Doubly charmed B decays with the LHCb experiment

Alison Maria Tully

of

Peterhouse

University of Cambridge

This dissertation is submitted to the University of Cambridge
for the degree of Doctor of Philosophy

June 2019

Abstract

CP violation is a necessary ingredient to create a matter-antimatter asymmetry. However, the amount of *CP* violation incorporated in the Standard Model is orders of magnitude too small to explain the observed asymmetry. This thesis documents two studies of *CP* violation with doubly charmed B decays. Both were performed using Run I data collected by the LHCb experiment in pp collisions at the LHC. This corresponds to $1.0\,\mathrm{fb}^{-1}$ collected at a centre-of-mass energy of 7 TeV and $2.0\,\mathrm{fb}^{-1}$ collected at a centre-of-mass energy of 8 TeV.

The first analysis is a search for B_c^+ decays to two charm mesons, $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$, where the second charm meson is either D^0 or $\overline{D}^0$, which are sensitive to the CKM angle γ. No evidence for signal is found and limits are set on twelve decay modes. The second analysis is the measurement of the *CP* asymmetry in $B^+ \to D_{(s)}^+ \overline{D}^0$ decays. The results are

$$\mathcal{A}^{CP}(B^+ \to D_s^+ \overline{D}^0) = (-0.4 \pm 0.5 \pm 0.5)\%,$$
$$\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0) = (+2.3 \pm 2.7 \pm 0.4)\%,$$

where the first uncertainty is statistical and the second systematic. This is the first measurement of $\mathcal{A}^{CP}(B^+ \to D_s^+ \overline{D}^0)$ and the most precise measurement of $\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0)$ to date. The results are consistent with *CP* symmetry.

Declaration

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

Alison Maria Tully

Acknowledgements

Completing a PhD, particularly as a member of a large collaboration such as LHCb, is never just the work of one, but uncountably many. First, I would like to thank my supervisor, Val Gibson, for giving me the opportunity to study at Cambridge University and for providing me with numerous helpful suggestions throughout my PhD. I owe a huge debt of gratitude to Rolf Oldeman, without whose patient guidance and seemingly endless knowledge of particle physics the work in this thesis would not have been completed.

The Cambridge LHCb group has been a stimulating environment to work in for the past three-and-a-half years. The feedback provided during group meetings improved the results presented in this thesis immeasurably. I am particularly thankful to those I had the pleasure of working with directly: Matt Kenzie, Blaise Delaney, Ana Trišović, Susan Haines, and Chris Jones. The other members of the Cambridge HEP group, particularly those in the infamously-named WhatsApp group, made lunches, coffee breaks and formals most enjoyable. Special thanks to Will Fawcett for the brief stint we had as lifting partners and for voluntarily reading parts of this thesis.

Throughout my PhD, I have been a member of the B to open charm physics working group within LHCb. I'm grateful for the insightful questions asked on my talks in working group meetings and the helpful comments received on the documentation of the analyses. For two years, I oversaw the monstrous B2OC stripping code alongside Alessandro Bertolin. Thanks to Alessandro, this was (relatively) painless.

I had the good fortune to spend a year-and-a-half of my PhD based at CERN. The LTA community (quite literally) made it feel like a home away from home. Special thanks go to James Cowley, Mike Nelson, Amal Vaidya and Matt Tilley; your boundless enthusiasm for after-work beers in R1 was inspiring. I hope we continue to make Oxbridge pilgrimages for May Week until well after we're too old to be there.

I would like to thank the Science and Technology Facilities Council for not only funding my PhD, but also a brief interlude into the world of policy via a three month long fellowship at the Parliamentary Office of Science and Technology.[1] Thanks also to the best POST supervisor, Lorna Christie, and fellow party cluster inhabitant, Rowena Bermingham, for enabling my caffeine addiction and inspiring a new-found love for bagels.

Thank you to my college, Peterhouse, for providing such a welcoming home for the two years I spent in Cambridge and for awarding me generous grants to tour the East Coast of the United States and attend the CERN Latin-American school of HEP in Argentina. Special thanks to my eclectic college squad (Jo Neville, Michael Waters, James Sinclair and Mohamed Derbal), for providing excellent dinner company and countless golden photography opportunities.

To my parents, Brenda and Richard, thank you for your support and encouragement over the years. Escaping to Glasgow during the holidays was always a welcome break. Thanks to the rest of my friends in Cambridge and London. To name just a few: Mel Singh, for always making me laugh and for introducing me to a passion of mine: climbing; Matteo Sbroscia, for the muffins and bike repairs; Mandy Choi, for the terrible complements and always providing a place to crash; and Jasmine Walter, for everything.

[1] https://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-0584

Preface

This thesis documents my contribution to finding an answer to one of the most fundamental questions in physics: *why is there an asymmetry between matter and antimatter*? The key to this question could lie in the study of *CP* violation. Doubly charmed B decays provide a rich laboratory for the study of *CP* violation due to the potential for large effects, both within and beyond the Standard Model. Two analyses of these decays are documented: the search for B_c^+ decays to two charm mesons and the measurement of the *CP* asymmetry in $B^+ \to D_{(s)}^+ \overline{D}^0$ decays. The results have been published in references [1] and [2], respectively.

