and normalisation channels are listed in Tables 5.5 and 5.6.

Type	Parameter	Fit result
POI	$N(B^+ \rightarrow D_s^+ \bar{D}^0)$	$1091 \pm 34$
Shape	Mass shift $\delta$ ( MeV/ $c^2$ )	$-10.6 \pm 1.3$
	Relative peak heights $\xi$	$0.0 \pm 0.2$
	Mean $B^+$ mass $\mu$ ( MeV/ $c^2$ )	$5278.4 \pm 0.4$
	Signal width $\sigma_1$ ( MeV/ $c^2$ )	$12.4 \pm 0.3$
	Combinatorial slope $c$	$(-9.4 \pm 1.1) \times 10^{-3}$
Yields	$N_{\text{comb}}$	$250 \pm 50$
	$N(B^+ \rightarrow D_s^+ \bar{D}^{*0})$	$330 \pm 70$
	$N(B^+ \rightarrow D_s^{*+} \bar{D}^0)$	$750 \pm 50$

Table 5.5: Normalisation fit result

Type	Parameter	Fit result
POI	$N(B^+ \rightarrow D_s^+ K^+ K^-)$	$442 \pm 29$
Shape	Mass shift $\delta$ ( MeV/ $c^2$ )	$4 \pm 12$
	Mean $B^+$ mass $\mu$ ( MeV/ $c^2$ )	$5278.9 \pm 0.9$
	Signal width $\sigma_1$ ( MeV/ $c^2$ )	$12.3 \pm 0.9$
	Combinatorial slope $c$	$(-2.8 \pm 0.3) \times 10^{-3}$
Yields	$N_{\text{comb}}$	$1110 \pm 70$
	$N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})$	$1100 \pm 500$
	$N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0})$	$300 \pm 120$
	$N(B_s^0 \rightarrow D_s^+ D_s^-)$	$200 \pm 500$
	$N(B_s^0 \rightarrow D_s^+ D^-)$	$470 \pm 150$
	$N(B_s^0 \rightarrow D_s^{*+} D_s^-)$	$0 \pm 400$

Table 5.6: Signal fit result

Two of the common parameters between the two modes, namely the mean  $B^+$  mass and width, have consistent values between the two fits. The combinatorial slopes vary between the two, likely because of the different background levels in the two datasets. Similarly the mass shifts vary between the two.

## 5.6 Efficiency corrections

The branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is determined by correcting the yields of signal and background decays by their corresponding selection efficiencies. These account for each step of the selection process and ultimately quantifies the extent to which the signal and normalisation channels are affected differently. As described in Sec. 5.1 the signal yield is corrected by the efficiency as a function of the candidate’s kinematic properties. Therefore the relative signal efficiency is determined as a function of the two-dimensional Dalitz plot coordinates  $m^2(D_s^+ K^-)$  and  $m^2(K^+ K^-)$

$$\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-)) = \frac{\epsilon_{B^+ \rightarrow D_s^+ K^+ K^-}(m^2(D_s^+ K^-), m^2(K^+ K^-))}{\epsilon_{B^+ \rightarrow D_s^+ \bar{D}^0}}. \quad (5.11)$$

The normalisation decay efficiency  $\epsilon_{B^+ \rightarrow D_s^+ \bar{D}^0}$  is a pseudo-two-body decay and therefore the efficiency has no dependence on the phase-space coordinates. The kinematic phase-space is split into bins and the efficiencies determined in each.

The efficiencies are all determined using samples of simulated signal and normalisation decays. The PID and MVA efficiencies require additional input from external calibration samples to correct for the imperfect modelling of the PID variables in the simulations.

### 5.6.1 Efficiencies from simulations

**Acceptance:** this accounts for the likelihood for the five final state tracks to be within the LHCb detector’s acceptance. The charged tracks are required to be in the range  $10 < \cos \theta < 400$  mrad, where  $\theta$  is the angle between the beam direction and the track. The relative efficiency is shown as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  Dalitz plot in Fig.5.11a. The relative efficiency is close to one for most of the phase-space.

**Reconstruction:** the reconstruction efficiency accounts for the fraction of decays in which all five final state tracks have been correctly reconstructed and combined into a  $B^+$  candidate that passes all selection requirements outlined in

the  $B^+ \rightarrow D_s^+ K^+ K^-$  *Stripping Line*. The *Stripping Line* reconstruction also explicitly requires that at least one trigger has fired for the event to be reconstructed, therefore this efficiency also accounts for part of the trigger efficiency. The distribution of the relative reconstruction efficiency between the signal and normalisation mode is shown in Fig. 5.11b. This is greater than one across the whole phase-space as a result of the choice to reconstruct the  $B^+ \rightarrow D_s^+ \bar{D}^0$  normalisation decays using the same selection as the  $B^+ \rightarrow D_s^+ K^+ K^-$  signal. The latter requires the  $D_s^+$ ,  $K^+$  and  $K^-$  candidates to form a well reconstructed vertex where the  $B^+$  meson decays. However, for the normalisation channel the  $\bar{D}^0$  meson travels away from this vertex, so some of the longer lived  $\bar{D}^0$  candidates do not pass the vertex quality requirement. This reduces the efficiency for the normalisation channel relative to the signal.

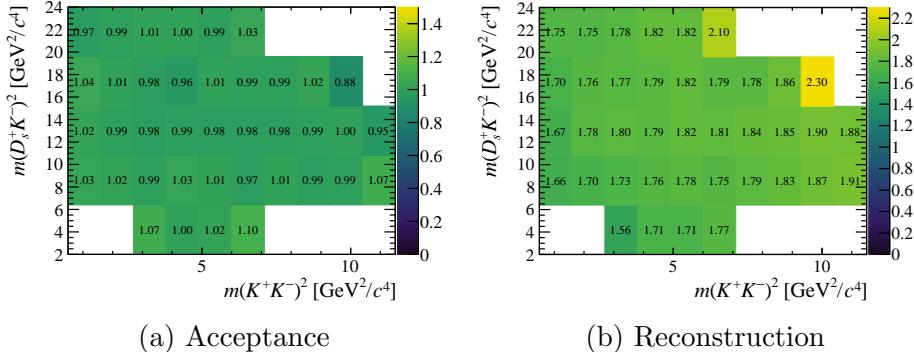


Figure 5.11: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

**Trigger:** the trigger efficiency accounts for the fraction of decays for which the signal candidate meets the TIS and TOS requirements outlined in Sec. 4.2, given that the trigger fired in that event. For the signal and normalisation simulation samples it is very likely that the signal candidate was the cause of the trigger if a trigger fired, therefore these are typically around 95%. The phase-space distribution of the relative efficiency is very uniform and shown in Fig. 5.12a.

**Veto**: this efficiency accounts for the fraction of decays passing the kinematic and normalisation vetoes detailed in Sec. 4.3.5 and 4.3.6. The vetoes to remove misidentified  $D$  and  $\Lambda_c^+$  hadrons are only applied to the  $D_s^+$  meson and therefore assumed to affect the signal and normalisation mode equally. The systematic uncertainty arising from this choice is discussed in Sections 5.7.1. Although the veto for misidentified normalisation decays proceeding via  $\bar{D}^0 \rightarrow K^+ \pi^-$  decays is included, the veto to remove correctly reconstructed normalisation decays with  $\bar{D}^0 \rightarrow K^+ K^-$  is not included. This is because the width of this veto is less than the binning width so it would therefore lead to the wrong efficiency for candidates within in the same bin but not in the window. This doesn't affect the final result, it simply means areas of the efficiency Dalitz plot are populated that will never be sampled. The distribution of the relative efficiency is shown in Fig. 5.12b. A vertical band is clear in the centre as a result of the veto for misidentified normalisation decays with  $\bar{D}^0 \rightarrow K^+ \pi^-$ .

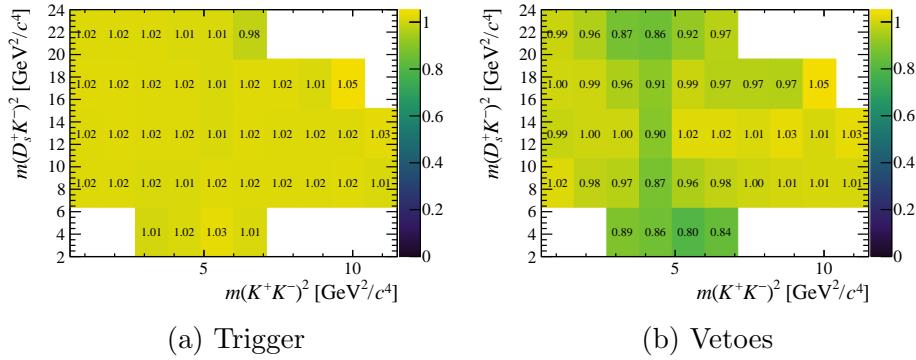


Figure 5.12: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

**Charmless**: this accounts for the fraction of decays failing the requirements on the  $D$  meson flight distance requirements aimed at removing charmless and single-charm backgrounds. The distribution shown in Fig. 5.13a shows a fairly flat distribution with the exception of some of the edge bins.

$\chi_{\text{IP}}^2$ : this accounts for the fraction of decays that pass the requirements on the  $B^+$  and  $D_s^+$  meson impact parameter significance. The relative efficiency is shown in Fig. 5.13b and shows a strong dependency on phase-space at high  $m^2(K^+K^-)$ . The lower relative efficiencies at high  $m^2(K^+K^-)$  may be because the  $D_s^+$  meson is produced almost at rest in this region. Therefore these  $D_s^+$  mesons are likely to have a lower transverse momentum and may fail the minimum impact parameter significance requirement.

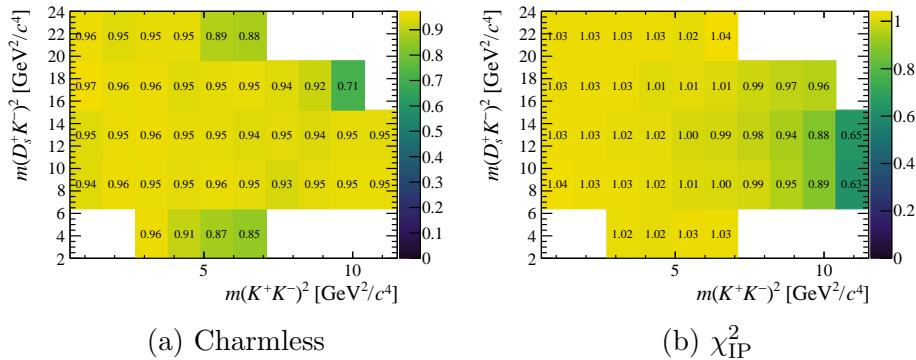


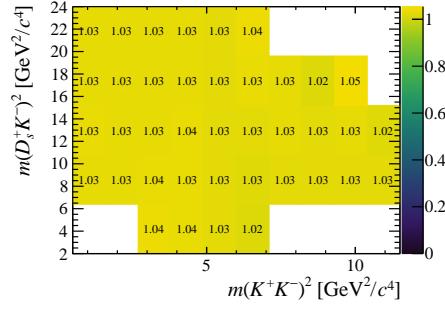
Figure 5.13: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+K^-), m^2(K^+K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

**Mass windows:** this accounts for the fraction of decays that pass the mass windows around the  $D_s^+$  and  $\bar{D}^0$  masses. No requirement is placed on the  $K^+K^-$  pair, therefore the relative efficiency shown in Fig. 5.14a is slightly greater than one across the whole phase-space.

## 5.6.2 Efficiencies requiring calibration samples

### PID efficiency

The efficiencies for the particle identification requirements as described in Sec. 4.3.2 are calculated by correcting the PID variables in simulation using a package called PIDCALIB [121]. This uses calibrations samples for the different particle species to determine the distribution of the PID variables in data. The calibration samples are background-subtracted to isolate the distributions of the PID variables for the tracks



(a) Mass windows

Figure 5.14: The relative efficiency  $\epsilon_{\text{ratio}}(m^2(D_s^+ K^-), m^2(K^+ K^-))$  as a function of the  $B^+ \rightarrow D_s^+ K^+ K^-$  kinematics.

of interest. The calibration samples for both  $K^+$  and  $\pi^+$  mesons are collected from samples of  $D^{*+} \rightarrow (D^0 \rightarrow K^+ \pi^-)\pi^+$  decays, using the decay products of the  $D^0$  decay. Additionally, samples of protons are collected from  $\Lambda \rightarrow p\pi^-$  decays. The PID variable distributions depend on both the kinematics of the track in question and the occupancy of the detector as a whole. These can both affect the characteristics of the hits in the RICH sub-detectors and therefore result in different PID variable distributions. The calibrations samples are characterised using three variables; the transverse momentum of the track  $p_T$ , the pseudo-rapidity  $\eta$ , and the number of tracks in an event  $n_{\text{Tracks}}$ . The calibration PID variables are parametrised using an unbinned approach, unlike in the search for  $B^+ \rightarrow D_s^+ \phi$  where a binned approach is used (Sec. 6.6.2). This unbinned approach creates four dimensional PDFs ( $\text{PID}, \eta, p_T, n_{\text{Tracks}}$ ) from the calibration samples using a kernel density estimation implemented with the MEERKAT package [122]. Similar four dimensional PDFs are created for simulation samples of the calibration modes. These two PDFs are used to calculate the transformation needed for each species as a function of  $(\text{PID}, \eta, p_T, n_{\text{Tracks}})$  to correct the distribution in simulation. This transformation preserves the correlations between the different PID variables and between the PID and kinematic properties. This correction is applied to both the DLL PID variables and ProbNNx variables used in the MVA selection. The phase-space distribution of the relative PID efficiency between the signal and normalisation channel is shown in Fig. 5.15a.

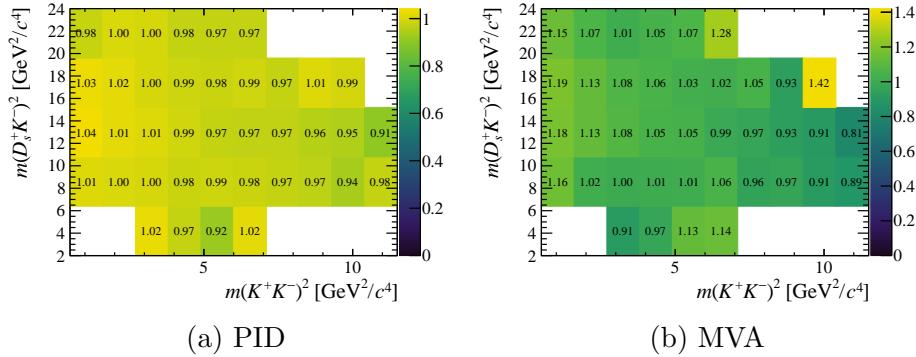


Figure 5.15: Efficiencies

### MVA efficiency

The MVA efficiency is determined from samples for signal and normalisation simulations in which the PID distributions have been corrected as described previously. This is different to the approach used in the search for  $B^+ \rightarrow D_s^+ \phi$  decays (Sec. 6.6.2) and is required to account for the variation in the efficiency across the phase-space. The corrected ProbNNx PID variables are used to determine the  $D_s^+$  and  $\phi$  MVA classifiers and the efficiency of the MVA requirements determined in different positions in the Dalitz plot as shown in Fig. 5.15b. There is a strong dependence on the position, with decays at low  $m^2(K^+ K^-)$  values having a larger relative efficiency than those at higher values. This is likely to be because the MVA methods were trained using a sample of  $\phi \rightarrow K^+ K^-$  decays, so the selection favourably selects candidates with an invariant mass near the  $\phi$  meson mass. A cross check is performed to compare the MVA efficiencies determined directly using the classifier produced with the corrected PID variables and the method described in Sec. 4.3.7 that uses the MVA training samples to calculate the efficiency. The values of the efficiency determined for each mode are compared in Fig. 5.16 as a function of the  $m(K^+ K^-)$  mass. The MVA training modes only exist for two discreet values;  $m(K^+ K^-) = m(\phi)$  and  $m(K^+ K^-) = m(\bar{D}^0)$  represented by the red points. These are compared to the  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s \bar{D}^0$  efficiencies from the corrected simulations in black and blue respectively. A good agreement is found.

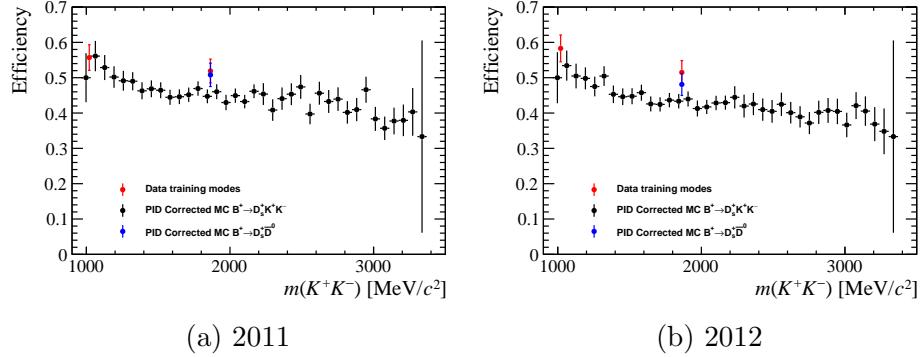


Figure 5.16: The PID efficiency variation as a function  $m(K^+K^-)$  as determined from the MVA training modes (red) and the  $B^+ \rightarrow D_s^+ K^+ K^-$  and  $B^+ \rightarrow D_s \bar{D}^0$  simulation samples (black and blue respectively). A good agreement is found between the red and black point at  $m(K^+K^-) = m(\phi)$ , and the red and blue point at  $m(K^+K^-) = m(\bar{D}^0)$ .

### 5.6.3 Total efficiency

The total relative efficiency between the signal and normalisation decays is calculated from the products of each contributing efficiency

$$\epsilon^{\text{Tot.}} = \epsilon^{\text{Accp.}} \times \epsilon^{\text{Reco.}|\text{Accp.}} \times \epsilon^{\text{Trig.}|\text{Reco.}} \times \epsilon^{\text{Mass.}|\text{Trig.}} \times \epsilon^{\text{Veto.}|\text{Mass.}} \times \epsilon^{\text{FD}|\text{Veto.}} \\ \times \epsilon^{\text{IP}|\text{FD}} \times \epsilon^{\text{PID}|\text{IP}} \times \epsilon^{\text{MVA}|\text{PID}}, \quad (5.12)$$

where each relative efficiency  $x$  is defined relative to the previous selection step  $y$  as  $\epsilon^{x|y}$ . This is similarly determined as a function of the two-dimensional  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  space and shown in Fig. 5.17a. This two dimensional histogram has a rather coarse binning and is susceptible to variations from statistical fluctuation.

This could lead to biases when correcting the weights of signal candidates, especially if the candidates are concentrated within a one or a few bins. To reduce this the two dimensional histogram is made into a smoothly varying distribution by using a cubic spline interpolation as shown in Fig. 5.17b. The final distribution used to correct the yields of signal decays is the weighted sum of the splines for each year of data taking.

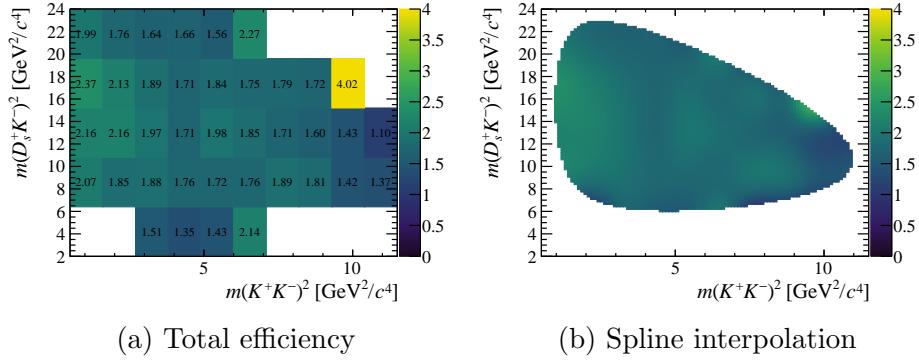


Figure 5.17: The total relative efficiency between the signal and normalisation mode as a function of the two-dimensional  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  (right) and a cubic spline interpolation of the same distribution.

## 5.7 Systematic uncertainties

A number of sources of systematic uncertainty are considered when determining the branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays.

### 5.7.1 Relative efficiencies

The yields of signal and normalisation decays are corrected by the ratio of selection efficiencies. The limitations and assumptions surrounding this ratio contribute to the systematic uncertainty.

**Simulation statistics:** the relative efficiencies used to correct the signal and normalisation yields in this analysis are determined from samples of simulated decays. These are of a limited size, therefore the relative efficiencies have an associated statistical uncertainty. The size of the simulation samples results in a 2% systematic uncertainty.

**Particle identification:** the efficiencies for the particle identification requirements are determined from calibration samples using the PIDCALIB package. The procedure used to determine the efficiencies from these samples contains sources of systematic uncertainty, for example the choice of binning scheme used to match the calibration samples to the signal decays and the background-subtraction procedure used to isolate the distributions. The presence of the same particle

species in the signal and normalisation mode helps to reduce this affect in the ratio of efficiencies to some extent, but not completely. A relatively conservative estimate of the systematic uncertainty of 0.5% per track is assigned to account any residual differences. This is assumed to be uncorrelated for the five tracks contributing to the signal channel and five track contributing to the normalisation channel, resulting in a total relative uncertainty of 2.0%.

**Veto efficiency:** the efficiency of the misidentified  $D$  and  $\Lambda_c^+$  hadron veto applied to the  $D_s^+$  meson is assumed to be the same for the signal and normalisation channel. However, slight differences in the kinematics of the  $D_s^+$  mesons in each case might lead to this approximation not being valid when the particle identification requirements are tightened on the ambiguous track. The relative veto efficiency could be calculated by determining the relative PID efficiency for the sub-set of particles that would be affected by the tightening of the cuts. Rather than calculating this ratio, the systematic uncertainty that would be associated with it is assigned to account for any bias. This totals 1.4%, calculated in a similar way to the PID relative efficiency, except with three tracks instead of five.

**MVA efficiency:** the relative efficiency of the MVA selection receives contributions from a number of sources. Primarily, the relative efficiencies are determined from simulation samples in which the PID variable distributions have been corrected as described in Sec. 5.6.2. This method of correcting the PID distributions such that the simulation samples could be used directly was only available for Run I simulations at the time of analysis. The phase-space distributions of the Run II samples were assumed to be the same, but the overall scale adjusted using the MVA efficiencies derived from the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decay modes used to train the MVAs. The systematic uncertainty arising from using the Run I variations in Run II are estimated by calculating the largest difference between the signal and normalisation efficiencies in

Run II and subtracting the smallest differences in Run I. This corresponds to a variation of 4.4%, assigned as the systematic uncertainty.

As they contribute to the scaling of the phase-space variations, all sources of systematic uncertainty contributing to the data-driven efficiencies are also considered. The training modes are binned in four bins of  $p_T$  and  $\chi_{\text{FD}}^2$ . The choice of binning scheme is varied and the resulting variation in efficiency determined. The yields of the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  decays used to obtain the data-driven efficiencies are limited, therefore the quantity  $1/\sqrt{N}$  associated to the smallest of these yields is added as a systematic uncertainty.

Differences are observed in the distribution of  $\chi_{\text{IP}}^2$  in simulations and data for the normalisation channel (Fig. 4.18). To quantify the effect this may have on the MVA efficiencies, the simulations are re-weighted to match the data distributions and the efficiencies recalculated. The resulting difference is included as a systematic uncertainty.

When training the MVA methods some discrepancies are observed between the MVA classifier response for the training and validation samples (Fig. 4.15). To assess whether this affects the final relative MVA efficiencies the training and validation samples are swapped and the efficiencies recalculated. The differences for each year of data taking are weighted according to their contribution to the final data set and assigned as a systematic uncertainty.

The total systematic uncertainty associated to the relative MVA efficiency is therefore 7.6%.

### 5.7.2 Signal and normalisation PDFs

Some parameters in the signal and normalisation PDFs are fixed to values obtained from simulation. These include the tail parameters, relative widths, and fractional amounts of the two CB functions that make up the PDFs. The values obtained from simulation have associated uncertainties arising from the limited simulation sample sizes. The nominal fits are repeated with the fixed parameters modified to

values sampled from Gaussian distributions, with a width given by the parameter uncertainties. All parameters are changed simultaneously. The resulting variation of  $0.036 \times 10^{-6}$  is assigned as the associated systematic uncertainty.

### 5.7.3 Background PDFs

In the signal mode the PDFs for the partially reconstructed background modes are taken directly from simulated events using one-dimensional kernel estimations [119]. In the nominal fit, these are smeared to account for the differences in the mass resolution between data and simulation. To account for any systematic uncertainty arising from the choice of resolution difference, the fit is repeated, randomly varying the smearing resolution each time. The resulting variation in the branching fraction is assigned as a systematic uncertainty.

In the normalisation mode a number of choices are made about the partially reconstructed backgrounds, namely the kinematic limits of the  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$ , and the external branching fractions of  $D_s^{*+} \rightarrow D_s^+ [\gamma/\pi^0]$ ,  $\bar{D}^{*0} \rightarrow \bar{D}^0 [\gamma/\pi^0]$  decays. These assumptions are simultaneously varied within the relevant uncertainties, resulting in a change of  $0.015 \times 10^{-6}$  in the branching fraction.