A brief introduction to the Standard Model and its shortcomings is given in **Chapter 1**. **Chapter 2** describes the theoretical motivation for the study of doubly charmed B decays, namely the sensitivity to the CKM angle γ that arises in the interference of $B_c^+ \to D_{(s)}^{(*)+} \overline{D}^{(*)0}$ and $B_c^+ \to D_{(s)}^{(*)+} D^{(*)0}$ decays and the new physics effects that can greatly enhance the *CP* asymmetry in $B^+ \to D_{(s)}^+ \overline{D}^0$ decays. The analyses are based on data collected by the LHCb experiment, an overview of which is given in **Chapter 3**. The selection that has been developed to identify the most signal-like events is documented in **Chapter 4**. Most of the selection requirements are shared between B_c^+ and B^+ decays due to similarities between the two. **Chapter 5** describes the search performed for the B_c^+ decays and **Chapter 6** the measurement of the *CP* asymmetry in $B^+ \to D_{(s)}^+ \overline{D}^0$ decays. Finally, the results are summarised in **Chapter 7**.

Throughout this dissertation charge conjugates of each decay mode are implied, except where explicitly stated. For example, $B_c^+ \to D_s^+ \overline{D}^0$ refers to both $B_c^+ \to D_s^+ \overline{D}^0$ and $B_c^- \to D_s^- D^0$.

Contents

"A scientist in his laboratory is not a mere technician: he is also a child confronting natural phenomena that impress him as though they were fairy tales."

— Marie Skłodowska-Curie

Chapter 1

Introduction

“Not only is the Universe stranger than we think, it is stranger than we can think.”
— Werner Heisenberg

The rather unfairly named “Standard Model of particle physics” (SM) is really quite remarkable. It describes the basic building blocks from which everything in the universe is made (the “fundamental particles”) and their interactions under the electromagnetic, weak and strong forces. The SM contains twelve matter particles of spin-1/2, known as fermions, four force particles of spin-1, known as gauge bosons, and the Higgs boson, which is related to the mechanism that grants particles mass. The fermions are made up of six flavours of quarks (u, d, c, s, t, b), which interact under all three forces, and leptons $(e^-, \nu_e, \mu^-, \nu_\mu, \tau^-, \nu_\tau)$, which do not interact under the strong force. The photon mediates the electromagnetic interaction, the W and Z bosons the weak interaction and the gluon the strong interaction. Each of the particles in the SM has an antiparticle counterpart, which, aside from charge, is otherwise identical.[1] The particle zoo is summarised in table 1.1.

The formation of the SM in the latter half of the 20th century was the combined brain power of a myriad of physicists across the globe, 55 of whom were awarded the Nobel Prize in Physics. Every particle the SM has predicted to exist has been discovered, culminating with the discovery of the Higgs boson in 2012 [4, 5] after a 40 year-long search effort. The precision to which some of its predictions have been verified is nothing short of astonishing. For example, the prediction of the anomalous magnetic moment of

[1]Some particles, such as the photon, are their own antiparticle.

Particle type	Name	Symbol	Mass (MeV/c^2)	Spin	Charge (e)
Quarks	Up	u	$2.16^{+0.49}_{-0.26}$	$\frac{1}{2}$	$+\frac{2}{3}$
	Down	d	$4.67^{+0.48}_{-0.17}$	$\frac{1}{2}$	$-\frac{1}{3}$
	Charm	c	$1,270 \pm 20$	$\frac{1}{2}$	$+\frac{2}{3}$
	Strange	s	93^{+11}_{-5}	$\frac{1}{2}$	$-\frac{1}{3}$
	Top	t	$172,900 \pm 400$	$\frac{1}{2}$	$+\frac{2}{3}$
	Bottom	b	$4,180^{+30}_{-20}$	$\frac{1}{2}$	$-\frac{1}{3}$
Leptons	Electron	e^-	$0.5109989461 \pm 0.0000000031$	$\frac{1}{2}$	-1
	Electron neutrino	ν_e	< 0.000002	$\frac{1}{2}$	0
	Muon	μ^-	$105.6583745 \pm 0.0000024$	$\frac{1}{2}$	-1
	Muon neutrino	ν_μ	< 0.19	$\frac{1}{2}$	0
	Tau	τ^-	1776.86 ± 0.12	$\frac{1}{2}$	-1
	Tau neutrino	ν_τ	< 18.2	$\frac{1}{2}$	0
Bosons	Photon	γ	0	1	0
	W boson	$W^\pm$	$80,379 \pm 12$	1	± 1
	Z boson	Z	$91,187.6 \pm 2.1$	1	0
	Gluon	g	0	1	0
	Higgs boson	H^0	$125,100 \pm 140$	0	0

Table 1.1: The mass, spin and electric charge of the particles in the SM [3].

the electron agrees with the experimental measurement to about one part in a trillion [6]. With the operation of the Large Hadron Collider (LHC), the SM has been placed under an ever more powerful microscope. A task force of thousands of physicists at the LHC experiments have performed a comprehensive suite of measurements probing various corners of the theory, but so far, it has proved impregnable.

Although it may seem like the career prospects of a budding particle physicist look rather bleak, despite its success, the SM leaves many important phenomena unexplained. Of the four fundamental forces (electromagnetic, weak, strong and gravity), it incorporates all but gravity. Although this isn't a problem for the very light particles collided at the LHC, there is a disconnect between the description of physics on very small, atomic scales (quantum field theory) and very large, astronomical scales (Einstein's theory of General Relativity).

Of the total energy density of the universe, the particles in the SM only make up a mere $\sim 5\%$. [7]. Of the rest, $\sim 25\%$ is "dark matter", a strange form of matter that interacts gravitationally but is invisible when viewed through a telescope, and $\sim 70\%$ is "dark energy", a mysterious form of energy with anti-gravitational properties, thought to be responsible for the accelerating expansion of the universe [8]. The SM does not contain any dark matter candidates, and attempts to describe dark energy in terms of the vacuum energy of quantum field theory lead to a discrepancy of around 120 orders of magnitude.[2]

If the Big Bang had produced equal amounts of matter and antimatter, the two would annihilate in a burst of light, and planets, stars and galaxies would simply not exist. The clear evidence to the contrary indicates the universe's favouritism towards matter. One of the necessary ingredients for matter to win out over antimatter is the violation of charge-parity (*CP*) symmetry [10]. Although the SM incorporates *CP* violation, the excess of matter it generates falls many orders of magnitude short of the observed asymmetry [11].

Due to these shortcomings, the SM is widely regarded as a low-energy approximation of a more fundamental theory. However, as the SM is so frustratingly successful, the form this theory takes is unknown. A feature that is required in order to explain the matter-antimatter asymmetry is new sources of *CP* violation. As *CP* violation is small in the SM, new physics contributions may be able to stand on an equal footing. The study of *CP* violation therefore represents one of the most promising routes to the discovery of new physics.

One of the primary goals of the LHCb experiment at the LHC is to make precise measurements of *CP*-violating observables in the decays of particles containing heavy quarks. The decays of charged B mesons,[3] B^+ or B_c^+, to two charm mesons provide a particularly rich laboratory for the study of *CP* violation due to the potential for large effects. Moreover, the origin of these effects differs between the two B meson species, with the B_c^+ decays expected to have large SM contributions [12–17], while the B^+ decays are sensitive to new physics [18, 19]. The study of these decays therefore provides a two-pronged probe into the nature of *CP* violation.

While the B^+ decays to two charm mesons are experimentally well-established, the B_c^+ decays have never before been observed. The study of B_c^+ decays is challenging for several reasons. First, the B_c^+ meson is composed of two heavy quarks, so its production is much rarer than that of the B^+ meson. Second, its lifetime is short, resulting in a lower reconstruction efficiency. Third, the branching fractions of the relevant decays are

[2]This is known as the cosmological constant problem and has been described as the "worst theoretical prediction in the history of physics" [9].

[3]Mesons are bound states of quark-antiquark pairs. Three-quark combinations are known as baryons.

Meson	Quark content	Mass (MeV/c^2)	Lifetime (ps)
B^+	$\bar{b}u$	5279.33 ± 0.13	1.638 ± 0.004
B^0	$\bar{b}d$	5279.64 ± 0.13	1.519 ± 0.004
B_c^+	$\bar{b}c$	6274.9 ± 0.8	0.510 ± 0.009
B_s^0	$\bar{b}s$	5366.88 ± 0.17	1.527 ± 0.011

Table 1.2: The quark content, mass and lifetime of the different B mesons [3].

predicted to be two orders of magnitude smaller for B_c^+ than for B^+ [16, 17, 20, 21]. The properties of the different B mesons are summarised in table 1.2.

Chapter 2

Theoretical motivation

"I like to say that while antimatter may seem strange, it is strange in the sense that Belgians are strange. They are not really strange; it is just that one rarely meets them."
— Lawrence M. Krauss

CP symmetry implies that if a particle is replaced by its antiparticle (*C* symmetry) and its spatial coordinates are inverted *i.e.* $(t, x, y, z) \to (t, -x, -y, -z)$ (*P* symmetry) then the physics will remain unchanged. The violation of this symmetry is known as *CP* violation. In this chapter, the role of *CP* violation in the generation of a baryon asymmetry is described (section 2.1), followed by a discussion of *CP* violation in the SM (section 2.2). Measurements of the CKM angle γ are then detailed (section 2.3) and lastly, the measurement of the *CP* asymmetry in $B^+ \to D^+_{(s)}\overline{D}^0$ decays is motivated (section 2.4).

2.1 Baryon asymmetry

The baryon asymmetry is the observed dominance of matter over antimatter in the universe. As there is no known explanation for this imbalance, the baryon asymmetry is one of the great mysteries of physics. In this section, antimatter is introduced (section 2.1.1), followed by a discussion of the mechanism responsible for generating a baryon asymmetry in the early universe (section 2.1.2). The conditions necessary for baryogenesis to occur are described (section 2.1.2.1), followed by a review of some of the models that have been proposed, both within (section 2.1.2.2) and beyond (section 2.1.2.3) the SM and finally, the experimental prospects are discussed (section 2.1.2.4).

2.1.1 Antimatter

The father of antimatter is English physicist and renowned eccentric, Paul Dirac. Dirac elegantly unified the theories of quantum mechanics and special relativity in 1928 with his eponymous equation [22], for which he received the Nobel Prize in Physics just 5 years later.[1] In natural units, the equation takes the form,

$$(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0, \tag{2.1}$$

where ψ is the wave function of a spin-1/2 fermion of mass m, ∂_{μ} is the 4-gradient and γ^{μ} are 4×4 unitary matrices. Repeated indices indicate summation.

Although ironically not fully realised by Dirac himself, this equation postulates the existence of antimatter. This is due to a peculiar feature, whereby for each solution with a positive energy E, there is a corresponding, equally valid, solution with negative energy, $-E$. Since there is no upper limit to the energy an electron can have, then correspondingly, there would be no lower limit and electrons would continuously shed photons as they descend into increasingly negative states. In practice, electrons don't do this so Dirac enlisted the help of the Pauli exclusion principle to come up with an explanation [23]. Since two fermions can't share the same quantum state, Dirac hypothesised that all of the negative energy states were already filled; an infinite number of electrons of infinite negative charge and infinite negative energy, known as the "Dirac sea".