### 5.7.4 Charmless contribution

The residual charmless and single charm contribution is only a significant fraction of the measured yield for the normalisation mode as detailed in Table 4.9. The residual yield of about eight candidates corresponds to a relative difference of 0.7% of the measured normalisation yield. This, however, would be a slight overestimation as the refitting of the decay chain with a constrained  $D_s^+$  mass would alter the resolution of the normalisation and charmless backgrounds differently. This has the effect of smearing the charmless candidates over a wider range. Rather than trying to include a PDF for this contribution or correcting the measured yield the expected size of the effect is used as the systematic uncertainty as it would propagate directly to the branching fraction.

### 5.7.5 Total systematic uncertainty

The systematic uncertainties contributed from each source discussed is detailed in Table 5.7. The total systematic uncertainty is also shown, along with the uncertainty arising from the externally measured normalisation channel branching fractions.

Source of Uncertainty	Systematic Uncertainty	
	Relative	Absolute ( $\times 10^{-6}$ )
MVA relative efficiency	7.6%	0.53
Simulation statistics	2.0%	0.14
PID relative efficiency	2.0%	0.14
Veto relative efficiency	1.4%	0.10
Charmless contribution	0.7%	0.05
Signal PDF parametrisation	-	0.036
Background PDF parametrisation	-	0.015
Total		0.59
Normalisation		0.70

Table 5.7: Contributions to the total systematic uncertainty of the  $B^+ \rightarrow D_s^+ K^+ K^-$  branching fraction measurement. The contributions are listed in descending order.

## 5.8 Results

The fit to  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates finds a total yield of  $N(B^+ \rightarrow D_s^+ K^+ K^-) = 443 \pm 29$  candidates. This constitutes the first observation of this decay mode. The branching fraction is calculated as

$$\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-) = \frac{N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)}{N(B^+ \rightarrow D_s^+ \bar{D}^0)} \times \mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) \times \mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-) \quad (5.13)$$

where  $N(B^+ \rightarrow D_s^+ \bar{D}^0)$  is the yield of normalisation decays, and  $N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)$  is defined to be

$$N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-) = \sum_i \frac{W_i}{\epsilon_i^{\text{ratio}}}, \quad (5.14)$$

where  $W_i$  is the per-candidate weight, as determined by the *sPlot* technique for candidate  $i$ ; and  $\epsilon_i^{\text{ratio}}$  represents the relative efficiency of the signal and normalisation modes  $\epsilon_i(B^+ \rightarrow D_s^+ K^+ K^-)/\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)$  in the relevant bin of the  $B^+ \rightarrow D_s^+ K^+ K^-$  Dalitz plot as calculated in Sec. 5.6.

The uncertainty on the corrected yield,  $N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)$ , is in principle given by

$$\sigma(N_{\text{corr}}) = \sqrt{\sum_i \left( \frac{W_i}{\epsilon_i^{\text{ratio}}} \right)^2}. \quad (5.15)$$

However, the fit used to determine the *sWeights* only allows the yields to float, as opposed to the nominal fit that has additional floating parameters including the signal position and width. This means that this estimate of the corrected yield uncertainty could neglect the uncertainty due to the shape parameters: *i.e.* the uncertainty calculated from the weights,  $\sigma_{\text{yields only}}(N) = \sqrt{\sum W_i^2}$ , can be less than the uncertainty returned by the nominal fit,  $\sigma_{\text{fit}}(N)$ .

To correctly account for this possibility, the uncertainty from the shape parameters is separated from the total uncertainty:  $\sigma_{\text{shape}}(N) = \sqrt{\sigma_{\text{fit}}(N)^2 - \sigma_{\text{yields only}}(N)^2}$ . This extra uncertainty is then scaled by the corrected yield to give the total uncertainty

$$\sigma_{\text{corr}}(N_{\text{corr}}) = \sqrt{\sigma(N_{\text{corr}})^2 + \left( \frac{N_{\text{corr}}}{N} \sigma_{\text{shape}}(N) \right)^2}. \quad (5.16)$$

For the nominal fit the uncertainties are summarised in Table 5.8.

Uncertainty	Value
$\sigma(N_{\text{corr}})$	13.0
$\sigma_{\text{fit}}(N)$	29.4
$\sigma_{\text{yields only}}(N)$	25.4
$\sigma_{\text{shape}}(N)$	14.7
$\sigma_{\text{corr}}(N_{\text{corr}})$	14.9

Table 5.8: The various uncertainties as detailed in Section 5.8 and their values in the nominal fit.

The corrected yield ratio can be expressed as the ratio of signal and normalisation branching fractions using Eq. 5.13. The value is measured to be

$$\frac{N_{\text{corr}}(B^+ \rightarrow D_s^+ K^+ K^-)}{N(B^+ \rightarrow D_s^+ \bar{D}^0)} = \frac{\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)}{\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) \mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)} = 0.197 \pm 0.015 \pm 0.017, \quad (5.17)$$

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

The branching fraction for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays is determined to be

$$\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-) = (7.1 \pm 0.5 \pm 0.6 \pm 0.7) \times 10^{-6}, \quad (5.18)$$

where the first uncertainty is statistical, the second is systematic and the third from the branching fractions of  $\bar{D}^0 \rightarrow K^+ K^-$  and of the normalisation mode  $B^+ \rightarrow D_s^+ \bar{D}^0$ . The values used for the branching fractions are  $\mathcal{B}(D^0 \rightarrow K^+ K^-) = (4.01 \pm 0.07) \times 10^{-3}$  and  $\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) = (9.0 \pm 0.9) \times 10^{-3}$  [28].

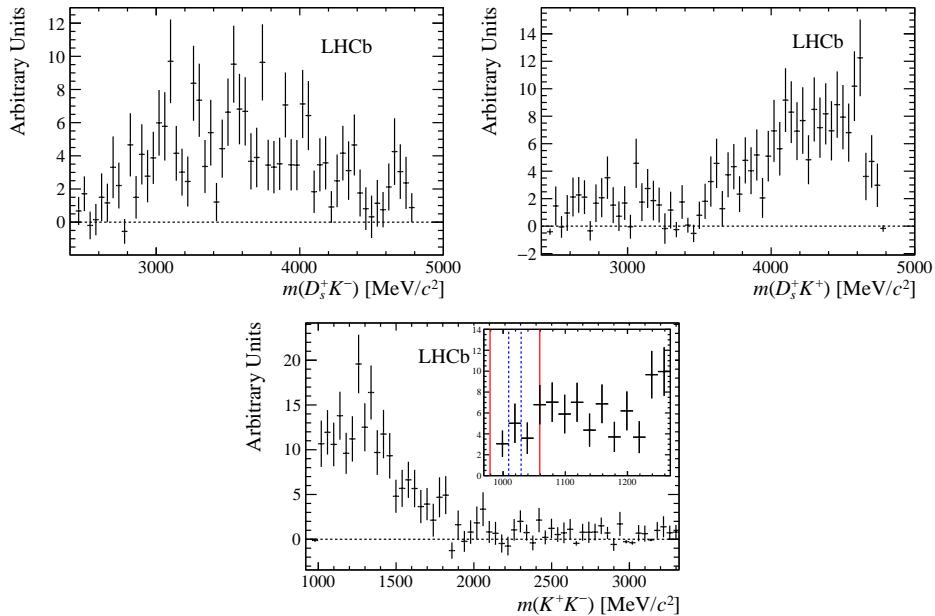


Figure 5.18: Two-body mass projections of the background-subtracted efficiency-corrected  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates. An inset has been included expanding the region around the  $\phi$  meson mass using the same ranges defined in the search for  $B^+ \rightarrow D_s^+ \phi$  decays in Fig. 6.1.

The two-body projections  $m(D_s^+ K^-)$  and  $m(K^+ K^-)$  are obtained for the signal component using the *sPlot* technique, shown in Fig. 5.18. No significant peak is observed in the  $\phi$  region of the  $m(K^+ K^-)$  plot; rather a broad distribution of candidates is found in the region up to  $m(K^+ K^-) \simeq 1900 \text{ MeV}/c^2$ . These plots have been corrected to account for the variation in the relative efficiencies as a function of the phase-space.

Additionally, the distribution of background-subtracted efficiency-corrected  $B^+ \rightarrow D_s^+ K^+ K^-$  decays are shown as a function of the two-dimensional space  $m^2(D_s^+ K^-)$  vs.  $m^2(K^+ K^-)$  in Fig. 5.19. The candidates can be seen to be localised to a small area of the phase-space.

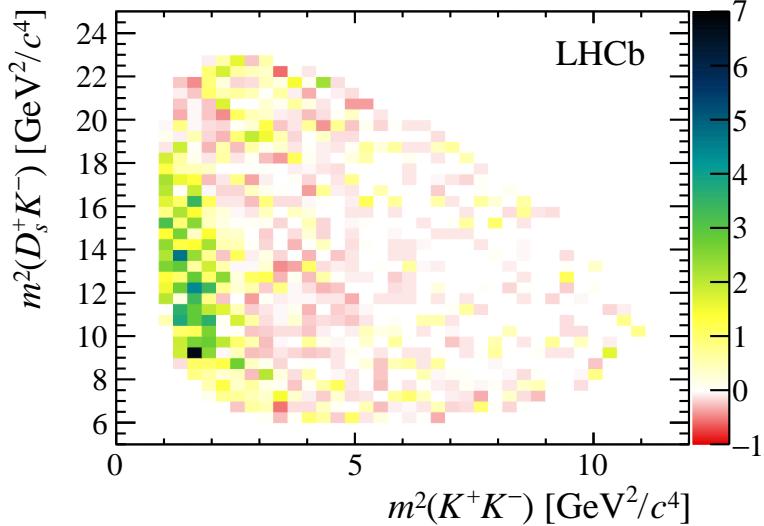


Figure 5.19: The Dalitz plot distribution of background-subtracted efficiency-corrected  $B^+ \rightarrow D_s^+ K^+ K^-$  decays.

The background-subtracted distribution of the angle,  $\cos \theta_K$ , defined to be the angle between the direction of the  $K^+$  meson and  $B^+$  momentum vector in the  $K^+ K^-$  rest frame is shown in Fig. 5.20. The distribution of this angle can change depending of the spin of the  $K^+ K^-$  system and therefore may help identify the contributing processes.

Additionally, the distribution of the angle  $\theta_K$  is plotted in bins of  $m(K^+ K^-)$  in Fig. 5.21. The distribution is observed to change as a function of  $m(K^+ K^-)$ , which may be a result of interference between competing processes.

## 5.9 Outlook

This analysis provides an initial insight to the possible processes that might contribute to the  $B^+ \rightarrow D_s^+ K^+ K^-$  decay. However, a full amplitude analysis would be needed to rigorously study the competing decays and account for any possible interference

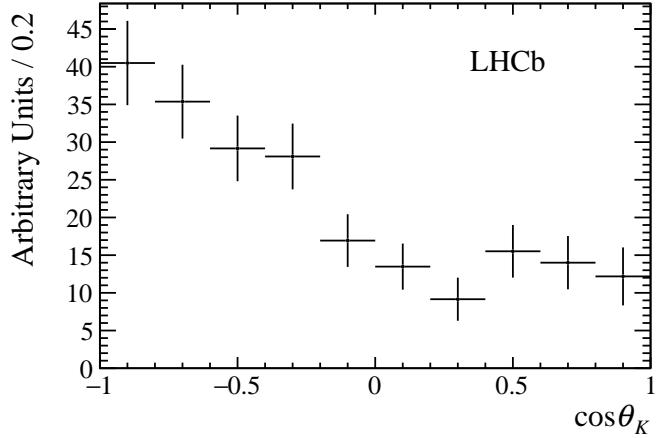


Figure 5.20: The background-subtracted efficiency-corrected distribution of  $\cos\theta_K$  for  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates.

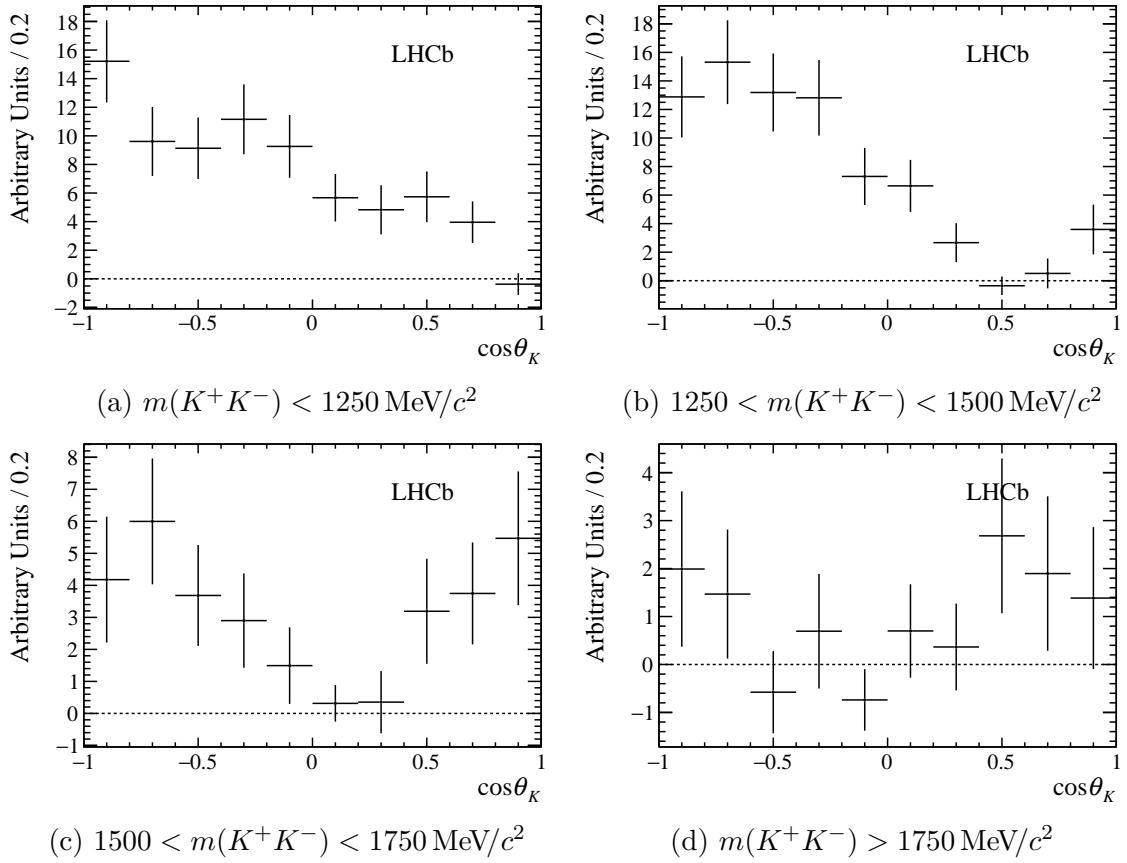


Figure 5.21: The distribution of  $\cos\theta_K$  in bins of  $m(K^+K^-)$  mass.

between the amplitudes. This search contains only a fraction of the Run II data sample; the inclusion of the 2017 and 2018 samples would provide a statistically solid

base to further this line of investigation.

# Chapter 6

## Mass fit to $B^+ \rightarrow D_s^+ \phi$ candidates

### Contents

---

<b>6.1</b>	<b>Fit strategy</b>	<b>146</b>
6.1.1	Simultaneous categories	147
6.1.2	$B^+ \rightarrow D_s^+ K^+ K^-$ model and assumptions	149
<b>6.2</b>	<b>Fit components</b>	<b>155</b>
6.2.1	Signal and normalisation decays	156
6.2.2	Partially reconstructed backgrounds	157
6.2.3	Combinatorial backgrounds	165
<b>6.3</b>	<b>Free and constrained parameters</b>	<b>165</b>
6.3.1	Parameter of interest	165
6.3.2	Shape parameters	166
6.3.3	Yields	167
6.3.4	Fractions	167
<b>6.4</b>	<b>Fit validation</b>	<b>169</b>
<b>6.5</b>	<b>Normalisation and signal fits</b>	<b>172</b>
<b>6.6</b>	<b>Efficiency corrections</b>	<b>175</b>
6.6.1	Efficiencies from simulation	175
6.6.2	Efficiencies requiring calibration samples	178
6.6.3	Total efficiencies	179
<b>6.7</b>	<b>Systematic uncertainties</b>	<b>180</b>
6.7.1	Relative efficiencies	180
6.7.2	Signal and normalisation PDFs	182
6.7.3	Background PDFs	182
6.7.4	Charmless contribution	183
6.7.5	$B^+ \rightarrow D_s^+ K^+ K^-$ model assumption	183

6.7.6	Total systematic uncertainty . . . . .	184
<b>6.8</b>	<b>Results . . . . .</b>	<b>184</b>
6.8.1	Limit setting . . . . .	186
6.8.2	Comparison to the previous measurement . . . . .	190

---

In this chapter the methodology used to search for  $B^+ \rightarrow D_s^+ \phi$  decays is described. The branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  is constructed by measuring the yield of  $B^+ \rightarrow D_s^+ \phi$  decays relative to the normalisation channel  $B^+ \rightarrow D_s^+ \bar{D}^0$ . This ratio is corrected by the ratio of selection efficiencies for the two modes. The branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  is then determined by multiplying this corrected ratio by externally measured values for the branching fractions  $\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0)$  and  $\mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)$ , and dividing by  $\mathcal{B}(\phi \rightarrow K^+ K^-)$ .

## 6.1 Fit strategy

The strategy used to search for  $B^+ \rightarrow D_s^+ \phi$  is more complicated than the method used to search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays as outlined in Chapter 5. This is necessary to allow various different signal and background components to be distinguished. In particular, the majority of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays now act as a background; those not proceeding via the  $\phi$  resonance must be distinguished from  $B^+ \rightarrow D_s^+ \phi$  decays. The ratio of  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  yields is determined using a simultaneous extended unbinned maximum likelihood fit. Three sets of categories are used, separating the candidates according to  $D_s^+$  meson decay mode, invariant mass of the  $K^+ K^-$  pair constituting the  $\phi$  meson,  $m(K^+ K^-)$ , and the cosine of an angle  $\cos \theta_K$ . The details and definitions of these categories are listed in Sec 6.1.1. The total extended NLL for this fit is created from the sum of each NLL in each of the categories

$$-\log \mathcal{L}(n_0 \dots n_j, \vec{p}) = \sum_{\alpha} \sum_{\beta} \sum_{\gamma} \left( -\log \mathcal{L}^{\alpha, \beta, \gamma}(n_0^{\alpha, \beta, \gamma} \dots n_j^{\alpha, \beta, \gamma}, \vec{p}) \right) \quad (6.1)$$

where  $\alpha$ ,  $\beta$  and  $\gamma$  represent indexes over the  $D_s^+$  mode,  $m(K^+K^-)$  and  $\cos\theta_K$  categories. The NLL for each category is defined as

$$-\log \mathcal{L}^{\alpha,\beta,\gamma}(n_0^{\alpha,\beta,\gamma} \dots n_j^{\alpha,\beta,\gamma}, \vec{p}) = -\sum_i^{N^{\alpha,\beta,\gamma}} \log \left( \sum_j n_j^{\alpha,\beta,\gamma} f_j^{\alpha,\beta,\gamma}(m = m_i, \vec{p}) \right) + \sum_j n_j^{\alpha,\beta,\gamma}. \quad (6.2)$$

As before,  $j$  represents the index over each contribution to the fit model, and  $i$  represents each of  $N^{\alpha,\beta,\gamma}$  entries in the data set for category  $\alpha, \beta, \gamma$ . The composite extended NLL is minimised with respect to the parameters  $\vec{p}$  to find the values for which the data is most likely.

The fit is implemented using the RooFit package within the ROOT framework. As the yields of candidates is likely to be small (especially in the  $D_s^+$  decay modes with smaller branching fractions) the error on the branching fraction is determined asymmetrically using the MINOS procedure. This determines the error intervals by finding the positions either side of the minimum that the NLL value increases by one unit. As the NLL distribution may not be symmetric this can lead to different upper and lower uncertainties. However, this calculation is computationally expensive, especially in this large fit with many degrees of freedom.

### 6.1.1 Simultaneous categories

The data sample is split into categories primarily to aid the differentiation of different signal and background components that contribute in similar invariant mass ranges.

#### $D_s^+$ meson decay mode

The three  $D_s^+$  decays modes used to reconstruct the signal and normalisation decays ( $D_s^+ \rightarrow K^+K^-\pi^+$ ,  $D_s^+ \rightarrow \pi^+\pi^-\pi^+$ , and  $D_s^+ \rightarrow K^+\pi^-\pi^+$ ) are fitted simultaneously in different categories. This allows the invariant mass distributions for the three modes to vary slightly in ways that could not be easily accounted for if the modes were combined in a single data set. In principle, the widths and resolutions of the  $B^+$  meson mass distributions could vary for the three different modes as a result of the different numbers of pions and kaons in the final state. The background levels also differ

between the modes as a result of the smaller branching fractions for  $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$  and  $D_s^+ \rightarrow K^+ \pi^- \pi^+$ . This leads to the background from combinations of unrelated tracks having a larger relative contribution.

The  $D_s^+ \rightarrow K^+ K^- \pi^+$  decay mode is additionally split into two further categories; candidates consistent with  $D_s^+ \rightarrow \phi \pi^+$  decays, and non- $\phi$  candidates. This exploits the high purity of  $D_s^+ \rightarrow \phi \pi^+$  decays.

### Invariant mass of $K^+ K^-$ pair, $m(K^+ K^-)$

Three distinct ranges of  $m(K^+ K^-)$  invariant mass are used to split the candidates. The first of these corresponds to normalisation  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates within the range  $|m(K^+ K^-) - m(\bar{D}^0)| < 25 \text{ MeV}/c^2$ . These are reconstructed separately to the signal decays, unlike in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays, as detailed in Chapter 4. The candidates reconstructed as  $B^+ \rightarrow D_s^+ \phi$  decays are split into two ranges; those within  $|m(K^+ K^-) - m(\phi)| < 10 \text{ MeV}/c^2$ , referred to as the *inner  $\phi$  mass category* and those candidates with  $10 < |m(K^+ K^-) - m(\phi)| < 40 \text{ MeV}/c^2$ , referred to as the *outer  $\phi$  mass category*. These two categories for the signal mode allow contributions from decays that do not proceed via a  $\phi$  meson to be distinguished from those that do. The *inner  $\phi$  mass category* contains 88% of signal  $B^+ \rightarrow D_s^+ \phi$  candidates, with the other 12% in the *outer  $\phi$  mass category*. The  $m(K^+ K^-)$  invariant mass distribution for simulated  $B^+ \rightarrow D_s^+ \phi$  is shown in Fig 6.1.

### Helicity angle, $\cos \theta_K$

The  $B^+ \rightarrow D_s^+ \phi$  decay involves the decay of a pseudoscalar particle to a pseudoscalar and vector particle. Therefore the  $\phi$  vector meson ( $J^P = 1^-$ ) must be produced longitudinally polarised. For a longitudinally polarised  $\phi$  meson decaying to  $K^+ K^-$ , the distribution of the angle  $\theta_K$ , defined as the angle that the kaon meson forms with the  $B^+$  momentum in the  $\phi$  rest frame (Fig. 6.2), is proportional to  $\cos^2 \theta_K$ . The distribution of  $\cos \theta_K$  for  $B^+ \rightarrow D_s^+ \phi$  as determined from simulated events is shown in Fig 6.1. Candidates are split into two categories;  $|\cos \theta_K| > 0.4$  and  $|\cos \theta_K| < 0.4$ . These categories contain 93% and 7% of the signal respectively.

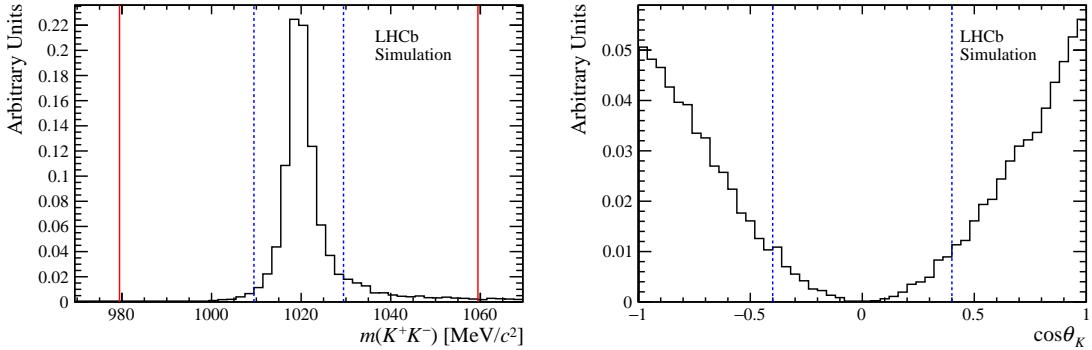


Figure 6.1: The distributions of  $m(K^+K^-)$  (left) and  $\cos\theta_K$  (right) in simulated  $B^+ \rightarrow D_s^+\phi$  decays. The vertical blue dashed lines represent the boundaries between categories defined in Sec. 6.1.1. The vertical red lines represent the mass window applied to candidates; those outside the red lines are not included in the data set.

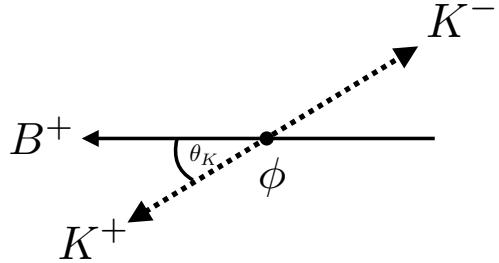


Figure 6.2: The angle  $\theta_K$  (referred to as the helicity angle) is defined to be the angle that the kaon mesons forms with the  $B^+$  meson momentum in the  $\phi$  rest frame.

This helicity angle is constructed using the momentum of the decay products calculated after the whole decay chain has been refitted with a  $D_s^+$  mass and  $B^+$  direction constraint. This significantly increases the fraction of signal events expected in the first of the two categories.

### 6.1.2 $B^+ \rightarrow D_s^+ K^+ K^-$ model and assumptions

The search for  $B^+ \rightarrow D_s^+\phi$  decays includes a component for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays that didn't proceed via a  $\phi$  meson. This is necessary as the search documented in Chapter 5 determined there is a non-zero contribution from these decays in the range of  $m(K^+K^-)$  invariant mass considered here (Fig 5.18). To avoid overestimating  $B^+ \rightarrow D_s^+\phi$  signal yield, separate components are included in the fit model for the

$ m(K^+K^-) - m_\phi $ ( MeV/ $c^2$ )	Helicity Category	
	$ \cos \theta_K  > 0.4$	$ \cos \theta_K  < 0.4$
< 10	82%	6%
(10, 40)	11%	1%

Table 6.1: Fractions of  $B^+ \rightarrow D_s^+ \phi$  candidates expected in the helicity and  $m(K^+K^-)$  invariant mass categories of the simultaneous fit.

$B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ K^+ K^-$  decays. Although the invariant mass distributions of these contributions are identical, they can still be disentangled by exploiting the different fractions of these decays expected in each of the helicity angle and  $m(K^+K^-)$  categories. The fractions for the  $B^+ \rightarrow D_s^+ \phi$  signal decays as listed in Table 6.1 show the decays are concentrated in the *inner  $\phi$  mass category* with  $|\cos \theta_K| > 0.4$ . To determine similar fractions for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays the LAURA++ package [123] is used to generate a number of simulation samples for different intermediate resonance models.

Only resonances in the  $K^+K^-$  system are considered as no significant structure is observed in the  $m(D_s^+ K^-)$  distribution in Fig. 5.18. As such, all resonances are neutral mesons. The models are generated separately, therefore the effect of interference between any combination of states has been entirely neglected. The generated samples are described in the following sections.

### The $\phi(1020)$ resonance

Decays proceeding via a  $\phi(1020)$  are produced as a crosscheck. As the simulations generated with LAURA++ have not been reconstructed with the full LHCb detector model, this sample is compared to the existing full simulation samples. The differences in between the fraction of the decays in the different  $m(K^+K^-)$  and  $\cos \theta_K$  categories in the two samples are taken as a proxy for the potential level of bias introduced by using these generator level samples instead of full simulation. The distribution of these simulated decays in  $m(K^+K^-)$  and  $\cos \theta_K$  are shown in Fig. 6.3. This figure also include the Dalitz plot distribution of the decays parametrised with the variables

$m^2(D_s^+ K^-)$  and  $m^2(K^+ K^-)$ . This resonance is generated with a Relativistic Breit-Wigner line shape [124].

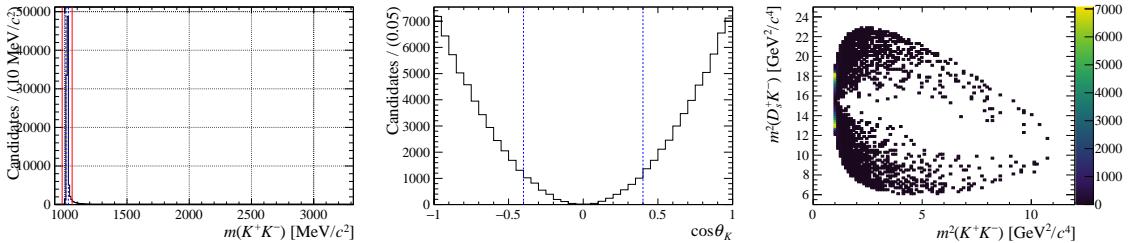


Figure 6.3: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos\theta_K$  for generated for the  $\phi(1020)$  resonance.

### Non-resonant decays

In addition to  $K^+K^-$  resonances, a non-resonant model is considered. This model is defined to have a uniform amplitude across the allowed phase-space. The distribution in  $m(K^+K^-)$  of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays in Fig. 5.18 is not consistent with this model as there are no candidates above  $m(K^+K^-) \sim 1900 \text{ MeV}/c^2$ . However, this component is included in this study for comparative purposes. The distributions of decays generated with this flat model are shown in Fig. 6.4.

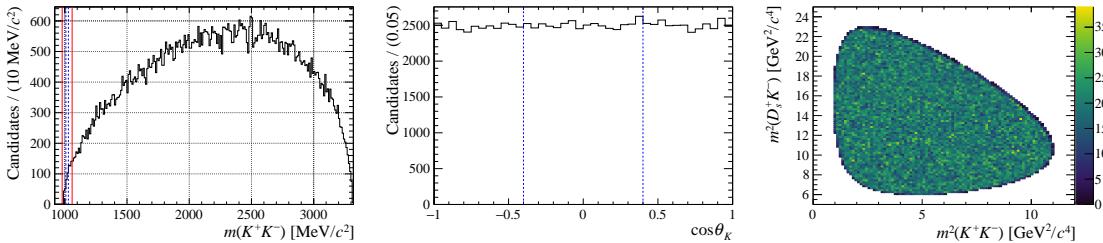


Figure 6.4: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos\theta_K$  for generated for non-resonant decays.

### The $f_0^0(980)$ resonance

The  $f_0^0(980)$  resonance is a light unflavoured  $J^P = 0^+$  state with mass  $990 \pm 20 \text{ MeV}/c^2$  and width  $10\text{--}100 \text{ MeV}/c^2$  [28]. It has been observed to decay to  $K^+K^-$  making it a suitable resonance to consider. Although its mass is at the lower end of the range considered here, its significant width allows it to contribute at higher invariant

masses. This component is modelled with the Flatté line shape [125] and the relevant distributions shown in Fig. 6.5.

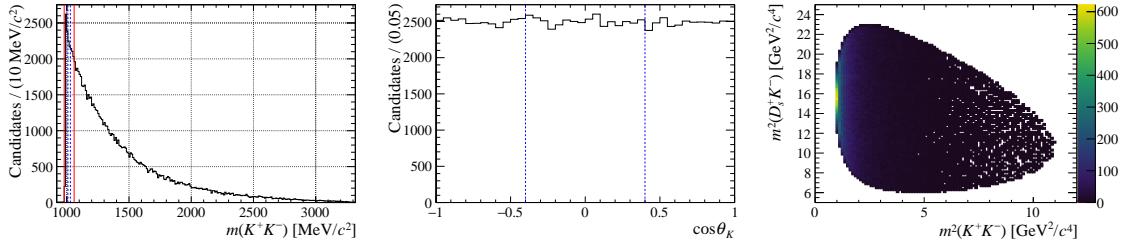


Figure 6.5: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos\theta_K$  for generated for the  $f_0^0(980)$  resonance.

### The $a_0^0(980)$ resonance

The  $a_0^0(980)$  resonance is a light unflavoured  $J^P = 0^+$  state with mass  $980 \pm 20 \text{ MeV}/c^2$  and width 50–100  $\text{MeV}/c^2$  and has been observed to decay to  $K\bar{K}$  final states [28]. This resonance is also modelled with the Flatté line shape and the relevant distributions are shown in Fig 6.6.

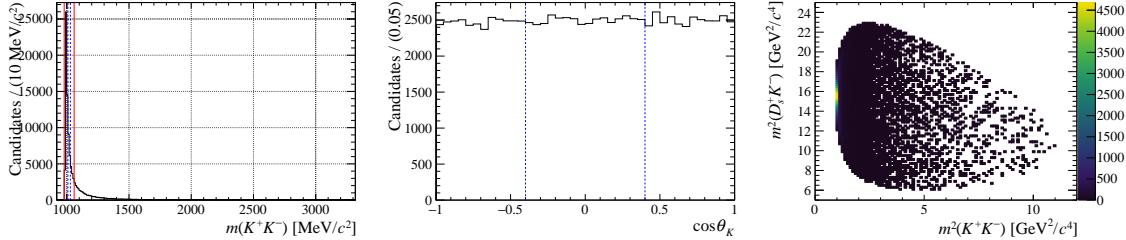


Figure 6.6: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos\theta_K$  for generated for the  $a_0^0(980)$  resonance.

### The $f_0^0(1370)$ resonance

The  $f_0^0(1370)$  resonance is a light unflavoured  $J^P = 0^+$  state with a mass in the range 1200–1500  $\text{MeV}/c^2$  and width in the range 200–500  $\text{MeV}/c^2$ . It has been observed to decay to the  $K\bar{K}$  final states. It is modelled with a Relativistic Breit-Wigner line shape and the relevant distributions are shown in Fig. 6.7.

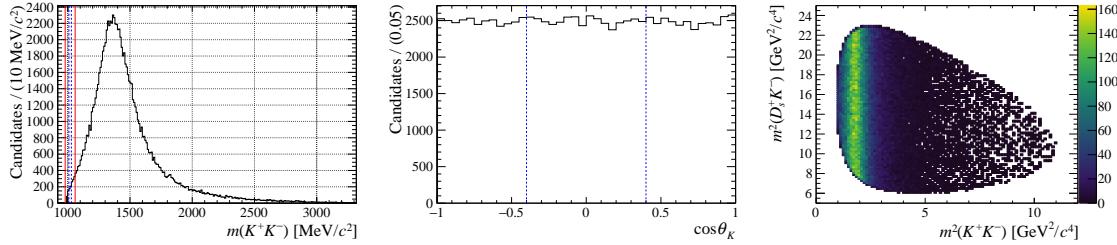


Figure 6.7: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos \theta_K$  for generated for the  $f_0^0(1370)$  resonance.

### The $f_2^0(1270)$ resonance

The  $f_2^0(1270)$  resonance is a  $J^P = 2^+$  state with mass  $1275.5 \pm 0.8$  MeV/c<sup>2</sup> and width  $186.7^{+2.2}_{-2.5}$  MeV/c<sup>2</sup> that has been observed to decay to  $K\bar{K}$  final states. This resonance is modelled with a Relativistic Breit-Wigner line shape as shown in Fig. 6.8.

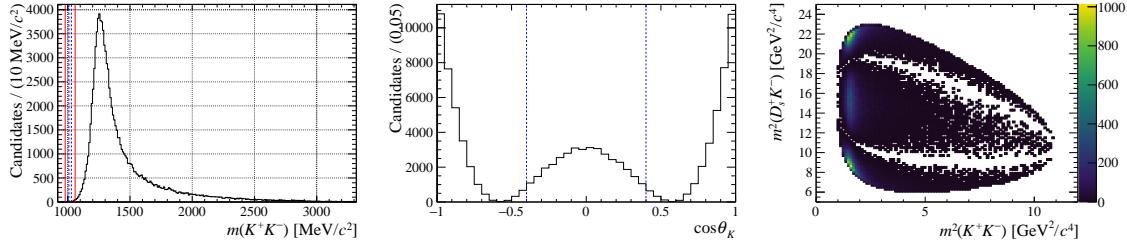


Figure 6.8: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos \theta_K$  for generated for the  $f_2^0(1270)$  resonance.

### The $a_2^0(1320)$ resonance

The  $a_2^0(1320)$  resonance is a  $J^P = 2^+$  state with a mass  $1318.1 \pm 0.7$  MeV/c<sup>2</sup> and width  $109.8 \pm 2.4$  MeV/c<sup>2</sup> (both measured in the  $K\bar{K}$  mode) observed decaying to the  $K\bar{K}$  final state. This resonance is modelled with a Relativistic Breit-Wigner line shape and shown in Fig. 6.9.

### Summary of models

The fraction of decays expected in each  $m(K^+K^-)$  and  $\cos \theta_K$  category for the different models considered are tabulated in Table 6.2. These are calculated by counting the numbers of entries in the corresponding ranges delineated by the vertical lines in Figs. 6.3-6.9. For reference the  $\phi(1020)$  fractions are included for both the fully

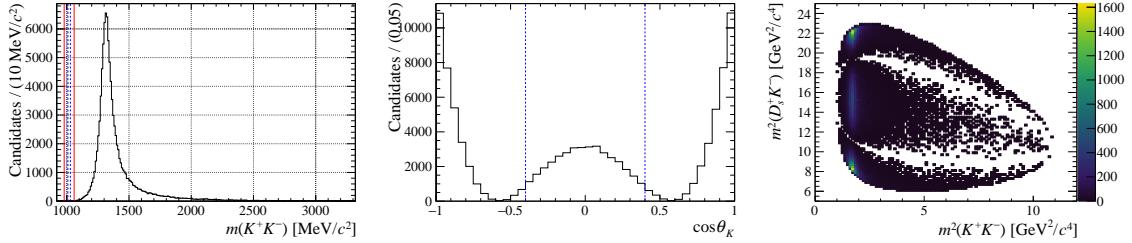


Figure 6.9: The distribution of  $m(K^+K^-)$  (left), Dalitz plot (middle) and the helicity angle  $\cos \theta_K$  for generated for the  $a_2^0(1320)$  resonance.

simulated decays and those generated with LAURA++. The maximum difference between these fractions is included as a source of systematic uncertainty in Sec. 6.7. For all of the models considered it is clear that the  $\phi$  resonance has significantly different fractions to all of the models considered, allowing the component to be distinguished.

Model	$ \Delta m  < 10 \text{ MeV}/c^2$		$10 <  \Delta m  < 40 \text{ MeV}/c^2$	
	$ \cos \theta_K  > 0.4$	$ \cos \theta_K  < 0.4$	$ \cos \theta_K  > 0.4$	$ \cos \theta_K  < 0.4$
$\phi(1020)$ LAURA++	83.5	5.7	10.2	0.7
$\phi(1020)$ Full	82.4	5.9	10.9	0.8
Non-resonant	16.3	11.1	45.4	27.2
$f_0^0(980)$	16.5	11.2	43.3	29.0
$a_0^0(980)$	12.7	8.7	47.0	31.5
$f_0^0(1370)$	16.1	9.6	45.0	29.3
$f_2^0(1270)$	8.6	6.5	59.1	25.8
$a_2^0(1320)$	9.7	9.7	51.6	29.0
Chosen fractions	$14.6 \pm 1.9$	$10.0 \pm 1.3$	$45.2 \pm 1.9$	$30.3 \pm 1.3$

Table 6.2: Fractions of decays expected in each  $m(K^+K^-)$  and  $\cos \theta_K$  category for the various resonance models considered in Sec. 6.1.2 ( $\Delta m = m(K^+K^-) - m(\phi)$ ).

In order to choose suitable fractions for the  $B^+ \rightarrow D_s^+ K^+ K^-$  fit component a number of points are considered;

- Very few events are observed above 2000 MeV/c<sup>2</sup> in the background-subtracted  $m(K^+K^-)$  distribution, therefore the non-resonant model is neglected.
- No significant peaking structure is observed in the  $m(K^+K^-)$  spectrum so on-shell resonances are neglected

- The helicity distribution shows no distinctive structure so spin zero states are favoured.
- As this is not a full amplitude analysis no attempt is made to include the effects of interference, either between the remaining off-shell resonances or between these and any possible  $B^+ \rightarrow D_s^+ \phi$  decays.

These considerations leave the  $f_0^0(980)$  and  $a_0^0(980)$  resonances. The fractions of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays that have been used in the fit model are fixed to the average of these two, listed in the final row of Table 6.2. Uncertainties are assigned that correspond to half the difference of the two values. These uncertainties are propagated to the  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  branching fraction in Sec. 6.7.

## 6.2 Fit components

The yields of  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays are extracted from the invariant mass distributions of the data sets by representing each component by probability density functions. The components are broadly very similar to those considered in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays detailed in Sec. 5.2. There are a number of necessary differences:

- The  $B^+$  invariant mass range considered for both the signal and normalisation channel is expanded to 4900–5900 MeV/ $c^2$ . This allows a more stable determination of the various backgrounds contributing in the vicinity of the signal decays. In particular it stabilises the fraction of decays assigned to the combinatorial and partially reconstructed backgrounds at low  $B^+$  invariant mass.
- More components are included in the model. This include both the signal mode,  $B^+ \rightarrow D_s^+ \phi$ , and a closely related additional background mode  $B^+ \rightarrow D_s^{*+} \phi$ .

### 6.2.1 Signal and normalisation decays

The invariant mass distributions of  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays are parametrised using the same DCB function as in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays (Sec. 5.2.1)

$$\text{DCB}(m|\mu, \sigma_1, \sigma_2, n, \alpha) = f_\sigma \times \text{CB}(m|\mu, \sigma_1, n, \alpha) + (1 - f_\sigma) \times \text{CB}(m|\mu, \sigma_2, n, \alpha), \quad (6.3)$$

where the CB function is defined as follows

$$\text{CB}(m|\mu, \sigma, n, \alpha) = \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2} & \text{if } \left(\frac{m-\mu}{\sigma}\right) < -|\alpha| \\ \frac{\left(\frac{n}{|\alpha|}\right)^n \times e^{-\frac{1}{2}|\alpha|^2}}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{m-\mu}{\sigma}\right)^n} & \text{otherwise.} \end{cases} \quad (6.4)$$

Again,  $\mu$ ,  $\sigma$ ,  $n$ ,  $\alpha$  and  $f_\sigma$  are adjustable parameters and  $m$  is the  $B$  meson invariant mass observable. The tail parameter  $\alpha$  is fixed to values determined from maximum likelihood fits to simulated candidates for the signal and normalisation decays. The parameter  $n$  is fixed to unity in the fits to both simulation and data to increase the stability of the tails. The two CB function are allowed have different widths,  $\sigma_1$  and  $\sigma_2$ , but the ratio  $\sigma_1/\sigma_2$  is fixed from the fits to simulations, as is  $f_\sigma$  that determines the fractional contribution of the narrower CB function ( $\sigma_1 < \sigma_2$ ). The values determined for each of these fixed parameters are tabulated in Table 6.3, along with the uncertainty obtained from the fits.

An extra constraint is added to the signal and normalisation DCB functions with respect to the configuration used in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays. As the number of signal candidates is likely to be small, the relative width of the narrower signal and normalisation CB functions,  $\sigma_1(D_s^+ \phi)/\sigma_1(D_s^+ \bar{D}^0)$ , is also fixed to values obtained from the fits to simulations. All fixed parameters are determined separately for the different  $D_s^+$  decay modes, and the results of the fits to simulated decays are shown in Fig. 6.10.

Parameter	Value		
	$D_s^+ \rightarrow K^+ K^- \pi^+$	$D_s^+ \rightarrow K^+ \pi^- \pi^+$	$D_s^+ \rightarrow \pi^+ \pi^- \pi^+$
$B^+ \rightarrow D_s^+ \phi$			
$\sigma_1/\sigma_2$	$0.49 \pm 0.01$	$0.47 \pm 0.01$	$0.46 \pm 0.01$
$f_\sigma$	$0.80 \pm 0.01$	$0.84 \pm 0.01$	$0.81 \pm 0.01$
$\alpha$	$2.76 \pm 0.07$	$3.06 \pm 0.16$	$3.71 \pm 0.23$
$n$	$1 \pm 0$	$1 \pm 0$	$1 \pm 0$
$B^+ \rightarrow D_s^+ \bar{D}^0$			
$\sigma_1/\sigma_2$	$0.43 \pm 0.01$	$0.42 \pm 0.01$	$0.40 \pm 0.01$
$f_\sigma$	$0.88 \pm 0.01$	$0.88 \pm 0.01$	$0.88 \pm 0.01$
$\alpha$	$2.91 \pm 0.06$	$3.36 \pm 0.26$	$3.53 \pm 0.25$
$n$	$1 \pm 0$	$1 \pm 0$	$1 \pm 0$
$\sigma_1(D_s^+ \phi)/\sigma_1(D_s^+ \bar{D}^0)$	$1.27 \pm 0.02$	$1.31 \pm 0.02$	$1.26 \pm 0.02$

Table 6.3: Fixed values obtained in fits to MC used in the model for the signal pdf.

### 6.2.2 Partially reconstructed backgrounds

The accurate parametrisation of partially reconstructed backgrounds is particularly important in the search for  $B^+ \rightarrow D_s^+ \phi$  decays as many different processes contribute to the low invariant mass range of the  $m(D_s^+ \phi)$  spectrum. These processes involve decays of  $B_s^0$ ,  $B^0$  or  $B^+$  mesons in which the five final state tracks reconstructed in the search for  $B^+ \rightarrow D_s^+ \phi$  decays are only a subset of the background modes final state. Typically, processes in which a low momentum pion or photon has not been reconstructed are found closest in mass to the signal decays. Decays of  $B_s^0$  mesons are particularly dominant as the  $B_s^0$  meson has a larger mass than the  $B^+$  meson.

#### Backgrounds to the normalisation channel

The modes  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  and  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  can both contribute as partially reconstructed backgrounds to the  $B^+ \rightarrow D_s^+ \bar{D}^0$  normalisation mode. These are parametrised using the same PDFs as in the search for  $B^+ \rightarrow D_s K^+ K^-$  decays, detailed in Sec. 5.2.2.

The invariant mass fit range is wider than in the fit to  $B^+ \rightarrow D_s^+ K^+ K^-$  candidates, therefore an extra contribution is included in the fit model to account for partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^{*0}$  decays at lower invariant masses.

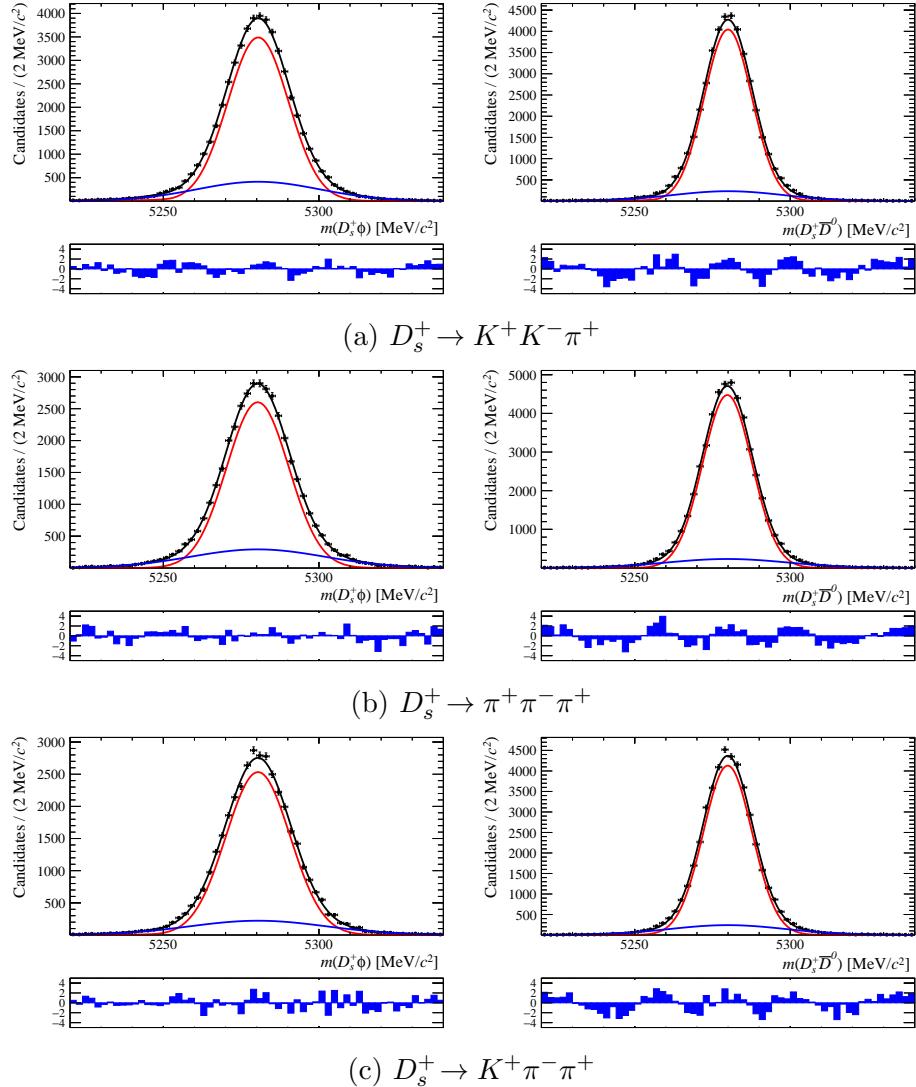


Figure 6.10: Invariant mass fits to simulated signal (left) and normalisation (right) decays. The results of maximum likelihood fits using the signal PDFs are overlaid, with the total function in black and the two contributing CB shapes in red and blue.

$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\pi^0])\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\pi^0])$ : these components are modelled by a parabola convolved with a resolution Gaussian. The parabola has a minimum in the centre and doesn't extend beyond endpoints  $a$  and  $b$

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b \left( \mu - \frac{a+b}{2} \right)^2 \left( \frac{1-\xi}{b-a} \mu + \frac{b\xi - a}{b-a} \right) e^{-\frac{-(\mu-(m-\delta))^2}{2\sigma^2}} d\mu. \quad (6.5)$$

These components are shown by the black lines in Fig. 6.11.