If a high energy photon hits a negative energy electron, it can be knocked into a positive energy state, leaving behind a "hole" in the sea (the same concept as appears in semiconductor physics). The sea has lost a negatively charged, negative energy electron, so its energy has increased by $|E|$ and its charge by $|e|$. Dirac published a paper in 1931 [24] predicting the hole would behave like a particle with these properties and voilà, just one year later Carl Anderson discovered the positron [25]. This was a triumphant moment for Dirac and one of the first discoveries of a new particle preceded by a theoretical prediction.

[1]The prize was jointly awarded to Dirac and Erwin Schrödinger "for the discovery of new productive forms of atomic theory".

However, there's a catch. Bosons have antiparticle counterparts too, such as W^+ and W^-. The problem of bosons infinitely radiating photons as they make a nightmarish, never-ending descent down energy levels cannot be salvaged by the Pauli exclusion principle. Hence, the Dirac sea interpretation is viewed today as merely a historical curiosity. Richard Feynman, and earlier Ernst Stückelberg, proposed an alternative interpretation [26], that negative energy particles are really particles with positive energy moving *backwards in time*. This interpretation ensures the physical observed energy is always positive and reformulates all the valid predictions of the Dirac sea, such as pair production (and conversely, electron-positron annihilation).

Since the discovery of the positron, it has been experimentally verified that all known particles have antiparticle counterparts. Some particles, such as the photon, are their own antiparticles. All physical properties of antimatter (mass, lifetime, *etc.*) other than charge are predicted to be identical to that of matter under the CPT theorem.[2] This has been tested and confirmed at high energies by experiments at the LHC and low energies by experiments such as ALPHA, which measured the hyperfine structure of antihydrogen [27].

2.1.2 Baryogenesis

If antimatter and matter have the same physical properties, the natural question to ask is *what happened to all the antimatter*? If there were regions of space filled with antimatter stars, solar systems, and galaxies, then there would be boundaries between these regions and our familiar matter-based world where annihilation occurs, forming photons. No such photon background is observed, which puts stringent limits on the fraction of antimatter in the universe [28].

The baryon asymmetry is defined as the difference in the number of baryons N_B and antibaryons $N_{\bar{B}}$ in the early universe divided by their sum. Since the product of annihilation is photons and there are no antibaryons observed today, the asymmetry can be estimated by the baryon to photon ratio [29],

$$\eta = \frac{N_B}{N_\gamma}, \tag{2.2}$$

[2] CPT theorem predicts that the combined operation of charge conjugation (C), parity transformation (P) and time reversal (T) is a fundamental symmetry of nature.

which has been measured using astronomical data to be $\sim 10^{-10}$. This means that as the universe cooled and matter and antimatter annihilated into photons, for every 10^{10} annihilations there was just one baryon left over. Hence, the total present-day asymmetry between matter and antimatter is in fact due to an asymmetry in the past smaller than one-in-a-billion.

There are two ways that a baryon asymmetry can arise. The first is that the asymmetry is built into the Big Bang and the "initial conditions" of the universe were set already favouring matter. However, it is thought that any initial asymmetry would be washed out by inflation.[3] The second, preferred, explanation is that everything started out symmetric and there was some dynamic process, known as "baryogenesis", that generated the asymmetry in the early universe. Baryogenesis implies that there must be a fundamental difference in the laws of physics for matter and antimatter.

2.1.2.1 Sakharov conditions

In 1967, Andrei Sakharov, a Russian scientist known as the "father of the Soviet hydrogen bomb", thankfully turned his attention to cosmology. He postulated that there are three conditions necessary for baryogenesis to occur [10]:

- **Baryon number violation**: this self-evident condition states that interactions that change the number of baryons must be allowed,

$$X + \bar{Y} \to B, \tag{2.3}$$

  where X is a baryon, $\bar{Y}$ is an antibaryon and B is a set of particles with baryon number sum $\neq 0$.

- ***C* and *CP* violation**: C symmetry violation is necessary so that interactions that produce more baryons than antibaryons are not counterbalanced by interactions that produce more antibaryons than baryons. CP symmetry violation is similarly necessary otherwise equal numbers of left-handed baryons and right-handed antibaryons (or right-handed baryons and left-handed antibaryons) would be produced.

[3] While this is generally accepted to be true, it has been shown that models with extreme fine-tuning can be constructed to overcome this [30].

Hence, the rates of the following interactions must be different,

$$X + \bar{Y} \rightarrow B, \tag{2.4}$$

$$\bar{X} + Y \rightarrow \bar{B}, \tag{2.5}$$

otherwise any change in baryon number would be cancelled out. *CP* violation is described in more detail in section 2.2.

- **Interactions out of thermal equilibrium**: in thermal equilibrium, the rate of a reaction is the same as its reverse,

$$X + \bar{Y} \rightarrow B, \tag{2.6}$$

$$B \rightarrow X + \bar{Y}, \tag{2.7}$$

again, cancelling out any change in baryon number.

2.1.2.2 Baryogenesis within the Standard Model

Remarkably, each of the three Sakharov conditions are satisfied by the SM, so in principle, it can incorporate baryogenesis. This scenario is known as "electroweak baryogenesis" [31].

Baryon number violation

Although baryon number is perturbatively conserved in the SM, *i.e.* all Feynman diagrams conserve baryon number, there are non-perturbative quantum effects, known as "sphalerons", that can violate baryon number conservation [32]. These effects are expected to be small at today's temperatures since they are due to quantum mechanical tunnelling. As such, baryon number violation has never been directly observed [3]. At high temperatures, there is no such suppression so these effects have been predicted to play a role in generating the baryon asymmetry, both in electroweak baryogenesis and beyond the SM (BSM) models.