$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\gamma])\bar{D}^0$  and  $B^+ \rightarrow D_s^+(\bar{D}^{*0} \rightarrow \bar{D}^0[\gamma])$ : these components are

modelled by a parabola convolved with a resolution Gaussian. The parabola has a maximum in the centre and doesn't extend beyond endpoints  $a$  and  $b$

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b -(\mu - a)(\mu - b) \left( \frac{1 - \xi}{b - a} \mu + \frac{b\xi - a}{b - a} \right) e^{-\frac{-(\mu - (m - \delta))^2}{2\sigma^2}} d\mu. \quad (6.6)$$

These components are shown by the blue lines in Fig. 6.11.

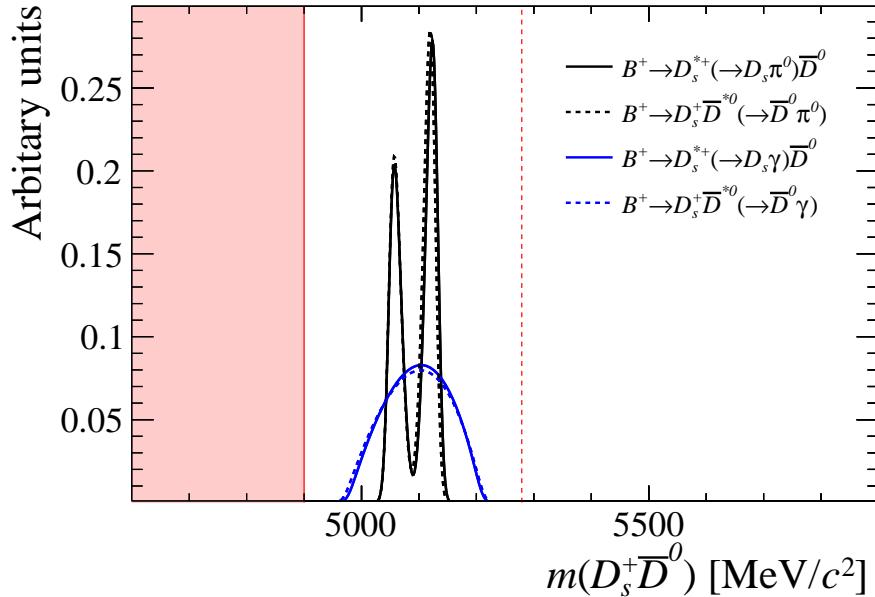


Figure 6.11: Partially reconstructed  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  and  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  decay parametrisations. The  $B^+$  mass is represented by a vertical dashed red line. Candidates in the area below  $4900 \text{ MeV}/c^2$  are not included in the fit.

$B^+ \rightarrow D_s^{*+} \bar{D}^{*0}$ : the lower invariant mass range used in this search necessitates including a PDF for  $B^+ \rightarrow D_s^{*+} \bar{D}^{*0}$  decays in which two soft particles have been missed, one from each of the excited  $D$  meson decays. It is possible for either a  $\pi^0$  or  $\gamma$  to be not reconstructed in the decays of both excited  $D$  mesons. Additionally, as this process involves a pseseudo-scalar meson decaying to two vector mesons, there should be two distinguishable helicity combinations for each process. This leads to a total of eight PDFs necessary to fully parametrise this contribution. Instead, however, this component is parametrised using a single function of the form given in Eq. 6.6. The endpoints  $a$  and  $b$  are estimated

by combining the effects of missing two neutral particles. The resulting distribution is shown in Fig. 6.12. The choice of PDF for this component is found to have negligible effect on the determination of the  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  branching fraction as detailed in Sec 6.7.

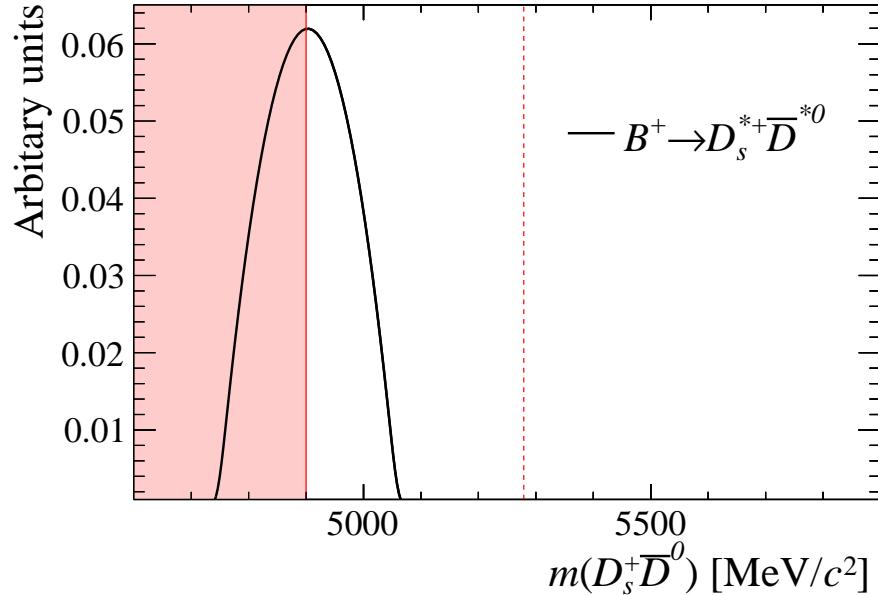


Figure 6.12: Partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^{*0}$  decay parametrisation. The PDF extends below the lower fit range at  $4900 \text{ MeV}/c^2$ .

### Backgrounds to the signal channel

$B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\gamma])\phi$  and  $B^+ \rightarrow (D_s^{*+} \rightarrow D_s^+[\pi^0])\phi$ : the decays of  $B^+$  mesons to an excited  $D_s^+$  meson and a  $\phi$  meson could contribute to the  $m(D_s^+ \phi)$  spectrum at low invariant masses when either a  $\pi^0$  or  $\gamma$  is missed from the excited meson decay. The resulting invariant mass distribution depends on the mass and spin of the non-reconstructed particle, as well as the helicity state of the  $D_s^{*+}$  meson. This background involves the decay of a pseudo-scalar meson to two vector mesons, hence, as a result of angular momentum conservation, there are three helicity states of the  $D_s^{*+}$  meson to consider. These are labelled *001*, *010* and *100*. The two transversely polarised states, *100* and *001*, have identical invariant mass distributions and are therefore referred collectively as *101*. This leads to

a total of four contributions to consider for  $B^+ \rightarrow D_s^{*+} \phi$  decays. These can be parametrised by parabolas convolved with resolution Gaussians in a similar way to the partially reconstructed backgrounds to the normalisation channel. These each share the same functional form

$$f(m|a, b, \sigma, \xi, \delta) = \int_a^b g(\mu, a, b) \left( \frac{1-\xi}{b-a} \mu + \frac{b\xi - a}{b-a} \right) e^{-\frac{-(\mu-(m-\delta))^2}{2\sigma^2}} d\mu. \quad (6.7)$$

where  $g(\mu, a, b)$  represents the parabola for each of the four components, listed in Table 6.4. The resulting invariant mass distributions are shown in Fig.6.7.

This decay is unobserved, therefore the relative contribution to this decay in nature from the  $101$  and  $010$  helicity states is not known. The total PDF for this contribution is made by weighting the  $\pi^0$  and  $\gamma$  contributions by their corresponding branching fractions, and by assuming that the  $101$  and  $010$  helicity states contribute with equal magnitudes. This assumption is varied and included as a source of systematic uncertainty in the  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  determination.

Missed particle	Helicity	$g(\mu, a, b)$	$a$ (MeV/c <sup>2</sup> )	$b$ (MeV/c <sup>2</sup> )
$\pi^0$	010	$(\mu - \frac{a+b}{2})^2$	5026.8	5124.8
$\pi^0$	101	$-(\mu - a)(\mu - b)$	5026.8	5124.8
$\gamma$	010	$-(\mu - a)(\mu - b)$	4936.4	5220.6
$\gamma$	101	$(\mu - \frac{a+b}{2})^2 + (\frac{a+b}{2})^2$	4936.4	5220.6

Table 6.4: The parabolas contributing to the partially reconstructed  $B^+ \rightarrow D_s^{*+} \phi$  decay parametrisation as defined in Eq. 6.7.

$\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0}$ : this decay can form a background to  $B^+ \rightarrow D_s^+ \phi$  decays when the soft pion from the  $K^{*0} \rightarrow K^+ \pi^-$  decay is not reconstructed. The lower bound of the fit range is wide enough that a significant fraction of these decays are retained in the fitted data set. The  $K^- K^{*0}$  is modelled as originating from the  $a_1(1260)$  resonance. This resonance has a width of  $250 - 600$  MeV [28], allowing it to decay to  $K^- K^{*0}$  even though its pole mass is below the  $K^- K^{*0}$  threshold. A PDF for this component is determined by reconstructing simulated  $\bar{B}_s^0 \rightarrow$

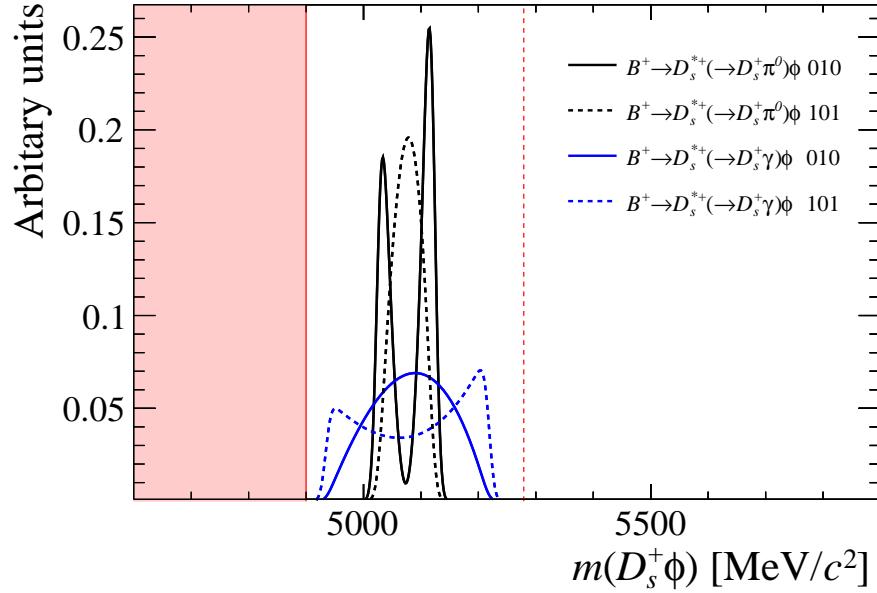


Figure 6.13: Partially reconstructed  $D_s^+ \phi$  shapes.

$D_s^+ K^- K^{*0}$  decays through the identical reconstruction and selection steps as the signal. The `ROOFIT` class `RooKeysPDF` is used to create a kernel estimation of the partially reconstructed  $B^+$  mass distribution for the candidates passing the selection. This is shown in Fig. 6.15. The fraction of these decays expected in each of the four  $m(K^+ K^-)$  and  $\cos \theta_K$  categories is determined from the simulation samples.

$\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$ : this decay similarly forms a background to the signal mode when both a low momentum neutral particle ( $\pi^0$  or  $\gamma$ ) is not reconstructed in the decay of the  $D_s^{*+}$  meson, in addition to the low momentum pion from the  $K^{*0}$  decay. The PDF is similarly determined using a kernel estimation from simulated decays passing the full selection as shown in Fig. 6.15. The fraction of candidates expected in each of the  $m(K^+ K^-)$  and  $\cos \theta_K$  categories are determined from these simulation samples.

$\bar{B}_s^0 \rightarrow D_s^+ D_s^-$ ,  $\bar{B}^0 \rightarrow D_s^+ D^-$  and  $\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$ : the decays of neutral  $B$  mesons to two charged  $D$  mesons can form a background to the signal decay if a low

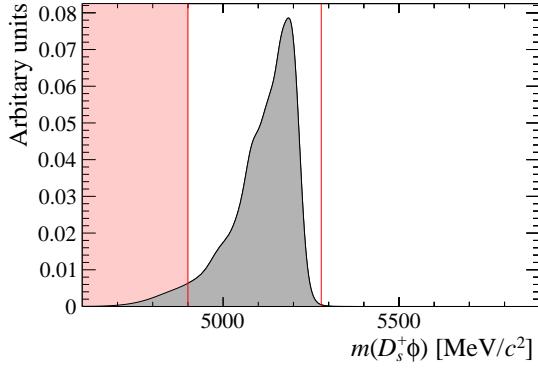


Figure 6.14: Partially reconstructed mass PDFs determined from samples of  $\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0}$  simulations processed with the same reconstruction and selection as the signal decays. The  $B^+$  meson mass is indicated by a vertical red line. The area below 4900 MeV/ $c^2$  is not included in the fit range, but included for reference. The PDF colours follow the same convention used in the final fit plots shown in Fig. 6.21.

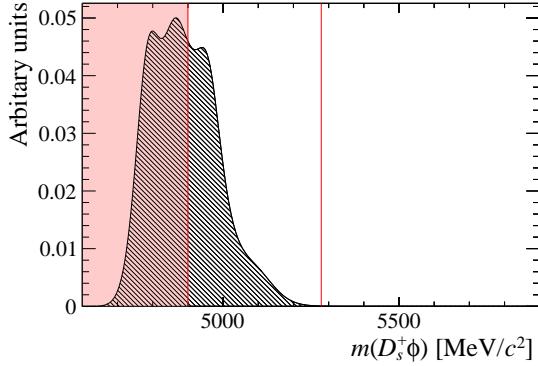


Figure 6.15: Partially reconstructed mass PDFs determined from samples of  $\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}$  simulations processed with the same reconstruction and selection as the signal decays. The  $B^+$  meson mass is indicated by a vertical red line. The area below 4900 MeV/ $c^2$  is not included in the fit range, but included for reference. The PDF colours follow the same convention used in the final fit plots shown in Fig. 6.21.

momentum pion is not reconstructed in one of the charm meson decays. In the case of  $\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$  decays an additional neutral particle is not reconstructed in the decay of the  $D_s^{*+}$  meson. The PDFs for these decays are determined from simulated decays that have been processed with the same reconstruction and selection steps as the signal. For  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  and  $\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$  decays, the PDF is created using a kernel estimation using the ROOFIT RooKeysPDF implementation, shown in Fig.6.16. Due to the similarities between  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$

and  $\bar{B}^0 \rightarrow D_s^+ D^-$  decays, the PDF for the latter is created by using the kernel estimation for  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  and shifting the shape down in mass by  $40 \text{ MeV}/c^2$  to account for the kinematic differences.

The branching fractions for these three decays are measured to be  $\mathcal{B}(\bar{B}^0 \rightarrow D_s^+ D^-) = (7.2 \pm 0.8) \times 10^{-3}$ ,  $\mathcal{B}(\bar{B}_s^0 \rightarrow D_s^+ D_s^-) = (4.4 \pm 0.5) \times 10^{-3}$  and  $\mathcal{B}(\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-) = (1.37 \pm 0.16)\%$ . Therefore the relative contributions for these three processes are fixed using these branching fractions, estimates of the relative efficiencies for each mode and the production fraction of  $\bar{B}_s^0$  mesons relative to  $\bar{B}^0$  mesons,  $f_s/f_d$ . This helps to add stability to the fit.

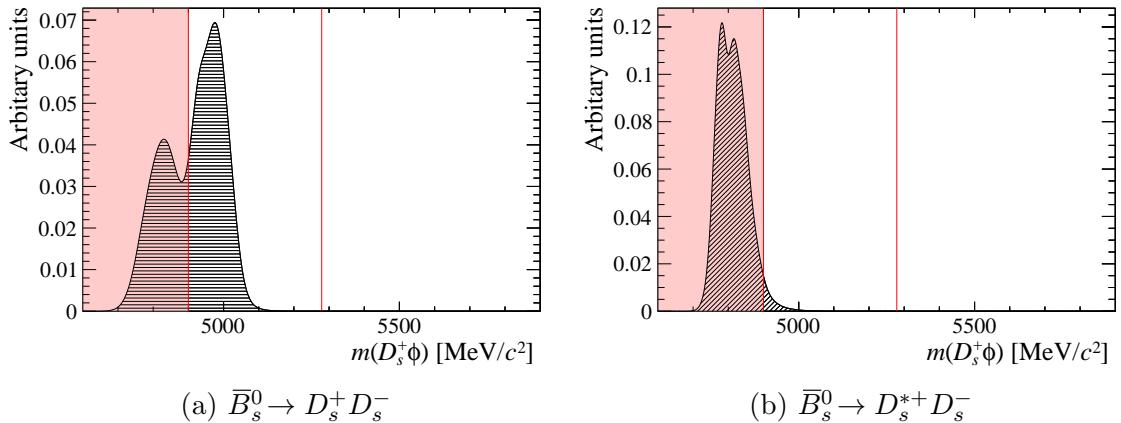


Figure 6.16: Partially reconstructed mass PDFs determined from samples of  $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$  and  $\bar{B}_s^0 \rightarrow D_s^{*+} D_s^-$  simulations processed with the same reconstruction and selection as the signal decays. The  $B^+$  meson mass is indicated by a vertical red line. The area below  $4900 \text{ MeV}/c^2$  is not included in the fit range, but included for reference. The PDF colours follow the same convention used in the final fit plots shown in Fig. 6.21.

To account for difference between the simulation and data samples all PDFs determined using the kernel estimation method are convolved with an additional Gaussian. The mean position of this Gaussian is given by  $\delta$ , the same offset used for the analytically described partially decays. The width of the Gaussian is increased to account for the difference in resolution between simulation and data.

### 6.2.3 Combinatorial backgrounds

Combinations of unrelated tracks form a background to the signal decays and extend across the entire fitted  $B^+$  meson mass range. This component is parametrised with a decaying exponential function

$$f(m|c) = e^{-m \times c}, \quad (6.8)$$

where the parameter  $c$  controls the slope and  $m$  is the observable  $B^+$  meson invariant mass. The combinatorial background is the dominant background under the signal decays, therefore it is important to accurately extrapolate distribution from the high and low mass ranges to the signal range.

The yields of combinatorial background is allowed to vary freely in each of the separate simultaneous fit categories. To improve the stability of the fit, the single slope parameter,  $c$ , is shared between all of the categories. In categories with lower statistics, for example the  $D_s^+ \rightarrow K^+ \pi^- \pi^+$  decay mode, it would be possible for the partially reconstructed backgrounds to be incorrectly assigned to the combinatorial component. This can happen if the exponential slope for that mode gets to large, causing it to ‘kick up’ at low invariant masses. This may bias the signal yield in this mode. Fixing the slope parameter to be the same between the different  $D_s^+$  modes and between the signal and normalisation channel allows the lower statistics modes to benefit from those with higher statistics.

## 6.3 Free and constrained parameters

The fit to  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays contains a total of 54 free parameters. These are broadly divided into four different categories detailed here.

### 6.3.1 Parameter of interest

The parameter of interest (POI) is the branching fraction for  $B^+ \rightarrow D_s^+ \phi$  decays. This is determined directly in the fit to the signal and normalisation decays. The

same single branching fraction is used for all of the  $D_s^+$  decays modes. The yield of signal candidates is calculated using the equation

$$N(B^+ \rightarrow D_s^+ \phi) = k_{D_s^+} \times \mathcal{B}(B^+ \rightarrow D_s^+ \phi) \times N(B^+ \rightarrow D_s^+ \bar{D}^0), \quad (6.9)$$

where  $k_{D_s^+}$  is a constant defined for each  $D_s^+$  decay mode

$$k_{D_s^+} = \frac{\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)}{\epsilon(B^+ \rightarrow D_s^+ \phi)} \times \frac{\mathcal{B}(\phi \rightarrow K^+ K^-)}{\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0) \mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)}. \quad (6.10)$$

Here, the efficiencies  $\epsilon$  are calculated for the specific  $D_s^+$  decay mode in question and external measurements are used for the additional branching fractions. The yield  $N(B^+ \rightarrow D_s^+ \phi)$  is the total signal yield in all four  $m(K^+ K^-)$  and  $\cos \theta_K$  categories. The yields in each category are constrained to be in the ratios previously listed in Table 6.1, *i.e.* the yield in the  $|m(K^+ K^-)| < 10 \text{ MeV}/c^2$  and  $|\cos \theta_K| > 0.4$  category is  $0.82 \times N(B^+ \rightarrow D_s^+ \phi)$ .

### 6.3.2 Shape parameters

As detailed in Sec 6.2, the fit included eight parameters governing the shapes of various PDFs.

- The combinatorial background PDF is controlled by a single slope parameter  $c$  for all categories.
- The mean  $B^+$  mass for the signal and normalisation mode is free to vary in the fit. The same value is used for all  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  PDFs.
- The invariant mass offset  $\delta$  shared between all partially reconstructed background PDFs is allowed to freely vary in the fit. The same value is used for the signal and normalisation modes.
- The relative heights of the two peaks in partially reconstructed  $B^+ \rightarrow D_s^{*+} \phi$ ,  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  decays,  $\xi$ , is allowed to vary. The same value is used for all modes.
- The smaller CB width parameter,  $\sigma_1$ , is allowed to vary for each of the  $D_s^+$  decay modes, leading to four more free parameters.

### 6.3.3 Yields

The fit model contains a total of 36 free parameters that are yields.

- The total yield of normalisation decays in each  $D_s^+$  decay mode is left floating in the fit. These variables are the sum of the yields in the two helicity categories,  $|\cos \theta_K| < 0.4$  and  $|\cos \theta_K| > 0.4$ . This results in four free parameters.
- The yield of certain partially reconstructed decays in the normalisation channel is left floating for each  $D_s^+$  decay mode independently. This yield is defined to be  $N(B^+ \rightarrow D_s^+ \bar{D}^{*0}) + N(B^+ \rightarrow D_s^{*+} \bar{D}^0)$ . The yield of the doubly excited  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  is not included in this total. This results in four free parameters.
- The yield of some partially reconstructed decays in the signal mode is left free in the fit. This is defined to be  $N(D_s^{*+} \phi) + N(D_s^+ K^- K^{*0}) + N(D_s^{*+} K^- K^{*0})$ . This leads to four free parameters.
- All yields of combinatorial backgrounds are left free in the fit. This results in 24 free parameters.

### 6.3.4 Fractions

The majority of the partially reconstructed backgrounds do not have unconstrained yields in each of the 24 simultaneous fit categories. Instead, a single partially reconstructed background yield per  $D_s^+$  decay mode is left free in the fit. The relative contributions of other partially backgrounds are determined by floating fraction parameters multiplied by this yield. This allows the different background contributions to vary relative to one another, but keeps the ratios of the relative contributions the same across the different  $D_s^+$  decay modes.

There are four free parameters controlling the backgrounds contributions in the signal mode:

1. Ratio of yields of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays to  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays

$$\frac{N(B^+ \rightarrow D_s^+ K^+ K^-)}{N(B^+ \rightarrow D_s^+ \bar{D}^0)}. \quad (6.11)$$

2. The fraction of  $\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0}$  decays to the total of  $\bar{B}_s^0 \rightarrow D_s^{(*)+} K^- K^{*0}$  decays

$$\frac{N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})}{N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}) + N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})}. \quad (6.12)$$

3. The fraction of  $B^+ \rightarrow D_s^{*+} \phi$  decays in the total of these and  $\bar{B}_s^0 \rightarrow D_s^{(*)+} K^- K^{*0}$  decays

$$\frac{N(B^+ \rightarrow D_s^{*+} \phi)}{N(B^+ \rightarrow D_s^{*+} \phi) + N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}) + N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})} \quad (6.13)$$

4. The ratio of yields of the partially reconstructed  $B_s^0 \rightarrow D_s^+ D_s^-$ ,  $B_s^0 \rightarrow D_s^+ D_s^{*-}$  and  $B^0 \rightarrow D_s^+ D^-$  decays to  $B^+ \rightarrow D_s^{*+} \phi$  and  $\bar{B}_s^0 \rightarrow D_s^{(*)+} K^- K^{*0}$  decays

$$\frac{N(B_s^0 \rightarrow D_s^+ D_s^-) + N(B_s^0 \rightarrow D_s^+ D_s^{*-}) + N(B^0 \rightarrow D_s^+ D^-)}{N(B^+ \rightarrow D_s^{*+} \phi) + N(\bar{B}_s^0 \rightarrow D_s^{*+} K^- K^{*0}) + N(\bar{B}_s^0 \rightarrow D_s^+ K^- K^{*0})} \quad (6.14)$$

Another five free parameters control the fractions of events in different categories for the normalisation mode:

1. Fraction of partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  decays in the total of  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  decays in the  $|\cos \theta_K| > 0.4$  category

$$\left( \frac{N(B^+ \rightarrow D_s^{*+} \bar{D}^0)}{N(B^+ \rightarrow D_s^{*+} \bar{D}^0) + N(B^+ \rightarrow D_s^+ \bar{D}^{*0})} \right)_{|\cos \theta_K| > 0.4} \quad (6.15)$$

2. Fraction of partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  decays in the total of  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  decays in the alternative  $|\cos \theta_K| < 0.4$  category

$$\left( \frac{N(B^+ \rightarrow D_s^{*+} \bar{D}^0)}{N(B^+ \rightarrow D_s^{*+} \bar{D}^0) + N(B^+ \rightarrow D_s^+ \bar{D}^{*0})} \right)_{|\cos \theta_K| < 0.4} \quad (6.16)$$

3. Fraction of partially reconstructed backgrounds in the  $|\cos \theta_K| > 0.4$  category

$$\frac{(N(B^+ \rightarrow D_s^{*+} \bar{D}^0) + N(B^+ \rightarrow D_s^+ \bar{D}^{*0}))_{|\cos \theta_K| > 0.4}}{(N(B^+ \rightarrow D_s^{*+} \bar{D}^0) + N(B^+ \rightarrow D_s^+ \bar{D}^{*0}))_{\text{Total}}} \quad (6.17)$$

4. The ratio of  $B^+ \rightarrow D_s^{*+} \bar{D}^{*0}$  decays to the total of  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  decays

$$\frac{N(B^+ \rightarrow D_s^{*+} \bar{D}^{*0})}{N(B^+ \rightarrow D_s^{*+} \bar{D}^0) + N(B^+ \rightarrow D_s^+ \bar{D}^{*0})} \quad (6.18)$$

5. Fraction of fully reconstructed normalisation decays in the  $|\cos \theta_K| > 0.4$  category

$$\frac{N(B^+ \rightarrow D_s^+ \bar{D}^0)_{|\cos \theta_K| > 0.4}}{N(B^+ \rightarrow D_s^+ \bar{D}^0)_{\text{Total}}}.$$
 (6.19)

This parameter is included as a cross check. The normalisation decays have a flat distribution in  $\cos \theta_K$  so the parameter would be expected to be around 0.6.