***C* and *CP* violation**

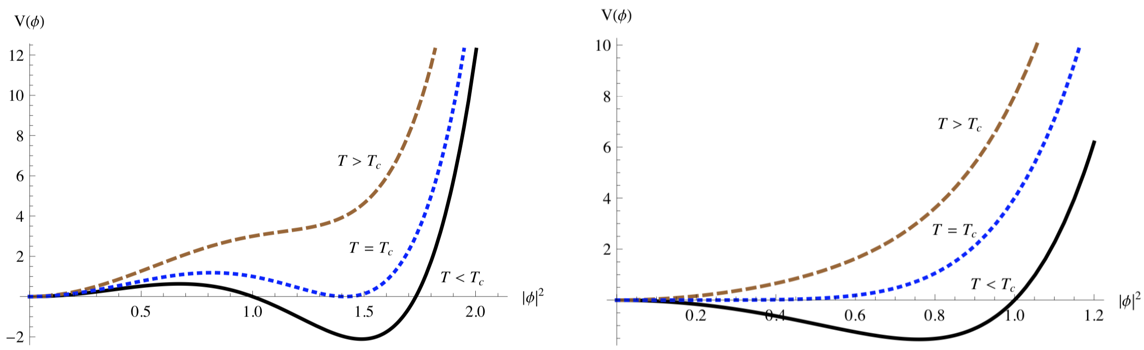
C symmetry is maximally violated in the weak interaction due to its chirality. When a left-handed particle is charge conjugated it becomes a left-handed antiparticle, which doesn't interact. Similarly, P symmetry is maximally violated since a left-handed particle becomes a right-handed particle under parity conjugation. The product of the two, CP symmetry, was thought to be conserved until Jim Christensen, James Cronin, Val Fitch and René Turlay unexpectedly observed CP violation in the decays of neutral kaons in 1964 [33]. CP violation has since been observed in the decays of D^0 [34], B^0, B^+ and B_s^0 mesons [3] and evidence has been found in Λ_b^0 decays [35]. All existing measurements of CP violation are elegantly explained by a single complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix. Present measurements of CP-violating observables imply that CP violation in the quark sector is small.

There is no experimental evidence of CP violation in the strong interaction, however, the QCD Lagrangian allows terms that are able to violate CP symmetry. If such terms existed, they would contribute to the neutron electric dipole moment, but current experimental limits place stringent limits on strong CP violation [36]. As there is no known reason for CP symmetry to be conserved in QCD, the amount of CP violation appears to be fine-tuned. This is known as the strong CP problem.

Interactions out of thermal equilibrium

In the SM, particles acquire their masses via the Higgs mechanism [37–39]. This involves a transition from a symmetric phase to a phase in which electroweak symmetry has been spontaneously broken. As the universe cools below the critical temperature at which the electroweak phase transition takes place, bubbles of the broken phase (the "true vaccum") start forming in the unbroken phase (the "false vacuum"). As the bubbles grow and nucleate, different regions of the universe pass through the phase boundary. As shown in figure 2.1, if the transition is first-order, there is an abrupt change in the position of the minimum of the Higgs potential at the critical temperature. This discontinuity is necessary to drive the universe out of thermal equilibrium. In a second-order transition, the change in the potential is continuous and would leave the universe in thermal equilibrium. Hence, a first-order electroweak phase transition is necessary in order to satisfy the last of the Sakharov conditions.

Although electroweak baryogenesis offers a convenient solution to the baryon asymmetry, it has recently been ruled out. This is largely because measurements of the Higgs mass don't allow a first-order electroweak phase transition [41]. However, even if there were a first-order phase transition, the levels of CP violation measured in the quark



(a) A first-order transition.

(b) A second-order transition

Figure 2.1: The Higgs potential for a first and second-order electroweak phase transition for various temperatures above and below the critical temperature T_c [40].

sector would be insufficient [11]. The SM is only capable of generating an asymmetry of $\eta \sim 10^{-27}$ [42], which is orders of magnitudes below the observed value of $\eta \sim 10^{-10}$.

2.1.2.3 Baryogenesis beyond the Standard Model

Electroweak baryogenesis is the only model for baryogenesis that exists that doesn't require an extension of the SM. Since this has been ruled out, the baryon asymmetry is widely regarded as strong evidence of physics beyond the SM. In this section, some of the leading new physics models for baryogenesis are reviewed.

Leptogenesis

In the original formulation of the SM, all three neutrinos are massless and left-handed. In 1998, evidence for neutrino oscillations was reported by the Super-Kamiokande collaboration [43]. The probability for such oscillations to occur is proportional to $\sin^2(\Delta m^2)$, where Δm is the difference in mass of the two neutrinos. Hence, the discovery of neutrino oscillations implies that at least two of the three neutrino generations have non-zero mass. This is the only evidence of physics beyond the SM in a laboratory experiment.

Neutrino oscillations are described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [44–46]. Analogously to the CKM matrix, the PMNS matrix may contain one or more complex phases responsible for CP violation in the lepton sector. The current generation of neutrino experiments are not yet sensitive enough to measure these phases, however, future experiments such as DUNE [47] and Hyper-Kamiokande [48] may be able to make an observation of one of them in the next few decades, depending on its value.

The "see-saw" mechanism is an extension of the SM that is able to explain why neutrinos have mass, but are much lighter than their quark counterparts. Remarkably, it simultaneously provides a model for baryogenesis known as "leptogenesis" [49]. This is because the see-saw mechanism requires the introduction of at least two heavy, right-handed neutrinos to the SM Lagrangian. The right-handed neutrinos decay into states containing leptons or antileptons, violating lepton number conservation and *CP* symmetry, and resulting in a departure from thermal equilibrium. The sphaleron process then converts some of this lepton asymmetry into a baryon asymmetry. Depending on the model used for the see-saw mechanism, the required *CP* violation may or may not originate from the PMNS matrix. Hence, leptogenesis can still be realised even if there are no *CP*-violating phases in the PMNS matrix and conversely, a discovery would not be a smoking gun for leptogenesis.