## 6.4 Fit validation

The simultaneous fitting framework is validated by generating pseudo-experiments (also referred to as *toys*). The total fit model PDF is randomly sampled to create a simulation sample with the same number of candidates as the nominal fit. These are then fitted using the same fit model, determining the best estimate and uncertainty of each parameter. The parameter values used to generate the pseudo-experiments are chosen to be the final parameter values as determined in a fit to data. This is known as the plug-in method [120]. The fitted value and associated error is used to determine the pull of each parameter of interest. As the errors are determined asymmetrically using MINOS the is defined conditionally to incorporate the appropriate error

$$g_{\text{pull}} = \begin{cases} \frac{x_{\text{gen}} - x_{\text{fit}}}{\sigma_+}, & \text{if } x_{\text{fit}} < x_{\text{gen}} \\ \frac{x_{\text{fit}} - x_{\text{gen}}}{\sigma_-}, & \text{otherwise,} \end{cases}$$
 (6.20)

where  $x_{\text{fit}}$  and  $x_{\text{gen}}$  are the fitted and generated values of the variable, and  $\sigma_+$  and  $\sigma_-$  are the high and low asymmetric errors.

The distributions of the values, errors and pulls for the yields of the normalisation decay in each of the  $D_s^+$  decays are shown in Fig. 6.17. The mean and widths are determined using simple fits to the pull distributions. The results and PDFs for these fits are overlaid on the distributions. The normalisation yield means and widths are found to be within  $2\sigma$  of zero and one respectively.

Similarly, the distributions of the signal yield, error and pull for each of the different  $D_s^+$  decay modes are shown in Fig. 6.19. The pull means and widths are all within  $2\sigma$  of zero and one respectively.

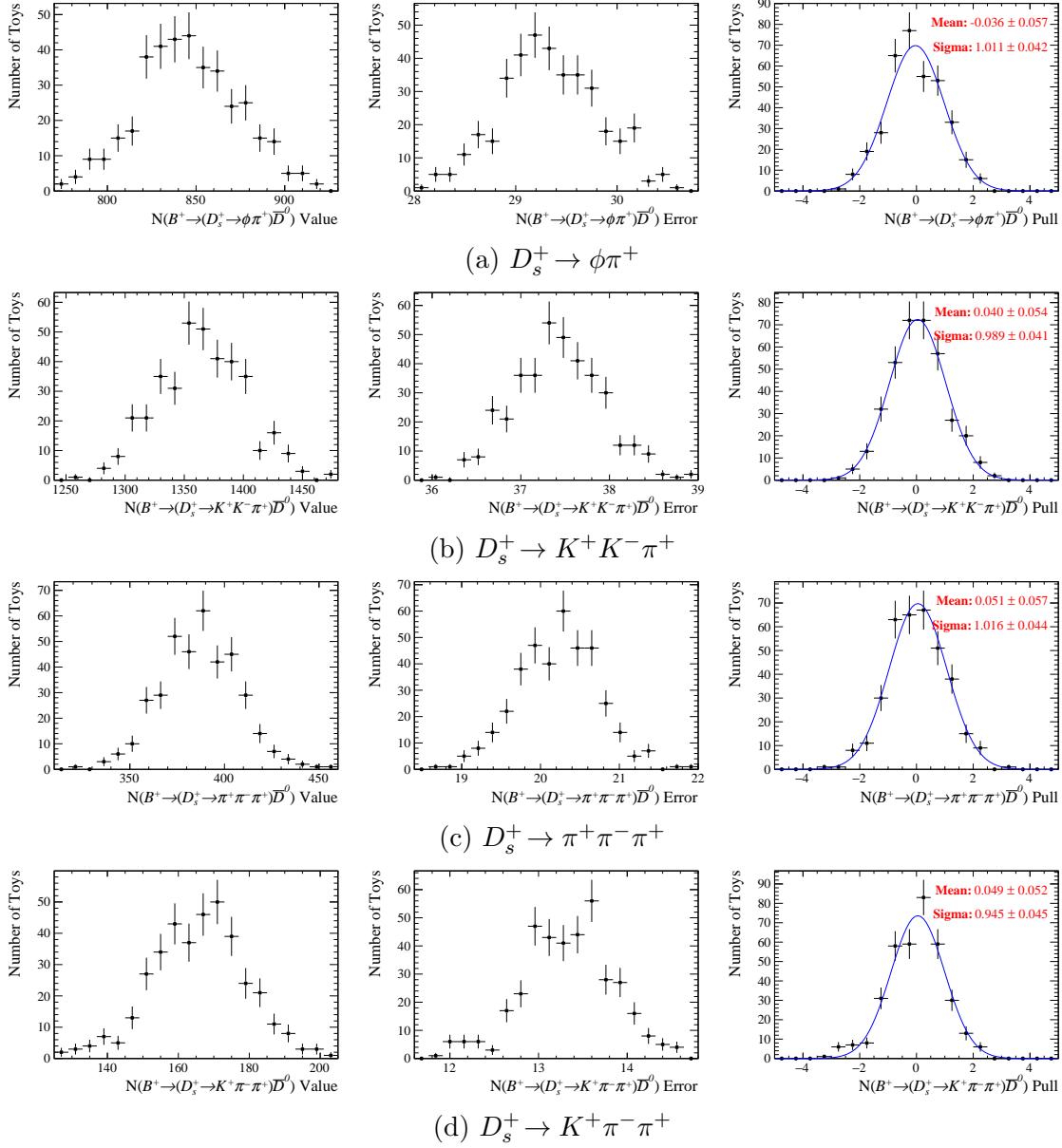


Figure 6.17: The yield, error and pull distributions for the normalisation channel.

These distributions were generated using pseudo-experiments in which the yields for each  $D_s^+$  mode were free variables rather than being determined using a single free branching fraction parameter.

The model is also studied using a single branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  that determines the yields in each signal mode. This parameter is calculated directly in the fit so it can be assessed for any possible bias. The distribution of the measured values, uncertainties and pulls are shown in Fig. 6.19.

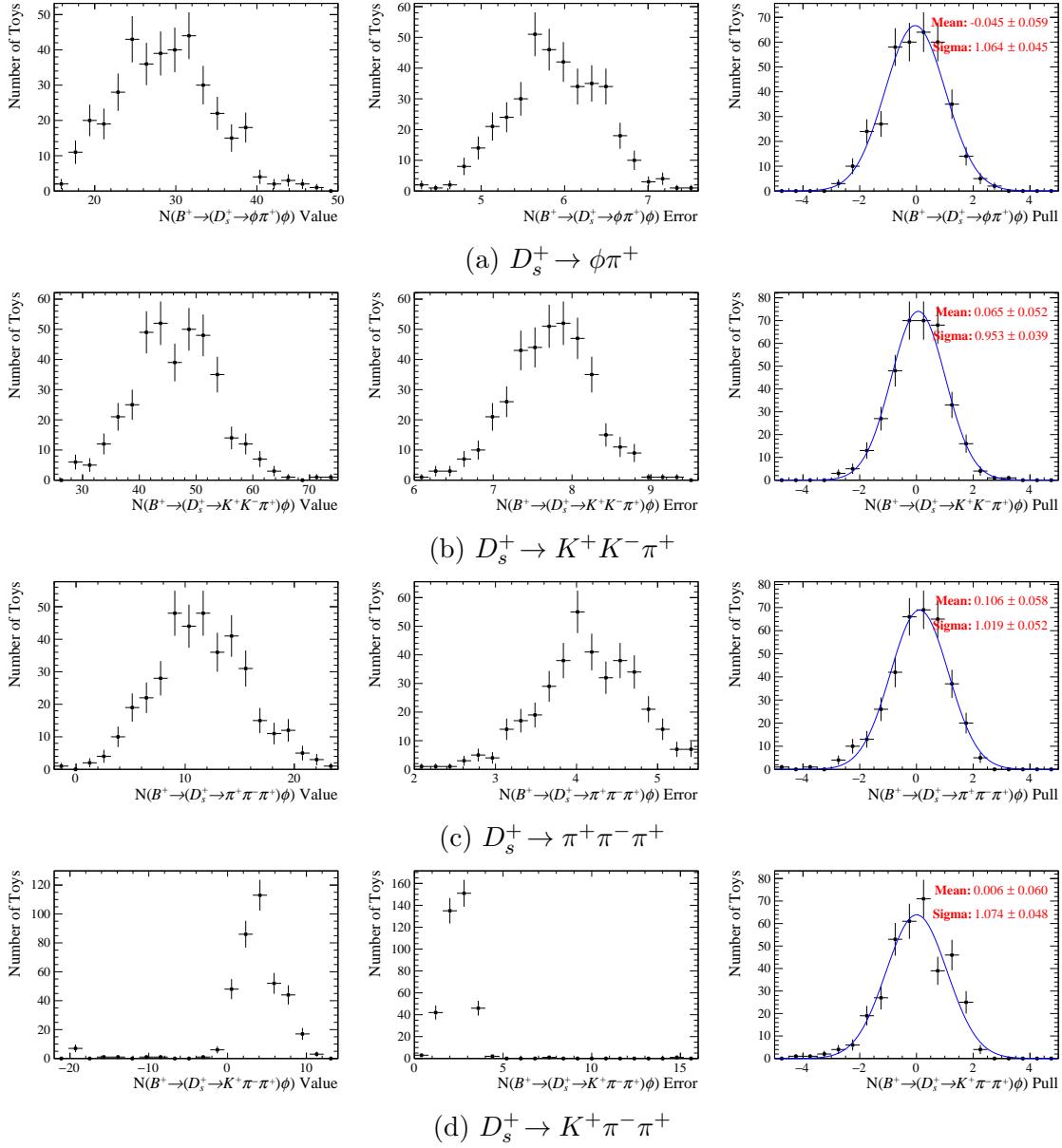


Figure 6.18: The yield, error and pull distributions for the signal channel.

The pull mean and width are both within  $2\sigma$  of zero and one respectively. As no significant biases are observed in the pulls of any of the yields or branching fraction no corrections are applied to the values determined fit.

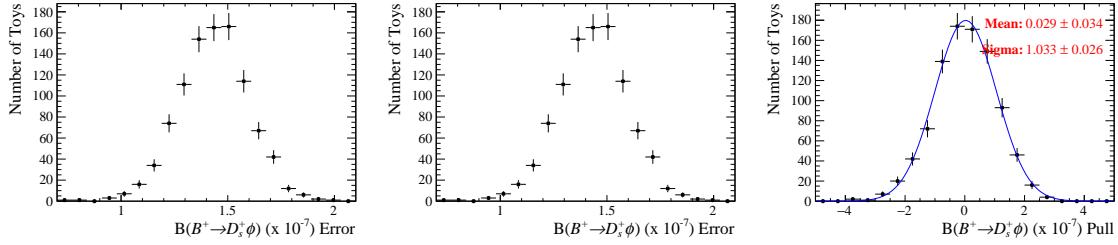


Figure 6.19: The branching fraction, error and pull distributions for the full simultaneous fit. The branching fraction is generated assuming the value obtained in the fit to data.

## 6.5 Normalisation and signal fits

The result of the simultaneous fit to the signal and normalisation decays is shown in Figs. 6.20 and 6.21. These figures show the distribution of candidates passing all selection requirements along with the total PDF model resulting from the minimisation of the log-likelihood. The four  $D_s^+$  decay mode categories have been merged into a single distribution. The fits in each different  $D_s$  decay mode category can be found in Appendix A.

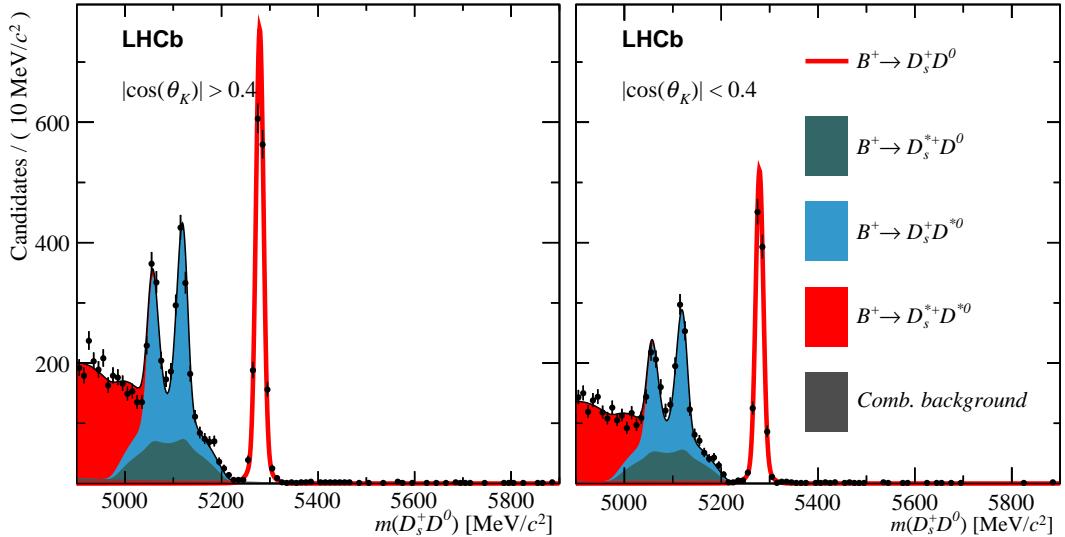


Figure 6.20: Invariant mass fits to  $B^+ \rightarrow D_s^+ \bar{D}^0$  candidates

The high purity of the normalisation mode reconstruction can be seen in Fig. 6.20, as the contribution from the combinatorial background shape is very small. Additionally, the double-peaked structure of the partially reconstructed  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and

$B^+ \rightarrow D_s^+ \bar{D}^{*0}$  decays is clearly visible.

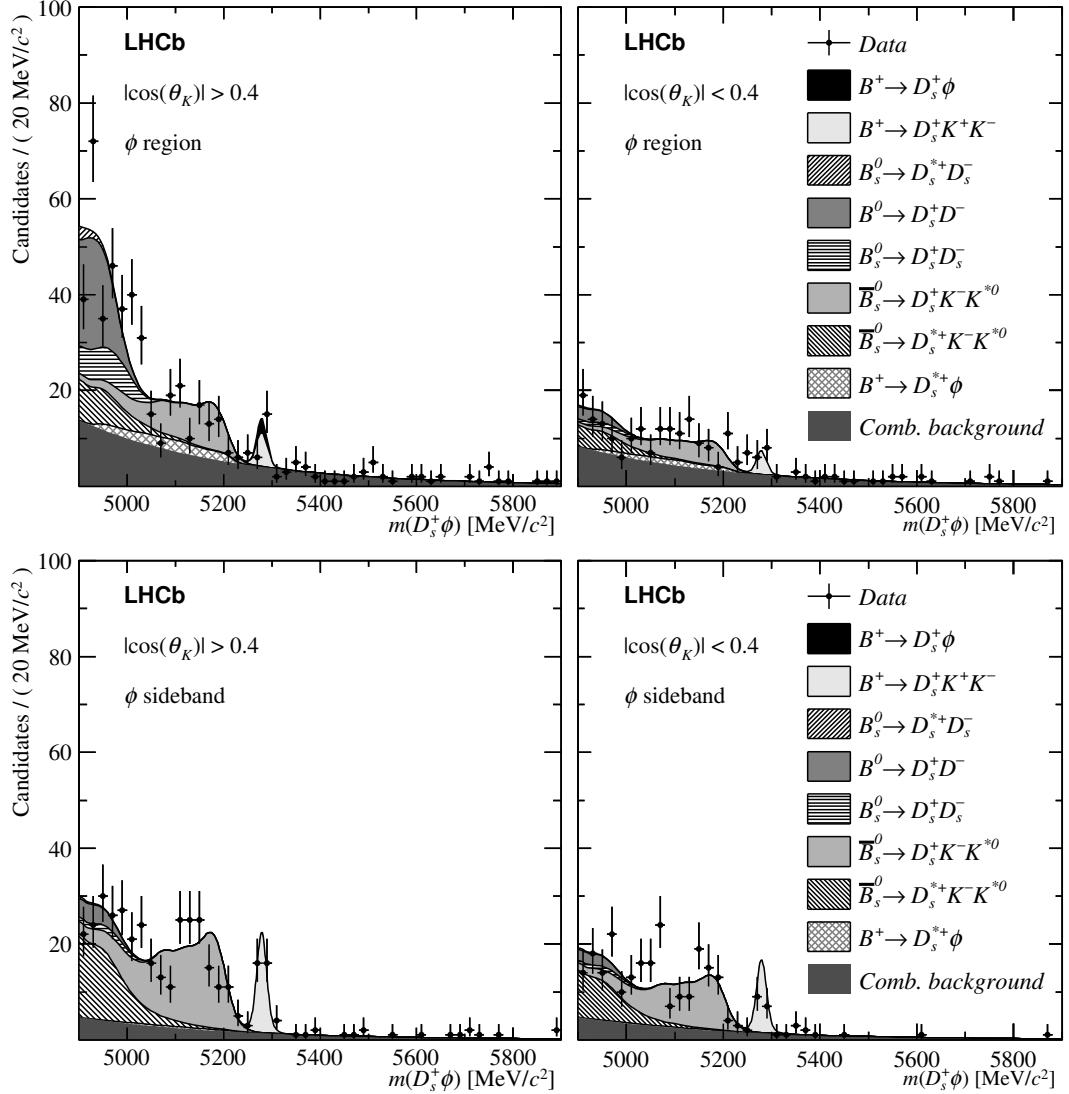


Figure 6.21: Invariant mass fits to  $B^+ \rightarrow D_s^+ \phi$  candidates

The fit to the signal mode is shown in Fig. 6.21. This figure includes the  $B^+ \rightarrow D_s^+ \phi$  candidates in both the  $|m(K^+ K^-)| < 10 \text{ MeV}/c^2$  ( $\phi$ -region) and  $10 < |m(K^+ K^-)| < 40 \text{ MeV}/c^2$  ( $\phi$ -sideband) categories. These are further split into the two  $\cos \theta_K$  categories in an analogous way to the normalisation decays.

The final values of the parameters as determined by the fit are tabulated in Tab. 6.5, including the branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ . This parameter includes a correction to account for the efficiencies of the signal and normalisation modes as discussed in Sec. 6.6.

Type	Parameter	Fit result
POI	Branching fraction $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)(\times 10^{-7})$	$1.17^{+1.56}_{-1.37}$
Shape	Combinatorial slope $c$	$(-3.3^{+0.4}_{-0.4}) \times 10^{-3}$
	Mean $B^+$ mass ( MeV/ $c^2$ )	$5279.10^{+0.17}_{-0.17}$
	Mass offset $\delta$ ( MeV/ $c^2$ )	$-1.975^{+0.362}_{-0.362}$
	Relative heights $\xi$	$0.68^{+0.05}_{-0.05}$
	$\sigma_1$ for $D_s^+ \rightarrow K^+ K^- \pi^+$ ( MeV/ $c^2$ )	$7.61^{+0.18}_{-0.18}$
	$\sigma_1$ for $D_s^+ \rightarrow K^+ \pi^- \pi^+$ ( MeV/ $c^2$ )	$8.77^{+0.59}_{-0.55}$
	$\sigma_1$ for $D_s^+ \rightarrow \phi \pi^+$ ( MeV/ $c^2$ )	$7.53^{+0.22}_{-0.21}$
	$\sigma_1$ for $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ ( MeV/ $c^2$ )	$8.30^{+0.36}_{-0.34}$
Yields	$N(B^+ \rightarrow (D_s^+ \rightarrow K^+ K^- \pi^+) \bar{D}^0)$	$1324^{+37}_{-36}$
	$N(B^+ \rightarrow (D_s^+ \rightarrow K^+ \pi^- \pi^+) \bar{D}^0)$	$182^{+14}_{-13}$
	$N(B^+ \rightarrow (D_s^+ \rightarrow \phi \pi^+) \bar{D}^0)$	$801^{+29}_{-28}$
	$N(B^+ \rightarrow (D_s^+ \rightarrow \pi^+ \pi^- \pi^+) \bar{D}^0)$	$369^{+20}_{-19}$
	$N(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$ in $D_s^+ \rightarrow K^+ K^- \pi^+$	$2827^{+52}_{-51}$
	$N(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$ in $D_s^+ \rightarrow K^+ \pi^- \pi^+$	$394^{+20}_{-20}$
	$N(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$ in $D_s^+ \rightarrow \phi \pi^+$	$1733^{+38}_{-37}$
	$N(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$ in $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$	$804^{+25}_{-25}$
	$N(D_s^{*+} \phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \rightarrow K^+ K^- \pi^+$	$139^{+19}_{-20}$
	$N(D_s^{*+} \phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \rightarrow K^+ \pi^- \pi^+$	$7^{+4}_{-4}$
	$N(D_s^{*+} \phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \rightarrow \phi \pi^+$	$67^{+10}_{-10}$
	$N(D_s^{*+} \phi + D_s^{(*)+} K^- K^{*0})$ in $D_s^+ \rightarrow K^+ \pi^- \pi^+$	$25^{+7}_{-6}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow K^+ K^- \pi^+$ in H1	$112^{+30}_{-25}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow K^+ K^- \pi^+$ in H2	$50^{+17}_{-14}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow K^+ \pi^- \pi^+$ in H1	$71^{+20}_{-17}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow K^+ \pi^- \pi^+$ in H2	$41^{+12}_{-10}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow \phi \pi^+$ in H1	$31^{+14}_{-11}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow \phi \pi^+$ in H2	$12^{+9}_{-6}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ in H1	$41^{+14}_{-11}$
	$N_{\text{comb}}(D_s^+ \bar{D}^0)$ $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ in H2	$35^{+14}_{-11}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow K^+ K^- \pi^+$ in H1	$107^{+23}_{-20}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow K^+ K^- \pi^+$ in H2	$52^{+12}_{-11}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow K^+ \pi^- \pi^+$ in H1	$19^{+7}_{-6}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow K^+ \pi^- \pi^+$ in H2	$18^{+5}_{-4}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow \phi \pi^+$ in H1	$38^{+14}_{-12}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow \phi \pi^+$ in H2	$27^{+8}_{-7}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ in H1	$42^{+10}_{-9}$
	$N_{\text{comb}}(D_s^+ \phi)$ $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ in H2	$25^{+7}_{-6}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow K^+ K^- \pi^+$ in H1	$33^{+13}_{-10}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow K^+ K^- \pi^+$ in H2	$26^{+11}_{-9}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow K^+ \pi^- \pi^+$ in H1	$18^{+7}_{-6}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow K^+ \pi^- \pi^+$ in H2	$14^{+5}_{-5}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow \phi \pi^+$ in H1	$5^{+6}_{-4}$

	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow \phi \pi^+$ in H2	$12^{+8}_{-7}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ in H1	$16^{+8}_{-6}$
	$N_{\text{comb}}(D_s^+ \phi)$ $\phi$ -sideband $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ in H2	$14^{+8}_{-6}$
Fractions	Ratio of $D_s^+ K^+ K^-$ to $D_s^+ \bar{D}^0$	$0.024^{+0.004}_{-0.004}$
	Fraction of $D_s^+ K^- K^{*0}$ in $(D_s^{*+} K^- K^{*0} + D_s^+ K^- K^{*0})$	$0.664^{+0.046}_{-0.044}$
	Fraction of $D_s^{*+} \phi$ in $(D_s^{(*)+} K^- K^{*0} + D_s^{*+} \phi)$	$0.172^{+0.103}_{-0.135}$
	Ratio of $D_s^+ D_{(s)}^{(*)-}$ to $(D_s^{*+} \phi + D_s^{(*)+} K^- K^{*0})$	$0.567^{+0.129}_{-0.107}$
	Fraction of $D_s^{*+} \bar{D}^0$ in $(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$ in H1	$0.302^{+0.031}_{-0.031}$
	Fraction of $D_s^{*+} \bar{D}^0$ in $(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$ in H2	$0.342^{+0.037}_{-0.037}$
	Fraction of normalisation part reco in H1	$0.593^{+0.005}_{-0.005}$
	Fraction of normalisation peak in H1	$0.592^{+0.010}_{-0.010}$
	Ratio of $D_s^{*+} \bar{D}^{*0}$ to $(D_s^{*+} \bar{D}^0 + D_s^+ \bar{D}^{*0})$	$0.607^{+0.016}_{-0.015}$

Table 6.5: The fit result with final values of all floating variables used in the fit model. Here H1 and H2 represent the  $|\cos \theta_K| > 0.4$  and  $|\cos \theta_K| < 0.4$  categories respectively.

## 6.6 Efficiency corrections

The branching fraction for  $B^+ \rightarrow D_s^+ \phi$  decays is determined by correcting the yields for the signal and normalisation channels by their respective efficiencies. In determining the efficiencies both the  $B^+ \rightarrow D_s^+ \phi$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays are assumed to be pseudo-two-body decays in which variations as a function of phase-space are negligible. This means the relative efficiency can be calculated as a simple ratio for each  $D_s^+$  decay mode, simplifying the correction considerably.

The efficiencies for each stage of selection are either determined from the appropriate simulation samples for the signal and normalisation decays, or from dedicated calibration data samples for the efficiencies involving particle identification requirements. Each set of efficiencies are determined with respect to the previous one, such that the total efficiency is given by the product of each of them.

### 6.6.1 Efficiencies from simulation

The efficiencies for each step of selection are listed separately for the different  $D_s^+$  decay modes in Table 6.6. These steps are closely related to the stages previously described in Chapter 4, however more specific details of what is included in each step

is described here. The efficiencies are calculated separately for the different years of data taking, however they are combined here, weighting according to the relative contributions to the total data set.

Requirement	$D_s^+$ mode	$\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)$	$\epsilon(B^+ \rightarrow D_s^+ \phi)$	Ratio
Acceptance	$K^+ K^- \pi^+$	$17.42 \pm 0.03$	$18.35 \pm 0.04$	$0.950 \pm 0.003$
	$K^+ \pi^- \pi^+$	$16.53 \pm 0.04$	$17.36 \pm 0.04$	$0.952 \pm 0.003$
	$\pi^+ \pi^- \pi^+$	$15.91 \pm 0.03$	$16.77 \pm 0.04$	$0.949 \pm 0.003$
Reconstruction	$K^+ K^- \pi^+$	$2.11 \pm 0.02$	$1.96 \pm 0.02$	$1.078 \pm 0.013$
	$K^+ \pi^- \pi^+$	$2.27 \pm 0.02$	$2.06 \pm 0.02$	$1.095 \pm 0.014$
	$\pi^+ \pi^- \pi^+$	$2.36 \pm 0.02$	$2.13 \pm 0.02$	$1.105 \pm 0.014$
Trigger	$K^+ K^- \pi^+$	$93.3 \pm 0.2$	$93.1 \pm 0.2$	$1.003 \pm 0.003$
	$K^+ \pi^- \pi^+$	$95.1 \pm 0.2$	$93.5 \pm 0.2$	$1.017 \pm 0.003$
	$\pi^+ \pi^- \pi^+$	$95.3 \pm 0.2$	$93.7 \pm 0.2$	$1.017 \pm 0.003$
Mass window	$K^+ K^- \pi^+$	$96.2 \pm 0.2$	$94.1 \pm 0.2$	$1.022 \pm 0.003$
	$K^+ \pi^- \pi^+$	$94.9 \pm 0.2$	$94.1 \pm 0.2$	$1.008 \pm 0.003$
	$\pi^+ \pi^- \pi^+$	$93.7 \pm 0.2$	$92.4 \pm 0.2$	$1.015 \pm 0.004$
Veto	$K^+ K^- \pi^+$	$99.9 \pm 0.0$	$95.2 \pm 0.2$	$1.049 \pm 0.002$
	$K^+ \pi^- \pi^+$	$99.9 \pm 0.0$	$92.4 \pm 0.3$	$1.080 \pm 0.004$
	$\pi^+ \pi^- \pi^+$	$99.9 \pm 0.0$	$90.3 \pm 0.3$	$1.105 \pm 0.004$
Charmless	$K^+ K^- \pi^+$	$72.6 \pm 0.4$	$100.0 \pm 0.0$	$0.726 \pm 0.004$
	$K^+ \pi^- \pi^+$	$65.8 \pm 0.4$	$63.0 \pm 0.5$	$1.044 \pm 0.011$
	$\pi^+ \pi^- \pi^+$	$66.0 \pm 0.4$	$82.9 \pm 0.4$	$0.796 \pm 0.006$
$\chi^2_{\text{IP}}$	$K^+ K^- \pi^+$	$100.0 \pm 0.0$	$96.2 \pm 0.2$	$1.039 \pm 0.002$
	$K^+ \pi^- \pi^+$	$100.0 \pm 0.0$	$95.7 \pm 0.3$	$1.045 \pm 0.003$
	$\pi^+ \pi^- \pi^+$	$100.0 \pm 0.0$	$95.9 \pm 0.3$	$1.043 \pm 0.003$

Table 6.6: Efficiencies (in %) determined from simulation samples for signal and normalisation decays and the ratio  $\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)/\epsilon(B^+ \rightarrow D_s^+ \phi)$  for each  $D_s^+$  decay mode. The errors are statistical.

**Acceptance:** this accounts for the fraction of generated decays in which all five final state tracks end up within the LHCb detector's acceptance. For all  $D_s^+$  decay modes the signal decay has a slightly higher efficiency than the normalisation channel. This may be because of the kinematics of the  $\phi \rightarrow K^+ K^-$  decay in which the two kaon tracks are typically close to one another and therefore more likely to both be in the acceptance.

**Reconstruction:** this efficiency determines the fraction of decays in which all tracks are well reconstructed and combined into a suitable candidate. The candidate

must pass all requirements outlined in the *Stripping Line* for the specific decay. The reconstruction selections used for both the signal and normalisation decays explicitly require that the event in which the candidate was found had fired the trigger. Therefore this efficiency includes some, but not all, of the trigger efficiency. This results in small efficiencies around 2% for each mode. Here, the efficiencies are slightly larger for the normalisation channel than the signal. This may be because the kaons from the  $\phi$  decay are typically lower momentum than those from the  $\bar{D}^0$ , and therefore less likely to be reconstructed.

**Trigger:** this efficiency effectively accounts for the likelihood that the candidate passed at least one of the requirements in Sec. 4.2, given a trigger fired in that event. These efficiencies have very high values around 94%.

**Mass windows:** this represents the efficiency for the candidates to be within the invariant mass windows for the  $D_s^+$  and  $\phi$  or  $\bar{D}^0$  mesons. Again, this is slightly high for the normalisation than the signal. The  $\phi$  meson invariant mass peak a fairly long tail extending to higher invariant masses that may be the cause.

**Vetoies:** the efficiency of the kinematic vetoes described in Sec.4.3.5 is included in this quantity. The misidentified  $D$  and  $A_c^+$  hadron vetoes target the  $D_s^+$  meson, present in both the signal and normalisation decays. The relative efficiency is assumed to be one for these specific vetoes therefore not included in this efficiency. The systematic uncertainty resulting from this assumption is discussed in Sec.6.7. Most of the kinematic vetoes are only applied to the signal mode, hence why the normalisation channel efficiencies are almost 100%.

**Charmless:** the requirements applied to the flight distance significance of the  $D_s^+$  meson is tuned differently for each  $D_s^+$  decay and signal and normalisation. As a result, the efficiencies have large variations between the different modes.

**$\chi_{\text{IP}}^2$ :** the efficiency of the  $\chi_{\text{IP}}^2$  requirements for the  $B^+$  and  $D_s^+$  candidates as detailed in Sec.4.3.8 are included in this quantity. These are only applied to the signal mode therefore the normalisation mode is 100% efficient.

### 6.6.2 Efficiencies requiring calibration samples

The efficiencies of the particle identification and MVA requirements are both determined using input from dedicated calibration samples as the distributions are known to be poorly represented in simulations. The method used here differs from that already described in Sec. 5.6.2 as the  $B^+ \rightarrow D_s^+ \phi$  decay can be assumed to be a pseudo-two-body decay in which phase-space dependent efficiencies are not required.

Requirement	$D_s^+$ mode	$\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)$	$\epsilon(B^+ \rightarrow D_s^+ \phi)$	Ratio
PID	$K^+ K^- \pi^+$	$88.1 \pm 0.1$	$89.9 \pm 0.0$	$0.980 \pm 0.001$
	$K^+ \pi^- \pi^+$	$86.7 \pm 0.2$	$88.8 \pm 0.1$	$0.977 \pm 0.002$
	$\pi^+ \pi^- \pi^+$	$85.3 \pm 0.1$	$87.0 \pm 0.0$	$0.980 \pm 0.001$
MVA	$K^+ K^- \pi^+$	$53.2 \pm 0.4$	$57.1 \pm 0.4$	$0.932 \pm 0.010$
	$K^+ \pi^- \pi^+$	$42.9 \pm 0.4$	$46.0 \pm 0.5$	$0.932 \pm 0.013$
	$\pi^+ \pi^- \pi^+$	$46.3 \pm 0.4$	$49.7 \pm 0.4$	$0.933 \pm 0.012$

Table 6.7: Efficiencies (in %) determined from the relevant calibration and validation samples using input from simulation and the ratio  $\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)/\epsilon(B^+ \rightarrow D_s^+ \phi)$  for each  $D_s^+$  decay mode. The errors are statistical.

#### PID efficiency

The efficiency of the particle identification requirements described in Sec. 4.3.2 are determined using a package called PIDCALIB [121]. This uses calibrations samples for the different particle species to determine the fraction of candidates passing the various PID variable requirements. The samples are background-subtracted to isolate the distributions of the PID variables for the tracks of interest. The calibration samples for both  $K^+$  and  $\pi^+$  mesons are collected from a sample of  $D^{*+} \rightarrow (D^0 \rightarrow K^+ \pi^-)\pi^+$  decays, using the decay products of the  $D^0$  decay.

Unlike in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays, the PID variable distributions of the calibration samples are parametrised using a binned approach, with three characterising variables; the momentum  $p$ , pseudo-rapidity  $\eta$ , and total number of tracks  $n_{\text{Track}}$ . Input from the signal and normalisation simulation samples is used to determine per-candidate efficiencies. The characteristics  $(p, \eta, n_{\text{Track}})$  of each simulated decay is used to find the corresponding calibration PID variable distribution. The

per-candidate efficiency is calculated by integrating this distribution above the PID requirement. The total efficiency is given by the average of the per-candidate efficiencies. The efficiencies of the PID requirements on the signal and normalisation for the different  $D_s^+$  decay modes are listed in Table 6.7.

### MVA efficiency

The efficiency of the MVA requirements is determined using the method outlined in Sec. 4.3.7. The validation samples of  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  and  $B_s^0 \rightarrow J/\psi \phi$  decays are binned in transverse momentum  $p_T$  and flight distance significance  $\chi_{\text{FD}}^2$ . The background-subtracted MVA responses are extracted for each bin. Samples of simulated decays are iterated through, finding the  $p_T$  and  $\chi_{\text{FD}}^2$  values for each candidate and calculating the MVA cut efficiency from the corresponding validation samples. The total efficiency for each candidate is the product of the  $D_s^+$  and  $\phi$  MVA selection efficiencies. The total efficiency for the whole sample is then given by the sum of the per-candidate efficiencies as listed in Table 6.7.

### 6.6.3 Total efficiencies

The total relative efficiency between the signal and normalisation decays is determined as the product of each contributing relative efficiency

$$\epsilon^{\text{Tot.}} = \epsilon^{\text{Accp.}} \times \epsilon^{\text{Reco.}|\text{Accp.}} \times \epsilon^{\text{Trig.}|\text{Reco.}} \times \epsilon^{\text{Mass.}|\text{Trig.}} \times \epsilon^{\text{Veto.}|\text{Mass.}} \times \epsilon^{\text{FD}|\text{Veto.}} \\ \times \epsilon^{\text{IP}|\text{FD}} \times \epsilon^{\text{PID}|\text{IP}} \times \epsilon^{\text{MVA}|\text{PID}}, \quad (6.21)$$

where each relative efficiency  $x$  is determined with respect to the previous selection step  $y$  as  $\epsilon^{x|y}$ . The total relative efficiencies are listed for each  $D^+$  decay mode in Table 6.8. These values are used as an input in the simultaneous fit to the signal and normalisation modes to correct the yields ratios for each  $D_s^+$  decay mode category.

$D_s^+$ decay mode	Ratio
$K^+ K^- \pi^+$	$0.757 \pm 0.014$
$K^+ \pi^- \pi^+$	$1.144 \pm 0.027$
$\pi^+ \pi^- \pi^+$	$0.907 \pm 0.019$

Table 6.8: The total relative efficiency  $\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)/\epsilon(B^+ \rightarrow D_s^+ \phi)$  determined for each  $D_s^+$  decay mode. The values for the individual years that contribute to the data set are weighted according to the size of their relative contribution. The errors are statistical.

## 6.7 Systematic uncertainties

The sources of systematic uncertainty are broadly similar to those considered in the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays. The more complex fit strategy and model requires most of these systematics to be reassessed and additional source included.

### 6.7.1 Relative efficiencies

The yields of signal and normalisation decays are corrected by their relative selection efficiencies, calculated separately for each  $D_s^+$  decay mode. A number of sources contribute to the systematic uncertainty in these values.

**Simulation statistics:** limited simulation samples are used to determine some of the selection efficiencies. A systematic uncertainty of 2% is assigned to account for the bias this might incur.

**Particle identification:** the efficiency of the PID requirements are calculated using calibration samples within the PIDCALIB package. These are also of a limited size and possibly affected by the choice of binning scheme. To first order these should affect the signal and normalisation decays to the same extent as the same final state is used. However, differences in the kinematics could invalidate this assumption. Therefore a systematic uncertainty of 0.5% per track is assigned to the use of the PIDCALIB efficiencies. This is assumed to be uncorrelated for the five signal and five normalisation tracks, leading to a total uncertainty in the relative efficiency of 2.0%.

**Veto efficiency:** similar to the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  the efficiency of the misidentified  $D$  and  $\Lambda_c^+$  hadron veto applied to the  $D_s^+$  meson is assumed to be the same for the signal and normalisation channel. A similar systematic uncertainty of 1.4% is assigned to account for the uncertainty incurred if the relative efficiency were to be fully calculated.

**MVA efficiency:** unlike the search for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays, the efficiencies for the MVA requirements are just calculated using the  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  samples used in the validation of the MVA methods. There is no need account for the phase-space dependence, so the simulation samples have not been corrected. These data-driven efficiencies are affected by the same sources as previously discussed in 5.7.1. The validation samples are binned in four bins of both  $p_T$  and  $\chi_{\text{FD}}^2$ . The choice of binning scheme is varied and the resulting variation in the efficiencies assigned as a systematic uncertainty. The yields of  $B_s^0 \rightarrow J/\psi \phi$  and  $\bar{B}_s^0 \rightarrow D_s^+ \pi^-$  are limited, which could lead to fluctuations in the efficiencies determined. The quantity  $1/\sqrt{N}$  for the smallest sample is assigned as the systematic uncertainty.

Some differences are observed in the distributions of  $\chi_{\text{IP}}^2$  between simulations and data for the normalisation mode (Fig. 4.18). The simulation samples are re-weighted to match the data distributions and the efficiencies recalculated. The resulting difference is included as a systematic uncertainty.

In training the MVA methods, the distribution of the MVA classifier shows some discrepancies between the training and validation samples at low classifier values. To quantify the effect this may have on the relative efficiencies the samples are swapped and the efficiencies recalculated. The resulting difference is assigned as the associated systematic uncertainty.

The total systematic uncertainty attributed to the relative MVA efficiencies is 6.2%.

### 6.7.2 Signal and normalisation PDFs

The PDFs used to describe the signal and normalisation channel invariant mass distributions are determined using input from fits to simulated decays. A similar approach to that already described in Sec. 5.7.1 is used to account for the systematic uncertainty that these fixed values contribute. The fit to data is repeated many times, each time the values of the fixed parameters are sampled from Gaussian distributions whose width is given by the associated statistical uncertainty from the fit to simulations. The resulting spread in the branching fraction obtained by re-sampling all fixed parameters simultaneously is  $0.036 \times 10^{-7}$ , assigned as the systematic uncertainty.

### 6.7.3 Background PDFs

Similarly the background PDFs require inputs determined from simulations. A large number of different properties of the PDFs are varied to determine the systematic uncertainty the fixed values incur. The resulting spread in the branching fraction is assigned as the systematic uncertainty.

The helicity fraction of  $B^+ \rightarrow D_s^{*+} \phi$  decays is not measured and therefore assumed to be 0.5. This assumption is varied across all allowed values. Additionally, the PDF for both  $B^+ \rightarrow D_s^{*+} \phi$  decay and the  $B^+ \rightarrow D_s^{*+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^{*0}$  decays that contribute to the normalisation mode require the kinematic limits  $a$  and  $b$  to be defined. These values are varied.

The backgrounds to the signal mode require the fraction of decays expected in each  $m(K^+ K^-)$  and  $\cos \theta_K$  category to be determined from simulations. These are each have an associated statistical uncertainty due to the limited simulation samples, therefore these fractions are varied by sampling a Gaussian whose width is determined by the uncertainty.

In the fit to the normalisation mode the partially reconstructed background  $B^+ \rightarrow D_s^{*+} \bar{D}^{*0}$  is approximated with a single PDF, rather than eight as required by the combinations of missed particles and helicity components. This function is replaced

with a DCB function with free widths mean and tail parameters. This results in a negligible change in the branching fraction.

The combinatorial background is parametrised with an exponential and constrained to have the same slope in all categories. To determine if these choices lead to systematic uncertainty the fit is repeated with the slope allowed to be different in the different  $D_s^+$  decay mode categories. Additionally an exponential plus constant offset is tried as a parametrisation instead. The resulting changes in the branching fractions are included as systematic uncertainties.

The total systematic uncertainty associated to the background PDFs is determined to be  $0.685 \times 10^{-7}$

#### 6.7.4 Charmless contribution

A residual yield of charmless and single-charm backgrounds are expected be present in the final dataset as detailed in Table 4.9. The nominal measurement for the branching fraction neglects these contributions. The branching fraction is recalculated assuming the expected yields contribute and the difference assigned as the systematic uncertainty. This is likely to be an overestimation of the possible difference as the charmless candidates are likely to have a wider distribution than the signal decays. This is a result of the  $D_s^+$  mass constraint applied to the candidates when calculating the fitted  $B^+$  mass.

The charmless contributions lead to a systematic uncertainty of 2%.

#### 6.7.5 $B^+ \rightarrow D_s^+ K^+ K^-$ model assumption

The search for  $B^+ \rightarrow D_s^+ \phi$  candidates includes a component for  $B^+ \rightarrow D_s^+ K^+ K^-$  decays that didn't proceed via a  $\phi$  meson. As detailed in Sec. 6.1.2 various assumptions go into the choice of  $B^+ \rightarrow D_s^+ K^+ K^-$  model used to determine the fraction of these decays expected in each  $m(K^+ K^-)$  and  $\cos \theta_K$  category. The final choice of fractions detailed in Table 6.2 are determined with associated uncertainties, calculated by taking the range of fractions in the models thought to be reasonable. The fit to data is performed many times changing these fixed fractions to values taken

from Gaussian distributions whose widths are given by these assigned uncertainties. The resulting spread in the measured branching fraction is used as a proxy for the systematic uncertainty resulting from this choice of model.

An additional systematic uncertainty is included because these fractions are taken directly from simulations produced by LAURA++, rather than full detector simulations. This is calculated by taking the difference in the fractions found for the  $B^+ \rightarrow D_s^+ \phi$  decay generated using LAURA++ and the full LHCb simulation.

### 6.7.6 Total systematic uncertainty

The sources of systematic uncertainty are listed in Table 6.9 in decreasing order. The total is also included, calculated by summing the contributions in quadrature. Additionally the uncertainty arising from the externally measured branching fractions is included.

Source of Uncertainty	Systematic Uncertainty	
	Relative	Absolute ( $\times 10^{-7}$ )
Background PDF parametrisation	-	0.685
Choice of $B^+ \rightarrow D_s^+ K^+ K^-$ model	-	0.379
MVA relative efficiency	6.2%	0.072
Signal PDF parametrisation	-	0.036
Charmless contribution	2.0%	0.023
Simulation statistics	2.0%	0.023
PID relative efficiency	2.0%	0.023
Veto relative efficiency	1.4%	0.016
Using LAURA++ rather full sim.	1.1%	0.013
Total		0.788
Normalisation		0.1

Table 6.9: Contributions to the total systematic uncertainty in the search for  $B^+ \rightarrow D_s^+ \phi$  decays. The contribution from the external measurements of the normalisation channel branching fraction is also included.

## 6.8 Results

The fit to  $B^+ \rightarrow D_s^+ \phi$  candidates finds a total yield of  $N(B^+ \rightarrow D_s^+ \phi) = 5.3 \pm 6.7$ , summed across all categories and  $D_s^+$  meson decay modes. A yield of  $N(B^+ \rightarrow$

$D_s^+ K^- K^+) = 65 \pm 10$  is found, consistent with the yield obtained from the full  $B^+ \rightarrow D_s^+ K^+ K^-$  measurement. The branching fraction for  $B^+ \rightarrow D_s^+ \phi$  decays is calculated as

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) = R \times \frac{\mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)}{\mathcal{B}(\phi \rightarrow K^+ K^-)} \times \mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0), \quad (6.22)$$

where the branching fraction  $\mathcal{B}(\phi \rightarrow K^+ K^-) = 0.489 \pm 0.005$  has been used [28].

The free variable  $R$  is defined to be the ratio of the signal and normalisation yields, corrected for the selection efficiencies. The yield of signal candidates in each  $D_s^+$  mode is constructed from  $R$  and the normalisation yield for the given  $D_s^+$  decay mode,  $N(B^+ \rightarrow D_s^+ \bar{D}^0)$ . The product of these two quantities is corrected by the ratio of selection efficiencies

$$N(B^+ \rightarrow D_s^+ \phi) = R \times N(B^+ \rightarrow D_s^+ \bar{D}^0) \times \frac{\epsilon(B^+ \rightarrow D_s^+ \phi)}{\epsilon(B^+ \rightarrow D_s^+ \bar{D}^0)}. \quad (6.23)$$

The simultaneous fit measures a single value of  $R$  for all  $D_s^+$  decay mode categories. From an ensemble of pseudoexperiments,  $R$  is distributed normally. It can be written as the ratio of signal and normalisation branching fractions using Eq. 6.22. The value is determined to be

$$R = \frac{\mathcal{B}(B^+ \rightarrow D_s^+ \phi)}{\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0)} \times \frac{\mathcal{B}(\phi \rightarrow K^+ K^-)}{\mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)} = (1.6_{-1.9}^{+2.2} \pm 1.1) \times 10^{-3}, \quad (6.24)$$

where the first uncertainty is statistical and the second systematic. This corresponds to a branching fraction for  $B^+ \rightarrow D_s^+ \phi$  decays of

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) = (1.2_{-1.4}^{+1.6} \pm 0.8 \pm 0.1) \times 10^{-7}, \quad (6.25)$$

where the first uncertainty is statistical, the second systematic, and the third results from the uncertainty on the branching fractions  $\mathcal{B}(B^+ \rightarrow D_s^+ \bar{D}^0)$ ,  $\mathcal{B}(\phi \rightarrow K^+ K^-)$  and  $\mathcal{B}(\bar{D}^0 \rightarrow K^+ K^-)$ . Considering only the statistical uncertainty, the significance of the  $B^+ \rightarrow D_s^+ \phi$  signal is 0.8 standard deviations ( $\sigma$ ).

The branching fraction for  $B^+ \rightarrow D_s^+ \phi$  decays is also determined separately for the different  $D_s^+$  decay modes included in the search. These are found to be

$$\begin{aligned} D_s^+ \rightarrow \phi\pi^+ & : \mathcal{B}(B^+ \rightarrow D_s^+ \phi) = +2.7^{+2.9}_{-2.3} \times 10^{-7} \\ D_s^+ \rightarrow K^+K^-\pi^+ & : \mathcal{B}(B^+ \rightarrow D_s^+ \phi) = +1.2^{+2.2}_{-1.8} \times 10^{-7} \\ D_s^+ \rightarrow \pi^+\pi^-\pi^+ & : \mathcal{B}(B^+ \rightarrow D_s^+ \phi) = -9.4^{+3.6}_{-2.8} \times 10^{-7} \\ D_s^+ \rightarrow K^+\pi^-\pi^+ & : \mathcal{B}(B^+ \rightarrow D_s^+ \phi) = +3.7^{+1.2}_{-7.6} \times 10^{-7}, \end{aligned} \quad (6.26)$$

where these results correspond to the following yields for each  $D_s^+$  decay mode

$$\begin{aligned} D_s^+ \rightarrow \phi\pi^+ & : N(B^+ \rightarrow D_s^+ \phi) = +3.9^{+4.2}_{-3.3} \\ D_s^+ \rightarrow K^+K^-\pi^+ & : N(B^+ \rightarrow D_s^+ \phi) = +2.7^{+5.0}_{-4.2} \\ D_s^+ \rightarrow \pi^+\pi^-\pi^+ & : N(B^+ \rightarrow D_s^+ \phi) = -5.2^{+2.0}_{-1.6} \\ D_s^+ \rightarrow K^+\pi^-\pi^+ & : N(B^+ \rightarrow D_s^+ \phi) = +0.8^{+2.6}_{-1.7}. \end{aligned} \quad (6.27)$$

A visual representation of these measurements are shown in Fig. 6.22 along with the value determined using all modes simultaneously.