Supersymmetry

The Minimal Supersymmetric Standard Model (MSSM) is an extension to the SM that predicts that every SM fermion (boson) has a boson (fermion) superpartner. Electroweak baryogenesis with the MSSM can generate the observed baryon asymmetry, although in a restricted parameter space [51]. Supersymmetric electroweak baryogenesis is attractive since it is potentially experimentally accessible, however, no evidence for supersymmetric particles has been discovered to date.

Grand Unified Theories

A Grand Unified Theory (GUT) is a model in which the three forces described by the SM are unified at very high energies. GUTs satisfy the three Sakharov conditions [50]: they predict tree-level baryon number violation, inherit *CP* violation from the SM as well as introducing new sources, and include new heavy particles with masses $M_{GUT} \sim 10^{16}$ GeV/c^2 whose decays can provide a departure from thermal equilibrium. A weakness of GUT baryogenesis is the lack of testability since GUT scales are difficult to reach in the laboratory.

2.1.2.4 Experimental prospects

Currently, there is no experimental evidence favouring one baryogenesis model over another. Some models introduce new physics entirely outwith the reach of any particle accelerator built today, or any that could conceivably be built in the future. The only Sakharov condition there is direct experimental evidence of is C and CP violation, however, there must be additional, new physics sources of CP violation to generate a large enough baryon asymmetry. Hence, the study of CP violation might be the key to unlocking the secrets of baryogenesis. As CP symmetry is not a symmetry of nature, almost every extension of the SM introduces additional sources of CP violation. Putting tight constraints on the SM contributions is vital to constrain new physics models.

2.2 Charge-parity violation

In this section, the formalism and classification of CP violation is introduced (section 2.2.1) followed by a discussion of the CKM paradigm of CP violation in the SM (section 2.2.2). An overview of the representations of the CKM matrix (section 2.2.2.1), the unitarity triangle (section 2.2.2.2) and the current experimental constraints (section 2.2.2.3) is given.

2.2.1 Types of *CP* violation

The amplitude for a particle P and its CP-conjugate $\overline{P}$ to decay to final state f and its CP-conjugate $\overline{f}$ is

$$A_f = \langle f|\mathcal{H}|P\rangle, \quad \overline{A}_f = \langle f|\mathcal{H}|\overline{P}\rangle, \tag{2.8}$$

$$A_{\overline{f}} = \langle \overline{f}|\mathcal{H}|P\rangle, \quad \overline{A}_{\overline{f}} = \langle \overline{f}|\mathcal{H}|\overline{P}\rangle, \tag{2.9}$$

where $\mathcal{H}$ is the weak Hamiltonian. If P is neutral, it can oscillate between particle and antiparticle states. The mass eigenstates are a superposition of the flavour eigenstates,

$$|P_1\rangle = p|P^0\rangle + q|\overline{P}^0\rangle, \tag{2.10}$$

$$|P_2\rangle = p|P^0\rangle - q|\overline{P}^0\rangle, \tag{2.11}$$

where p and q are complex constants representing the amount of mixing. With this formalism, three types of *CP* violation can be defined: *CP* violation in the decay, *CP* violation in mixing and *CP* violation in the interference between mixing and decay.

CP violation in the decay

CP violation in the decay occurs when the rate for the decay of $P \to f$ differs from that of $\overline{P} \to \overline{f}$,

$$|\overline{A}_{\overline{f}}/A_f| \neq 1. \tag{2.12}$$

This is easiest to identify in the decays of charged particles as mixing effects are absent. The *CP* asymmetry of a charged B meson is given by

$$\mathcal{A}^{CP} = \frac{\Gamma(B^- \to f^-) - \Gamma(B^+ \to f^+)}{\Gamma(B^- \to f^-) + \Gamma(B^+ \to f^+)} \tag{2.13}$$

$$= \frac{|\overline{A}_{f^-}/A_{f^+}|^2 - 1}{|\overline{A}_{f^-}/A_{f^+}|^2 + 1}, \tag{2.14}$$

where Γ is the decay rate. The convention is that $\mathcal{A}^{CP}$ is the difference in rate between the decay of the particle containing a heavy quark and the decay of the particle containing a heavy antiquark.

The decay amplitudes of $B^+ \to f^+$ and $B^- \to f^-$ can be written generally as the sum of all contributions,

$$A_{f^+} = \sum_i |a_i| e^{i(\delta_i + \phi_i)}, \quad \overline{A}_{f^-} = \sum_i |a_i| e^{i(\delta_i - \phi_i)}, \tag{2.15}$$

where a_i are the magnitudes, ϕ_i are the weak phases and δ_i are the strong phases. Weak phases arise from the couplings of the $W^\pm$ bosons and strong phases originate from final state interactions. The sign of the strong phases are the same for A_{f^+} and $\overline{A}_{f^-}$ since the strong interaction conserves *CP* symmetry. Only the relative difference in phases is physically meaningful.

In the case of two contributions, the *CP* asymmetry is given by

$$\mathcal{A}^{CP} = -\frac{2|a_1||a_2| \sin(\delta_2 - \delta_1) \sin(\phi_2 - \phi_1)}{|a_1|^2 + |a_2|^2 + 2|a_1||a_2| \cos(\delta_2 - \delta_1) \cos(\phi_2 - \phi_1)}, \tag{2.16}$$

which is non-zero only if the two contributions have different weak and strong phases. Often, the parameter of interest in studies of *CP* violation is the weak phase difference. Its extraction from $\mathcal{A}^{CP}$ requires knowledge of the amplitude ratio $|a_2/a_1|$ and the strong phase difference, both of which involve non-pertubative hadronic parameters that are difficult to calculate.