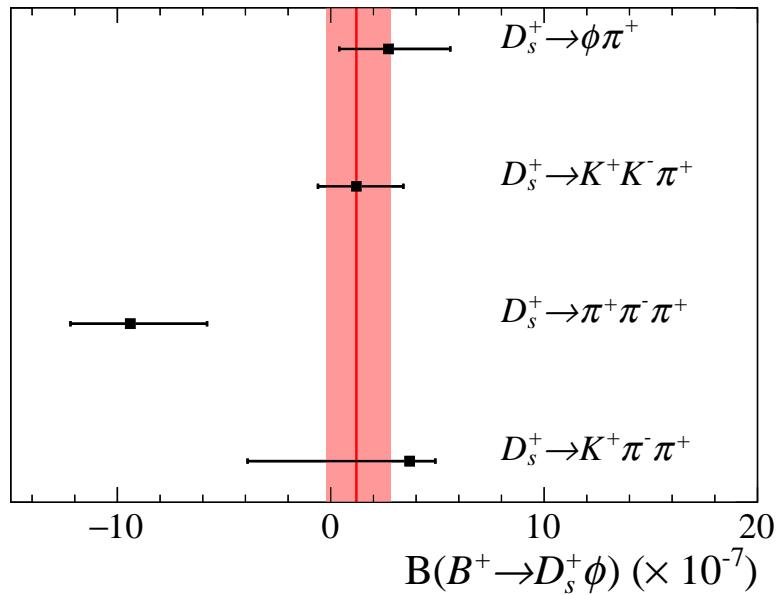


Figure 6.22: Results split for different  $D_s^+$  decay modes.

### 6.8.1 Limit setting

The measured branching fraction,  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi) = (1.2^{+1.6}_{-1.4} \pm 0.8 \pm 0.1) \times 10^{-7}$ , is not significant enough to constitute evidence or an observation for the  $B^+ \rightarrow D_s^+ \phi$  decay

and is consistent with a branching fraction of zero. Whilst this measurement is useful in itself, for example it could provide constraints in combination with other results, it is also useful to set a limit on the branching fraction for a more straightforward comparison with theoretical predictions. Three different methods of limit estimation are attempted. These methods make different assumptions are therefore applicable in slightly different situations.

### The $\text{CL}_S$ method

The first method tried is the  $\text{CL}_S$  method, widely used in the high energy physics community. This method tests the p-value of a signal plus background hypothesis,  $\text{CL}_{S+B}$ , against a background only hypothesis,  $\text{CL}_B$ ,

$$\text{CL}_S = \frac{\text{CL}_{S+B}}{\text{CL}_B}. \quad (6.28)$$

The free parameters in the fit other than the POI are considered nuisance parameters. This method is implemented using the RooStats package [126] within the ROOT framework.

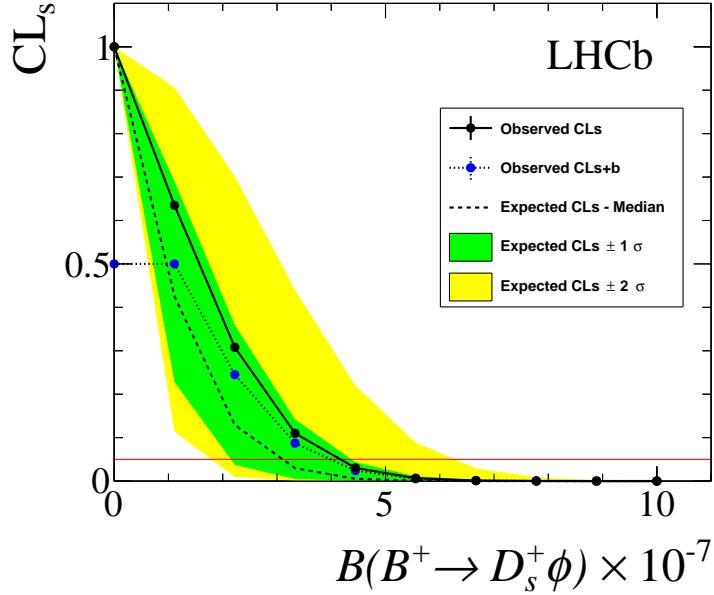
The  $\text{CL}_S$  distribution is shown in Fig. 6.23 which includes the  $1\sigma$  bands in green and  $2\sigma$  bands in yellow. This limit is determined in the asymptotic limit [127]. The 95% upper limit is determined as the point where the  $\text{CL}_S$  value falls below 5% as illustrated by the red line. This corresponds to a 95% limit of

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) < 4.2 \times 10^{-7}. \quad (6.29)$$

### The profile likelihood method

An upper limit at 95% confidence limit is determined for the branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$  using the profile likelihood method. This is calculated by determining the value of the branching fraction  $x_U$  that satisfies the equation

$$\frac{\int_0^{x_U} \mathcal{L}(x) dx}{\int_0^\infty \mathcal{L}(x) dx} = 0.95, \quad (6.30)$$

Figure 6.23: CL<sub>S</sub> limit determination.

where  $\mathcal{L}(x)$  is the profile likelihood as a function of the branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ . This Bayesian method integrates the prior knowledge about the branching fraction, namely that the value must be greater than or equal to zero; the profile likelihood is integrated from zero upwards. The 95% CL limit determined when considering only statistical uncertainties is

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) < 4.1 \times 10^{-7}. \quad (6.31)$$

To account for systematic uncertainty, the likelihood is convolved with a Gaussian distribution with a width given by the systematic uncertainty. The likelihood and difference in the log-likelihood are shown in Fig. 6.24, with and without the systematic uncertainty included. The limit at 95% CL including the systematic uncertainty is determined to be

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) < 4.4 \times 10^{-7}. \quad (6.32)$$

### The Feldman-Cousins method

Upper limits at 95% and 90% confidence levels (CL) are also determined using the Feldman-Cousins approach [128]. An ensemble of pseudo-experiments is generated

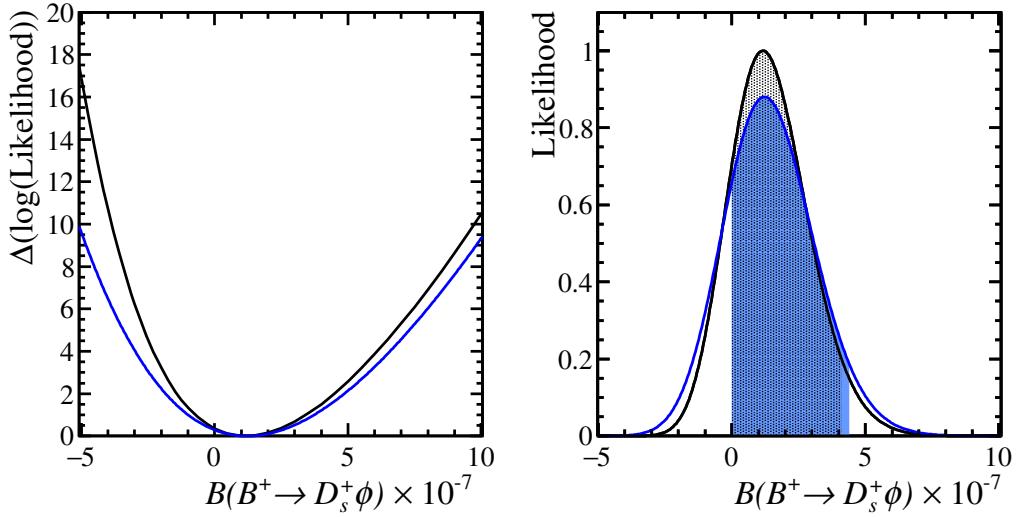


Figure 6.24: Bayesian profile likelihood limit determination: the (left) difference in the log-likelihood and (right) the likelihood as a function of the assumed  $B^+ \rightarrow D_s^+ \phi$  branching fraction. The black distributions include only the statistical uncertainty, whilst the blue also include the systematic uncertainty. The shaded regions represent the areas integrated to determine the 95% CL limits.

for different values of the branching fraction  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ . These generated pseudo-experiments are then fitted with the nominal fit model to calculate the fitted branching fraction and associated statistical uncertainty,  $\sigma_{\text{stat}}$ . This method constructs confidence bands based on a likelihood ratio method, calculating the probability of fitting a branching fraction for a given generated branching fraction. This probability is assumed to follow a Gaussian distribution with width  $\sigma = \sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{syst}}^2}$ , where  $\sigma_{\text{stat}}$  and  $\sigma_{\text{syst}}$  are the statistical and systematic uncertainties. The dominant source of systematic uncertainty in this measurement is from the background PDFs. As the size of this uncertainty is not expected to vary as a function of the generated branching fraction,  $\sigma_{\text{syst}}$  is assumed to be constant. Nuisance parameters are accounted for using the plug-in method [120]. The generated confidence bands are shown in Fig. 6.25, where the statistical-only 90% and 95% confidence level bands are shown, along with the 95% confidence level band with systematic uncertainty included. This corresponds to a statistical-only 95% (90%) confidence level of  $\mathcal{B}(B^+ \rightarrow D_s^+ \phi) < 4.4 \times 10^{-7}$  ( $3.9 \times 10^{-7}$ ),

and a 95% (90%) confidence level including systematic uncertainties of

$$\mathcal{B}(B^+ \rightarrow D_s^+ \phi) < 4.9 \times 10^{-7} \text{ (} 4.2 \times 10^{-7} \text{).} \quad (6.33)$$

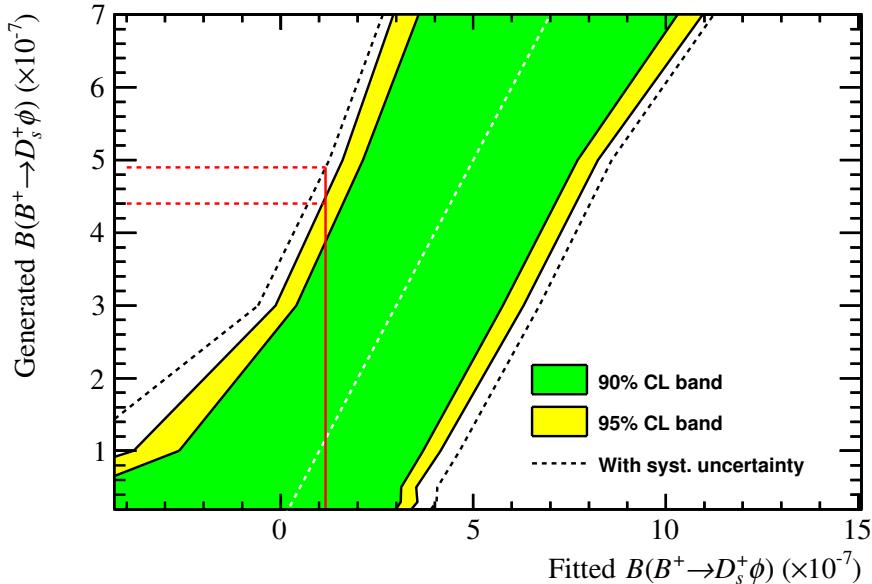


Figure 6.25: Confidence bands produced using the Feldman-Cousins approach. The green and yellow bands represent the statistical-only 90% and 95% confidence level bands and the black dotted line represents the 95% limit including systematic uncertainties. The measured value of the branching fraction is shown by the vertical red line, and the corresponding 95% confidence levels, with and without systematic uncertainties, are represented by the dotted red lines.

### 6.8.2 Comparison to the previous measurement

The limit on the  $B^+ \rightarrow D_s^+ \phi$  branching fraction presented here and in Ref [21] supersede the previous evidence reported by the LHCb collaboration in Ref. [129]. The updated analysis takes advantage of a much larger dataset now available at LHCb. This update determines that there is a sizeable contribution from  $B^+ \rightarrow D_s^+ K^+ K^-$  decays that contribute within the  $\phi$ -meson mass window that was previously not considered. The result is consistent with the prediction that rescattering contributions to  $B^+ \rightarrow D_s^+ \phi$  decays are small.

## **Appendix A**

**Invariant mass fit to  $B^+ \rightarrow D_s^+ \phi$   
candidates in each  $D_s^+$  decay mode  
category**

A. Invariant mass fit to  $B^+ \rightarrow D_s^+ \phi$  candidates in each  $D_s^+$  decay mode category 192

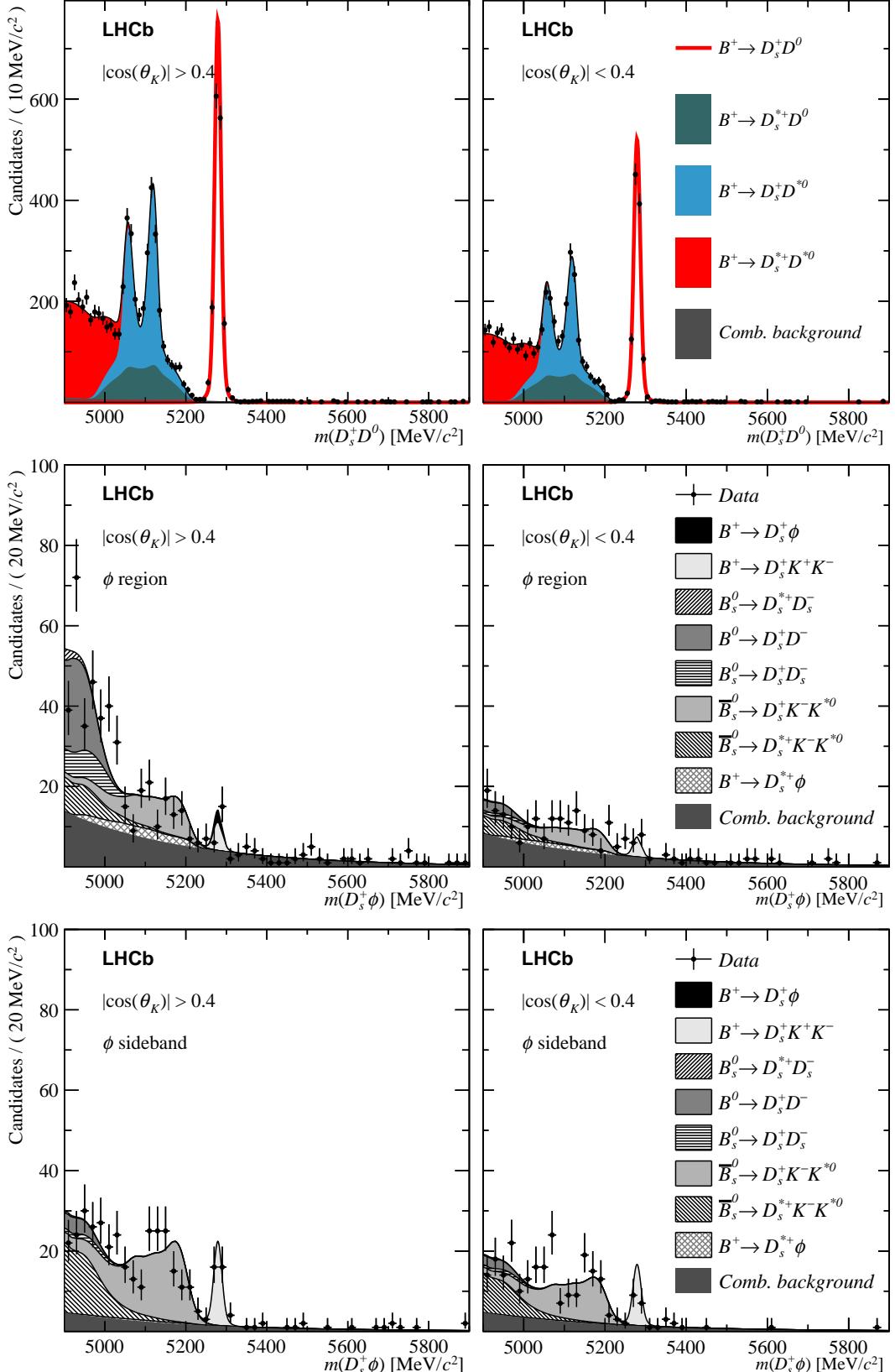


Figure A.1: Invariant mass fits to all  $B^+ \rightarrow D_s^+ \phi$  candidates

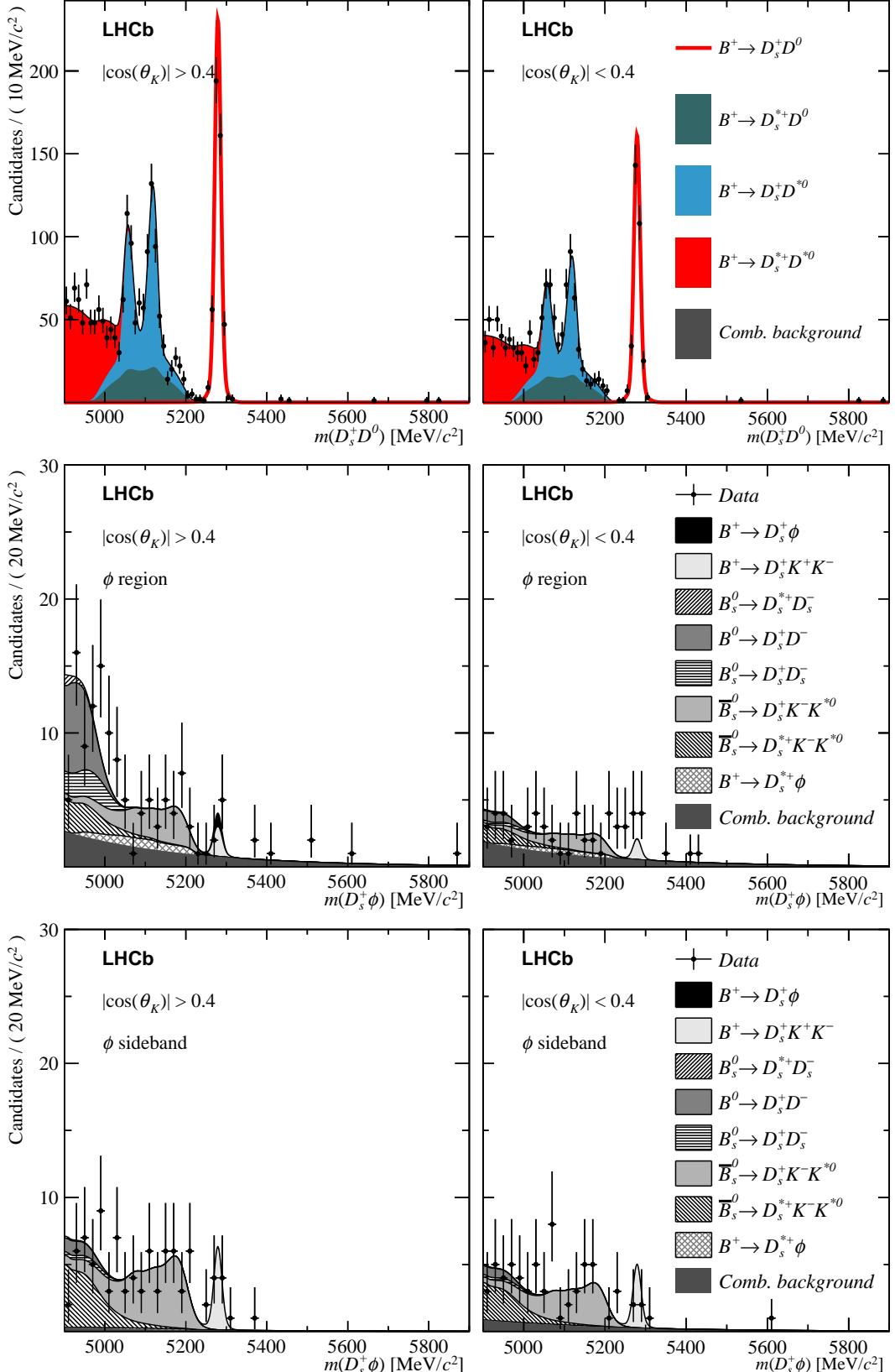


Figure A.2: Invariant mass fits to  $B^+ \rightarrow D_s^+ \phi$  candidates with  $D_s^+ \rightarrow \phi \pi^+$ .

A. Invariant mass fit to  $B^+ \rightarrow D_s^+ \phi$  candidates in each  $D_s^+$  decay mode category 194

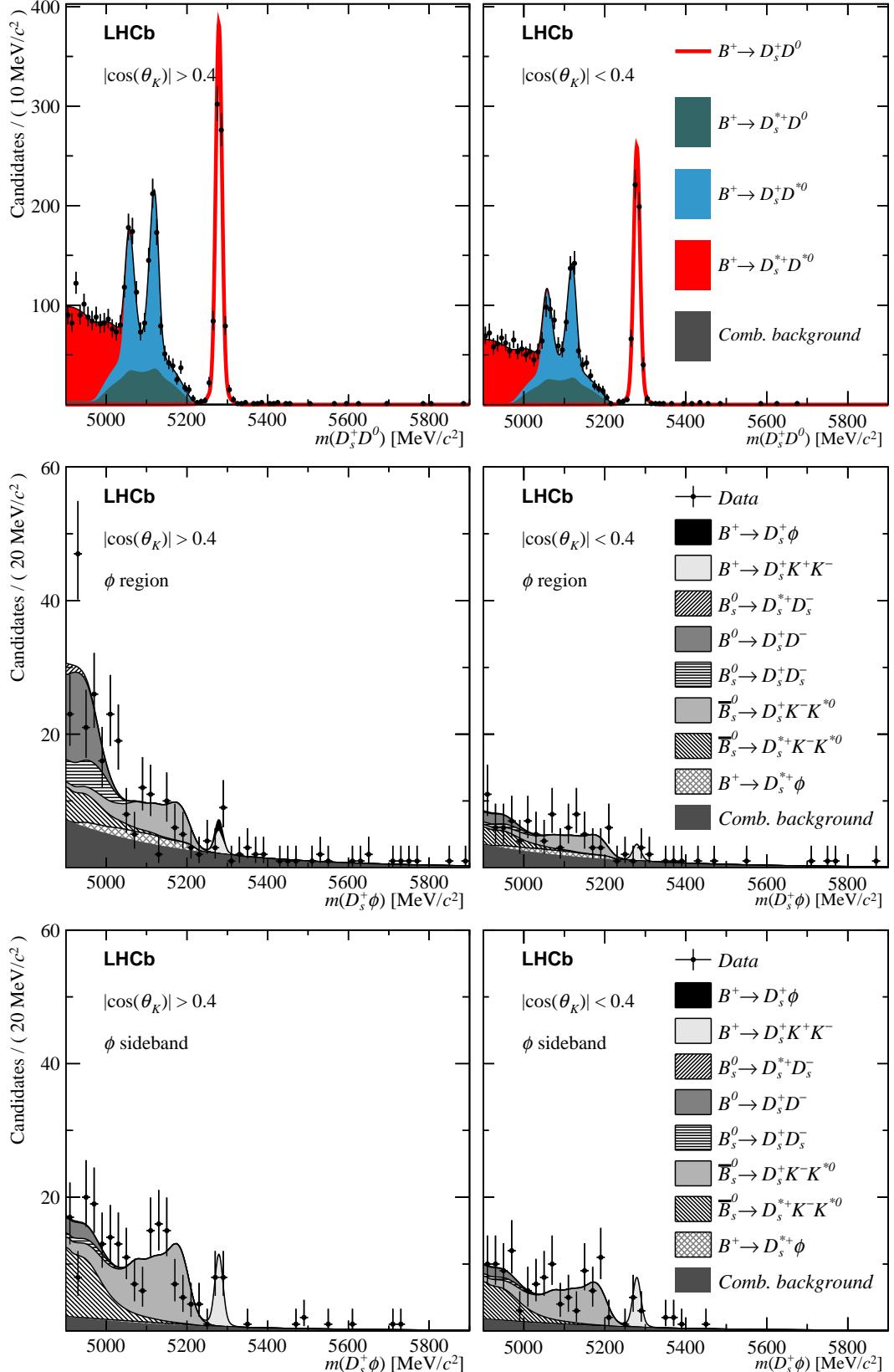


Figure A.3: Invariant mass fits to  $B^+ \rightarrow D_s^+ \phi$  candidates with  $D_s^+ \rightarrow K^+ K^- \pi^+$  (excluding  $D_s^+ \rightarrow \phi \pi^+$ ).