CP violation in mixing

CP violation in mixing occurs when the probability for $P^0 \to \overline{P}^0$ is not the same as $\overline{P}^0 \to P^0$,

$$|q/p| \neq 1. \tag{2.17}$$

Experimentally, this is searched for using semileptonic decays $B^0 \to l^+ X$, where the charge of the lepton determines the B flavour at decay. A $b\bar{b}$ pair that hadronises to form a B^0 and a $\overline{B}^0$ would decay to two oppositely charged leptons in the absence of mixing. If the leptons have the same charge, this implies one of the B mesons oscillated. The asymmetry is measured by comparing the number of events with two positively charged leptons with two negatively charged leptons,

$$\mathcal{A}_{SL}(t) = \frac{d\Gamma/dt(\overline{B}^0(t) \to l^+X) - d\Gamma/dt(B^0(t) \to l^-X)}{d\Gamma/dt(\overline{B}^0(t) \to l^+X) + d\Gamma/dt(B^0(t) \to l^-X)} \tag{2.18}$$

$$= \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{2.19}$$

This measurement provides a clean way of isolating the effects of *CP* violation in mixing.

CP violation in the interference of mixing and decay

CP violation in the interference of mixing and decay occurs when the rate for $P^0 \to \overline{P}^0 \to f$, where f is a non-flavour specific final state, differs from the rate for $\overline{P}^0 \to P^0 \to f$. For *CP* eigenstates, $f = \overline{f}$, this is defined as

$$\mathfrak{Im}\left(\frac{q}{p}\frac{\overline{A}_f}{A_f}\right) \neq 0. \tag{2.20}$$

The *CP* asymmetry is measured using the decays of neutral particles into *CP* eigenstates,

$$\mathcal{A}^{CP}(t) = \frac{d\Gamma/dt(\overline{B}^0(t) \to f) - d\Gamma/dt(B^0(t) \to f)}{d\Gamma/dt(\overline{B}^0(t) \to f) + d\Gamma/dt(B^0(t) \to f)}. \tag{2.21}$$

This type of *CP* violation can occur even when both $|q/p| = 1$ and $|A_f/\overline{A}_f| = 1$.

2.2.2 The CKM paradigm

CP violation in the SM occurs in the quark sector via the weak interaction. The down-type quark eigenstates that participate in the weak interaction, d', s' and b', are made up of a superposition of the mass eigenstates, d, s and b,

$$|d'\rangle = a_1|d\rangle + a_2|s\rangle + a_3|b\rangle, \tag{2.22}$$

$$|s'\rangle = b_1|d\rangle + b_2|s\rangle + b_3|b\rangle, \tag{2.23}$$

$$|b'\rangle = c_1|d\rangle + c_2|s\rangle + c_3|b\rangle, \tag{2.24}$$

where a_i, b_i and c_i, for $i = 1, 2, 3$, are complex constants. The weak interaction couples quarks within the same generation in the following doublets,

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \ \begin{pmatrix} c \\ s' \end{pmatrix}, \ \text{and} \ \begin{pmatrix} t \\ b' \end{pmatrix}, \tag{2.25}$$

which means that the u quark couples only to the d' superposition of down-type states and likewise for the two heavier generations. Therefore, the constants, a_i, b_i and c_i, can be interpreted physically as the coupling strength of the u, c, and t quarks to the d, s and b quarks. For example, a_1 is the coupling strength of the u quark to the d quark.

Equations 2.22, 2.23 and 2.24 represent a rotation between weak and mass eigenstates in the complex plane. Therefore, they can be rewritten in terms of a 3×3 unitary matrix, the CKM matrix, V_{CKM},

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.26}$$

where the elements of the CKM matrix, V_{ij}, represent the coupling of the i-th up-type quark to the j-th down-type quark. The couplings between antiquarks are given by V_{ij}^*.

For a general 3×3 unitary matrix, there are nine degrees of freedom parameterised in terms of three real angles and six complex phases. The freedom to redefine the phases of the quark mass eigenstates can be used to remove five of the phases. The remaining phase, known as the Kobayashi-Maskawa phase, is responsible for all *CP* violation in the SM. If the value of this phase was zero, *CP* symmetry would be conserved. This model provides an elegant explanation of *CP* violation; only one free parameter is required to quantify the amount of *CP* violation permitted in the quark sector yet this gives an accurate account of all *CP*-violating measurements made in particle physics to date.

2.2.2.1 Representations

Taking explicit advantage of the reduction in free parameters, the CKM matrix can be rewritten under different representations. The "Chau-Keung" representation [52] describes the CKM matrix in terms of a *CP*-violating phase, δ, and three angles, θ_{12}, θ_{13} and θ_{23},

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.27}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -s_{23}c_{12} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.28}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. The angles θ_{ij} can be chosen to lie in the first quadrant so that $c_{ij}, s_{ij} \geq 0$.

The Wolfenstein parameterisation [53] is a perturbative expansion of V_{CKM} in powers of λ, where $\lambda \equiv s_{12}$. Expanding to order λ^3,

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{2.29}$$

where the other parameters are defined as

$$A \equiv \frac{s_{23}}{s_{12}^2}, \ \rho \equiv \frac{s_{13}\cos\delta}{s_{12}s_{23}}, \ \text{and } \eta \equiv \frac{s_{13}\sin\delta}{s_{12}s_{23}}. \tag{2.30}$$

The relative size of the couplings is evident from the power of λ each term contains, where λ is approximately 0.22 [3]. In this representation, η is the term corresponding to the complex phase responsible for *CP* violation and it can be seen that η terms only enter V_{CKM} at $\mathcal{O}(\lambda^3)$. This indicates that the amount of *CP* violation in the quark sector is expected to be small.