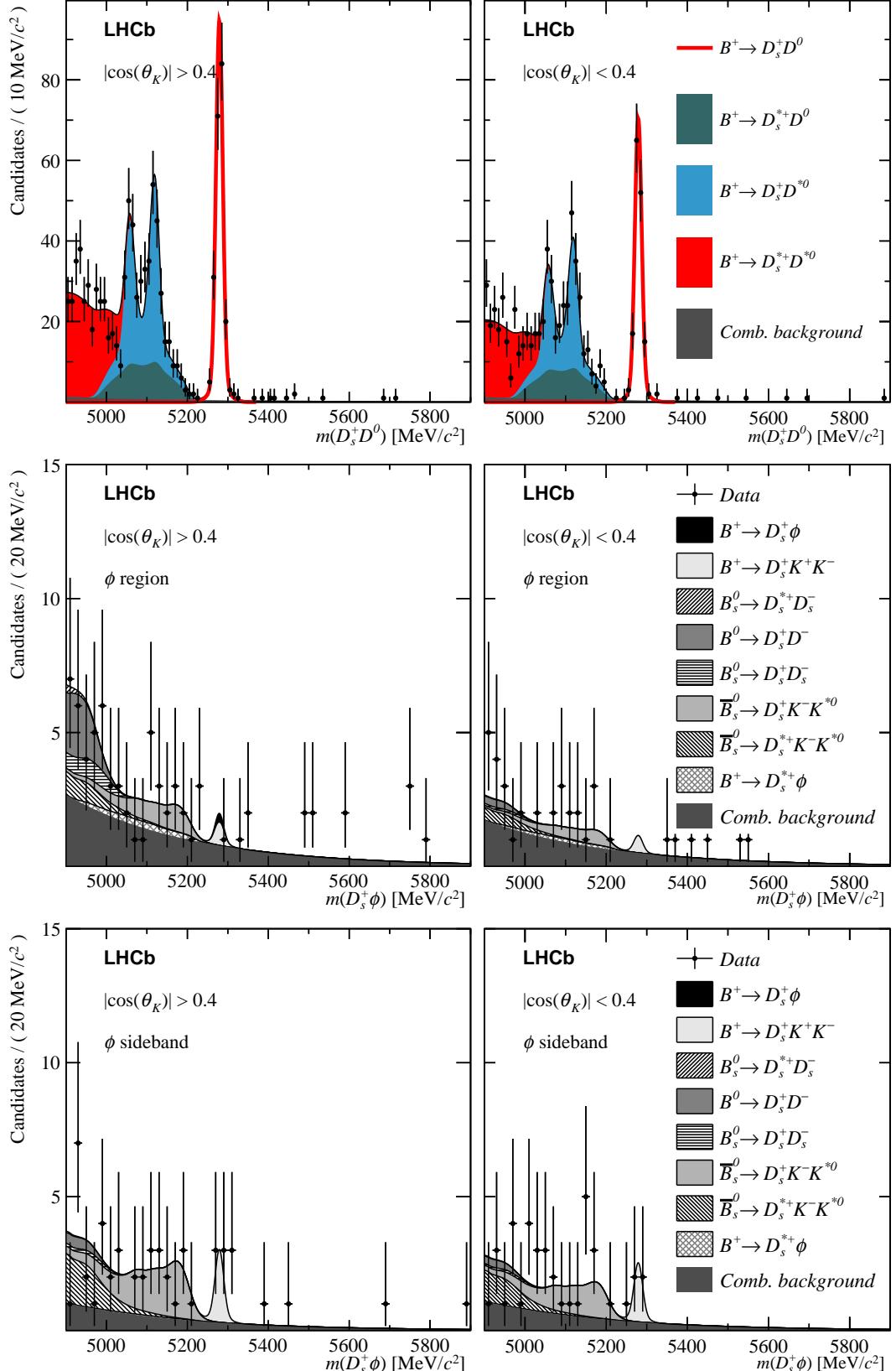


Figure A.4: Invariant mass fits to  $B^+ \rightarrow D_s^+ \phi$  candidates with  $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ .

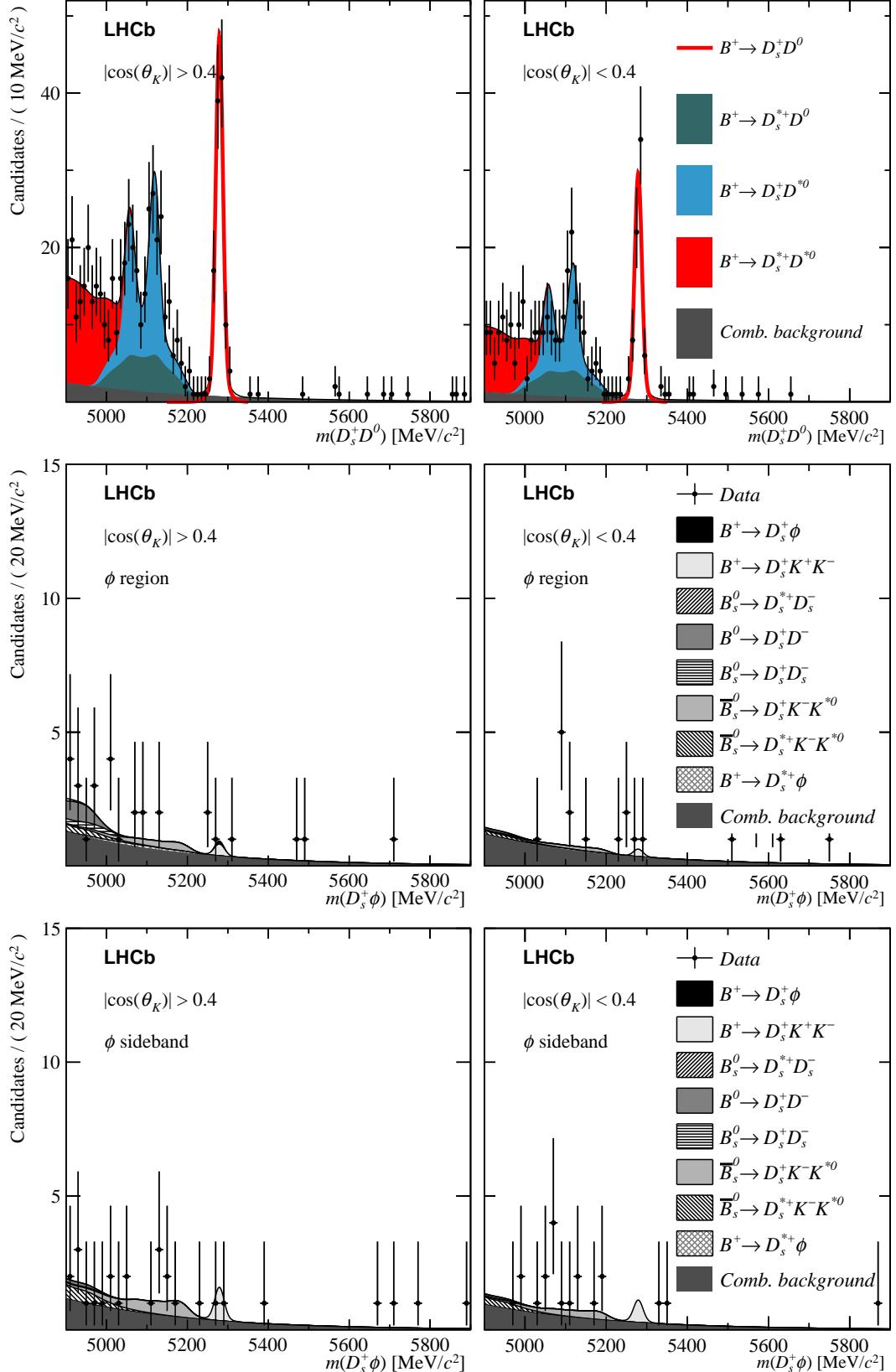


Figure A.5: Invariant mass fits to  $B^+ \rightarrow D_s^+ \phi$  candidates with  $D_s^+ \rightarrow K^+ \pi^- \pi^+$ .

# Bibliography

- [1] J. J. Thomson, *Cathode rays*, The Electrician **39** (1897) 104, (Friday evening meeting of the Royal Institution, 30 April 1897).
- [2] E. Rutherford, *The scattering of alpha and beta particles by matter and the structure of the atom*, Phil. Mag. Ser. 6 **21** (1911) 669.
- [3] J. Chadwick, *Possible Existence of a Neutron*, Nature **129** (1932) 312.
- [4] H. Yukawa, *On the Interaction of Elementary Particles. I*, Proceedings of the Physico-Mathematical Society of Japan. 3rd Series **17** (1935) 48.
- [5] C. M. G. Lattes, H. Muirhead, G. P. S. Occhialini, and C. F. Powell, *Processes involving charged mesons*, Nature **159** (1947) 694, [42(1947)].
- [6] C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell, *Observations on the Tracks of Slow Mesons in Photographic Emulsions. 1*, Nature **160** (1947) 453, [99(1947)].
- [7] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.
- [8] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, in *Developments in the quark theory of hadrons. VOL. 1. 1964 - 1978* (D. B. Lichtenberg and S. P. Rosen, eds.), pp. 22–101. 1964.
- [9] Belle Collaboration, S.-K. Choi *et al.*, *Observation of a Narrow Charmoniumlike State in Exclusive  $B^\pm \rightarrow K^\pm \pi^+ \pi^- J/\psi$  Decays*, Phys. Rev. Lett. **91** (2003) 262001.

- [10] LHCb Collaboration, R. Aaij *et al.*, *Observation of  $J/\psi p$  Resonances Consistent with Pentaquark States in  $\Lambda_b^0 \rightarrow J/\psi K^- p$  Decays*, Phys. Rev. Lett. **115** (2015) 072001.
- [11] J.-E. Augustin *et al.*, *Discovery of a Narrow Resonance in  $e^+e^-$  Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.
- [12] J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle  $J$* , Phys. Rev. Lett. **33** (1974) 1404.
- [13] M. L. Perl *et al.*, *Evidence for Anomalous Lepton Production in  $e^+ - e^-$  Annihilation*, Phys. Rev. Lett. **35** (1975) 1489.
- [14] S. W. Herb *et al.*, *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.
- [15] G. Arnison *et al.*, *Experimental observation of isolated large transverse energy electrons with associated missing energy at  $s=540$  GeV*, Physics Letters B **122** (1983), no. 1 103 .
- [16] M. Banner *et al.*, *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider*, Physics Letters B **122** (1983), no. 5 476 .
- [17] CDF Collaboration, F. Abe *et al.*, *Observation of Top Quark Production in  $\bar{p}p$  Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626.
- [18] D0 Collaboration, S. Abachi *et al.*, *Search for High Mass Top Quark Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV*, Phys. Rev. Lett. **74** (1995) 2422.
- [19] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012), no. 1 1 .

- [20] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012), no. 1 30 .
- [21] LHCb collaboration, R. Aaij *et al.*, *First observation of  $B^+ \rightarrow D_s^+ K^+ K^-$  decays and a search for  $B^+ \rightarrow D_s^+ \phi$  decays*, JHEP **01** (2018) 110, [arXiv:1711.05637](https://arxiv.org/abs/1711.05637).
- [22] R. P. Feynman, *The Theory of Positrons*, Phys. Rev. **76** (1949) 749.
- [23] *The quantum theory of the electron*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **117** (1928), no. 778 610, [arXiv:<http://rspa.royalsocietypublishing.org/content/117/778/610.full.pdf>](http://rspa.royalsocietypublishing.org/content/117/778/610.full.pdf).
- [24] C. D. Anderson, *The positive electron*, Phys. Rev. **43** (1933) 491.
- [25] ATLAS collaboration, M. Aaboud *et al.*, *Evidence for light-by-light scattering in heavy-ion collisions with the ATLAS detector at the LHC*, Nature Phys. **13** (2017), no. 9 852, [arXiv:1702.01625](https://arxiv.org/abs/1702.01625).
- [26] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [27] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [28] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.
- [29] D. Hanneke, S. Fogwell, and G. Gabrielse, *New Measurement of the Electron Magnetic Moment and the Fine Structure Constant*, Phys. Rev. Lett. **100** (2008) 120801.
- [30] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, *Erratum: Tenth-order electron anomalous magnetic moment: Contribution of diagrams without closed lepton loops [Phys. Rev. D 91, 033006 (2015)]*, Phys. Rev. D **96** (2017) 019901.

- [31] *Precision electroweak measurements on the Z resonance*, Physics Reports **427** (2006), no. 5 257 .
- [32] P. von Ballmoos, *Antimatter in the Universe: constraints from gamma-ray astrometry*, Hyperfine Interactions **228** (2014) 91.
- [33] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,no.5,61(1991)].
- [34] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with  $B^0 \rightarrow K^{*0}\ell^+\ell^-$  decays*, JHEP **08** (2017) 055, [arXiv:1705.05802](https://arxiv.org/abs/1705.05802).
- [35] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using  $B^+ \rightarrow K^+\ell^+\ell^-$  decays*, Phys. Rev. Lett. **113** (2014) 151601, [arXiv:1406.6482](https://arxiv.org/abs/1406.6482).
- [36] Y. Fukuda *et al.*, *Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days*, Phys. Rev. Lett. **81** (1998) 1158.
- [37] SNO Collaboration, Q. R. Ahmad *et al.*, *Measurement of the Rate of  $\nu_e + d \rightarrow p + p + e^-$  Interactions Produced by  ${}^8B$  Solar Neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301.
- [38] E. Corbelli and P. Salucci, *The extended rotation curve and the dark matter halo of M33*, Monthly Notices of the Royal Astronomical Society **311** (2000), no. 2 441.
- [39] D. Clowe, A. Gonzalez, and M. Markevitch, *Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657558: Direct Evidence for the Existence of Dark Matter*, The Astrophysical Journal **604** (2004), no. 2 596.
- [40] M. Markevitch *et al.*, *Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 065756*, The Astrophysical Journal **606** (2004), no. 2 819.

- [41] J. Kuckei *et al.*, *Strong CP violation and the neutron electric dipole form factor*, Physics of Atomic Nuclei **70** (2007) 349.
- [42] Y. Nambu, *Quasi-Particles and Gauge Invariance in the Theory of Superconductivity*, Phys. Rev. **117** (1960) 648.
- [43] J. Goldstone, *Field theories with superconductor solutions*, Il Nuovo Cimento (1955-1965) **19** (1961) 154.
- [44] A. Crivellin, D. Müller, A. Signer, and Y. Ulrich, *Correlating lepton flavor universality violation in B decays with  $\mu \rightarrow e\gamma$  using leptoquarks*, Phys. Rev. D **97** (2018) 015019.
- [45] T. Kaluza, *Zum Unitätsproblem der Physik*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys. ) **1921** (1921) 966.
- [46] O. Klein, *Quantentheorie und fünfdimensionale relativitätstheorie*, Zeitschrift für Physik **37** (1926) 895.
- [47] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **38** (1977) 1440.
- [48] S. Behrends *et al.*, *Observation of Exclusive Decay Modes of b-Flavored Mesons*, Phys. Rev. Lett. **50** (1983) 881.
- [49] C. S. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413.
- [50] S. L. Glashow, *The renormalizability of vector meson interactions*, Nucl. Phys. **10** (1959) 107.
- [51] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.
- [52] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

- [53] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.
- [54] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.
- [55] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973), no. 2 652.
- [56] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.
- [57] 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** (1964) 138.
- [58] P. del Amo Sanchez *et al.*, *Measurement of CP observables in  $B^\pm \rightarrow D_{CP}K^\pm$  decays and constraints on the CKM angle  $\gamma$* , Phys. Rev. D **82** (2010) 072004.
- [59] Belle Collaboration, A. Poluektov *et al.*, *Evidence for direct CP violation in the decay  $B^\pm \rightarrow D^{(*)}K^\pm$ ,  $D \rightarrow K_S^0\pi^+\pi^-$  and measurement of the CKM phase  $\phi_3$* , Phys. Rev. D **81** (2010) 112002.
- [60] BABAR Collaboration, B. Aubert *et al.*, *Observation of CP Violation in the  $B^0$  Meson System*, Phys. Rev. Lett. **87** (2001) 091801.
- [61] Belle Collaboration, K. Abe *et al.*, *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87** (2001) 091802.
- [62] LHCb Collaboration, R. Aaij *et al.*, *First observation of cp violation in the decays of  $B_s^0$  mesons*, Phys. Rev. Lett. **110** (2013) 221601.
- [63] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343.
- [64] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346.

- [65] Belle Collaboration, B. Kronenbitter *et al.*, *Measurement of the branching fraction of  $B^+ \rightarrow \tau^+ \nu_\tau$  decays with the semileptonic tagging method*, Phys. Rev. D **92** (2015) 051102.
- [66] BABAR Collaboration, B. Aubert *et al.*, *Search for the rare leptonic decays  $B^+ \rightarrow l^+ \nu_l$  ( $l = e, \mu$ )*, Phys. Rev. D **79** (2009) 091101.
- [67] Belle Collaboration, N. Satoyama *et al.*, *A search for the rare leptonic decays  $B^+ \rightarrow \mu^+ \nu_\mu$  and  $B^+ \rightarrow e^+ \nu_e$* , Physics Letters B **647** (2007), no. 2 67 .
- [68] LHCb Collaboration, R. Aaij *et al.*, *Observation of the Annihilation Decay Mode  $B^0 \rightarrow K^+ K^-$* , Phys. Rev. Lett. **118** (2017) 081801.
- [69] LHCb collaboration, R. Aaij *et al.*, *Observation of  $B_c^+ \rightarrow D^0 K^+$  decays*, Phys. Rev. Lett. **118** (2017) 111803, [arXiv:1701.01856](https://arxiv.org/abs/1701.01856).
- [70] LHCb collaboration, R. Aaij *et al.*, *First observation of the rare  $B^+ \rightarrow D^+ K^+ \pi^-$  decay*, Phys. Rev. D **93** (2016) 051101(R), Erratum ibid. **D93** (2016) 119902, [arXiv:1512.02494](https://arxiv.org/abs/1512.02494).
- [71] S. Descotes-Genon, J. He, E. Kou, and P. Robbe, *Nonleptonic charmless  $B_c$  decays and their search at LHCb*, Phys. Rev. D **80** (2009) 114031.
- [72] M. Gronau, D. London, and J. L. Rosner, *Rescattering contributions to rare  $B$ -meson decays*, Phys. Rev. D **87** (2013) 036008, [arXiv:1211.5785](https://arxiv.org/abs/1211.5785).
- [73] S. Okubo,  *$\phi$ -meson and unitary symmetry model*, Physics Letters **5** (1963), no. 2 165 .
- [74] M. Gronau and J. L. Rosner,  *$\omega - \phi$  mixing and weak annihilation in  $D_s$  decays*, Phys. Rev. D **79** (2009) 074006.
- [75] HPQCD collaboration, R. J. Dowdall *et al.*,  *$B$ -meson decay constants from improved lattice nonrelativistic QCD and physical  $u, d, s$  and  $c$  quarks*, Phys. Rev. Lett. **110** (2013) 222003, [arXiv:1302.2644](https://arxiv.org/abs/1302.2644).

- [76] ETM collaboration, A. Bussone *et al.*, *Mass of the b quark and B-meson decay constants from  $N_f = 2+1+1$  twisted-mass lattice QCD*, Phys. Rev. **D93** (2016) 114505, [arXiv:1603.04306](https://arxiv.org/abs/1603.04306).
- [77] Fermilab Lattice and MILC collaborations, A. Bazavov *et al.*,  *$B_{(s)}^0$ -mixing matrix elements from lattice QCD for the Standard Model and beyond*, Phys. Rev. **D93** (2016) 113016, [arXiv:1602.03560](https://arxiv.org/abs/1602.03560).
- [78] H. Zou, R.-H. Li, X.-X. Wang, and C.-D. Lu, *The CKM suppressed  $B(B_s) \rightarrow \bar{D}_{(s)}P, \bar{D}_{(s)}V, \bar{D}_{(s)}^*P, \bar{D}_{(s)}^*V$  decays in perturbative QCD approach*, J. Phys. **G37** (2010) 015002, [arXiv:0908.1856](https://arxiv.org/abs/0908.1856).
- [79] R. Mohanta, *Searching for new physics in the rare decay  $B^+ \rightarrow D_s^+\phi$* , Phys. Lett. **B540** (2002) 241, [arXiv:hep-ph/0205297](https://arxiv.org/abs/hep-ph/0205297).
- [80] R. Mohanta and A. K. Giri, *Possible signatures of unparticles in rare annihilation type B decays*, Phys. Rev. **D76** (2007) 057701, [arXiv:0707.3308](https://arxiv.org/abs/0707.3308).
- [81] C.-D. Lu, *Calculation of pure annihilation type decay  $B^+ \rightarrow D_s^+\phi$* , Eur. Phys. J. **C24** (2002) 121, [arXiv:hep-ph/0112127](https://arxiv.org/abs/hep-ph/0112127).
- [82] H. Georgi, *Another odd thing about unparticle physics*, Physics Letters B **650** (2007), no. 4 275 .
- [83] LHCb collaboration, R. Aaij *et al.*, *First evidence for the annihilation decay mode  $B^+ \rightarrow D_s^+\phi$* , JHEP **02** (2013) 043, [arXiv:1210.1089](https://arxiv.org/abs/1210.1089).
- [84] R. H. Dalitz, *Decay of  $\tau$  Mesons of Known Charge*, Phys. Rev. **94** (1954) 1046.
- [85] BaBar collaboration, B. Aubert *et al.*, *Evidence for the rare decay  $B^+ \rightarrow D_s^+\pi^0$* , Phys. Rev. Lett. **98** (2007) 171801, [arXiv:hep-ex/0611030](https://arxiv.org/abs/hep-ex/0611030).
- [86] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

- [87] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, [arXiv:1405.7808](https://arxiv.org/abs/1405.7808).
- [88] G. Haefeli *et al.*, *The LHCb DAQ interface board TELL1*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **560** (2006), no. 2 494 .
- [89] LHCb VELO, ST, K. Rinnert, *LHCb silicon detectors: the Run 1 to Run 2 transition and first experience of Run 2*, PoS **VERTEX2015** (2015) 004.
- [90] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](https://arxiv.org/abs/1412.6352).
- [91] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, [arXiv:1311.3893](https://arxiv.org/abs/1311.3893).
- [92] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](https://arxiv.org/abs/1211.6759).
- [93] A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **876** (2017) 221 , The 9th international workshop on Ring Imaging Cherenkov Detectors (RICH2016).
- [94] R. V. Gomez and the LHCb collaboration, *Commissioning of the Scintillator Pad Detector of LHCb with cosmic rays and first LHC collisions*, Journal of Physics: Conference Series **293** (2011), no. 1 012059.
- [95] I. Machikhiliyan, *The calibration of an LHCb electromagnetic calorimeter by recovering the invariant mass of neutral pions*, Moscow University Physics Bulletin **68** (2013), no. 5 360.
- [96] Y. Guz, *The LHCb Calorimeter system: design, performance and upgrade*, Journal of Instrumentation **12** (2017), no. 07 C07024.

- [97] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Journal of Basic Engineering **82** (1960) 35.
- [98] R. Frhwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 2 444 .
- [99] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The LHCb Turbo Stream*, Journal of Physics: Conference Series **664** (2015), no. 8 082004.
- [100] The LHCb collaboration, *Precision luminosity measurements at LHCb*, Journal of Instrumentation **9** (2014), no. 12 P12005.
- [101] S. van der Meer, *Calibration of the effective beam height in the ISR*, CERN-ISR-PO68-31.
- [102] *Transverse Emittance Measurements Using LHCb's Beam-Gas Interactions*, no. 8 in International Particle Accelerator Conference, (Geneva, Switzerland), JACoW, May, 2017. doi: 10.18429/JACoW-IPAC2017-MOPAB131.
- [103] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.
- [104] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.
- [105] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [106] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [107] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.

- [108] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [109] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [110] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, [arXiv:1210.6861](https://arxiv.org/abs/1210.6861).
- [111] LHCb collaboration, R. Aaij *et al.*, *First observations of  $\bar{B}_s^0 \rightarrow D^+ D^-$ ,  $D_s^+ D^-$  and  $D^0 \bar{D}^0$  decays*, Phys. Rev. **D87** (2013) 092007, [arXiv:1302.5854](https://arxiv.org/abs/1302.5854).
- [112] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](https://arxiv.org/abs/physics/0402083).
- [113] R. Brun and F. Rademakers, *ROOT An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** (1997), no. 1 81 , New Computing Techniques in Physics Research V.
- [114] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- [115] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. [arXiv:physics/0308063](https://arxiv.org/abs/physics/0308063).
- [116] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, [arXiv:physics/0503191](https://arxiv.org/abs/physics/0503191).
- [117] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, [arXiv:physics/0306116](https://arxiv.org/abs/physics/0306116).

- [118] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP observables in  $B^\pm \rightarrow D^{(*)}K^\pm$  and  $B^\pm \rightarrow D^{(*)}\pi^\pm$  decays*, Phys. Lett. **B777** (2017) 16, [arXiv:1708.06370](https://arxiv.org/abs/1708.06370).
- [119] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, [arXiv:hep-ex/0011057](https://arxiv.org/abs/hep-ex/0011057).
- [120] B. Sen, M. Walker, and M. Woodroffe, *On the unified method with nuisance parameters*, Statistica Sinica **19** (2009) 301.
- [121] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.
- [122] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, Journal of Instrumentation **10** (2015), no. 02 P02011, [arXiv:1411.5528](https://arxiv.org/abs/1411.5528).
- [123] J. Back *et al.*, *Laura++ : a Dalitz plot fitter*, [arXiv:1711.09854](https://arxiv.org/abs/1711.09854).
- [124] G. Breit and E. Wigner, *Capture of Slow Neutrons*, Phys. Rev. **49** (1936) 519.
- [125] S. M. Flatt, *Coupled-channel analysis of the  $\pi\eta$  and  $KK$  systems near  $KK$  threshold*, Physics Letters B **63** (1976), no. 2 224 .
- [126] L. Moneta *et al.*, *The RooStats Project*, [arXiv:1009.1003](https://arxiv.org/abs/1009.1003).
- [127] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011) 1554.
- [128] G. J. Feldman and R. D. Cousins, *A unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, [arXiv:physics/9711021](https://arxiv.org/abs/physics/9711021).
- [129] LHCb collaboration, R. Aaij *et al.*, *First evidence for the annihilation decay mode  $B^+ \rightarrow D_s^+\phi$* , JHEP **02** (2013) 043, [arXiv:1210.1089](https://arxiv.org/abs/1210.1089).