2.2.2.2 The unitarity triangle

The unitarity condition of the CKM matrix, $VV^\dagger = V^\dagger V = \mathbb{1}$, imposes constraints on its elements,

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}, \ \sum_j V_{ij}V_{kj}^* = \delta_{ik}. \tag{2.31}$$

The six vanishing combinations from these constraints can be represented geometrically as triangles in the complex plane. The areas of the triangles are all the same and quantify the amount of *CP* violation in the SM.

The most commonly used unitarity triangle arises from

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.32}$$

as each of the terms is of comparable size, so the corresponding triangle has interior angles of $\mathcal{O}(1)$. Dividing equation 2.32 by its best-known term $V_{cd}V_{cb}^*$,

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0, \tag{2.33}$$

gives a triangle with vertices at (0,0), (1,0) and ($\overline{\rho}$, $\overline{\eta}$), where $\overline{\rho} = \rho(1-\frac{1}{2}\lambda^2)$ and $\overline{\eta} = \eta(1-\frac{1}{2}\lambda^2)$. This triangle is shown in figure 2.2. The three interior angles are defined

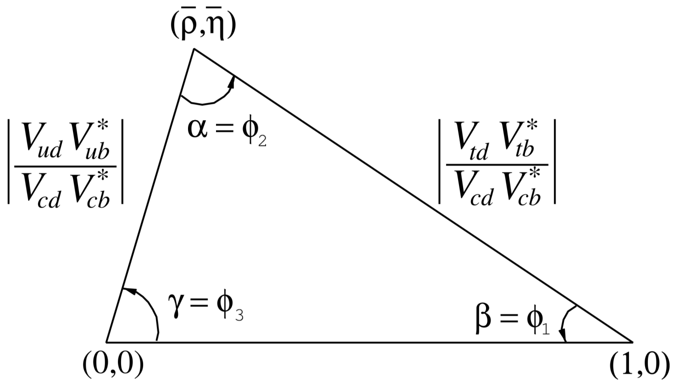


Figure 2.2: The unitarity triangle corresponding to equation 2.33 [3]. γ is the interior angle of the vertex at the origin.

as

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \tag{2.34}$$

$$\beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \tag{2.35}$$

$$\gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.36}$$

Within the SM, measurements of α, β, and γ, and the CKM elements should all be consistent with their interpretation in terms of the angles and sides of a triangle. In the presence of new physics, there may be some inconsistency in these measurements. The CKM angles α, β, and γ can all be measured independently so achieving precise measurements and "overconstraining" the unitarity triangle is not only an important test of the SM but also an indirect search for new physics.

From equation 2.36, it can be seen that γ doesn't depend on CKM elements involving the top quark. Therefore, unlike α and β, it can be measured using tree-level decays. This makes γ an important SM benchmark since it is unlikely to be affected by new physics.

2.2.2.3 Current experimental constraints

The sizes of the CKM matrix couplings are not predicted in the SM and have to be experimentally constrained. The independently measured values for the magnitudes are [3]

$$|V_{\text{CKM}}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97420 \pm 0.00021 & 0.2243 \pm 0.0005 & 0.00394 \pm 0.00036 \\ 0.218 \pm 0.004 & 0.997 \pm 0.017 & 0.0422 \pm 0.0008 \\ 0.0081 \pm 0.0005 & 0.0394 \pm 0.0023 & 1.019 \pm 0.025 \end{pmatrix}, \tag{2.37}$$

which display a hierarchical structure with dominant diagonal elements. There is no known reason for this hierarchy and interestingly, this structure is not mirrored in the PMNS matrix. The values measured are consistent with unitarity.

The CKM angles can also be directly constrained with *CP* violation measurements in B meson decays. Figure 2.3 illustrates the various experimental constraints on the $(\bar{\rho}, \bar{\eta})$ plane. The world-average values obtained by combining direct measurements of the CKM angles from LHCb and other experiments are [54]

$$\alpha = (86.4^{+4.5}_{-4.3})^\circ, \tag{2.38}$$

$$\beta = (22.14^{+0.69}_{-0.67})^\circ, \tag{2.39}$$

$$\gamma = (72.1^{+5.4}_{-5.7})^\circ. \tag{2.40}$$

Their sum, $\alpha + \beta + \gamma = (180 \pm 7)^\circ$, is consistent with the SM expectation, however, there is still room for new, sub-dominant sources of *CP* violation [55].

The least well-constrained of the CKM angles is γ. The experimental uncertainty is many orders of magnitude greater than the irreducible theoretical uncertainty $|\delta\gamma| \lesssim \mathcal{O}(10^{-7})$ [56]. Hence it is essential to both update existing γ measurements as more data is collected by the LHCb detector and to add new γ measurements using as of yet unexplored decay modes.

The apex of the unitarity triangle $(\bar{\rho}, \bar{\eta})$ can be determined using tree-level or loop-level processes, as shown in figure 2.4. The tree-level determination combines γ and the side opposite β, since these are the only parameters that don't depend on V_{tq}. The loop-level determination combines several well-known measurements including β and the neutral B mixing parameters Δm_d and Δm_s. Under the SM, the apex should be the same in both of these determinations and any deviation would point towards new physics. They are currently consistent, however, the comparison is limited by the precision of the tree-level measurements.

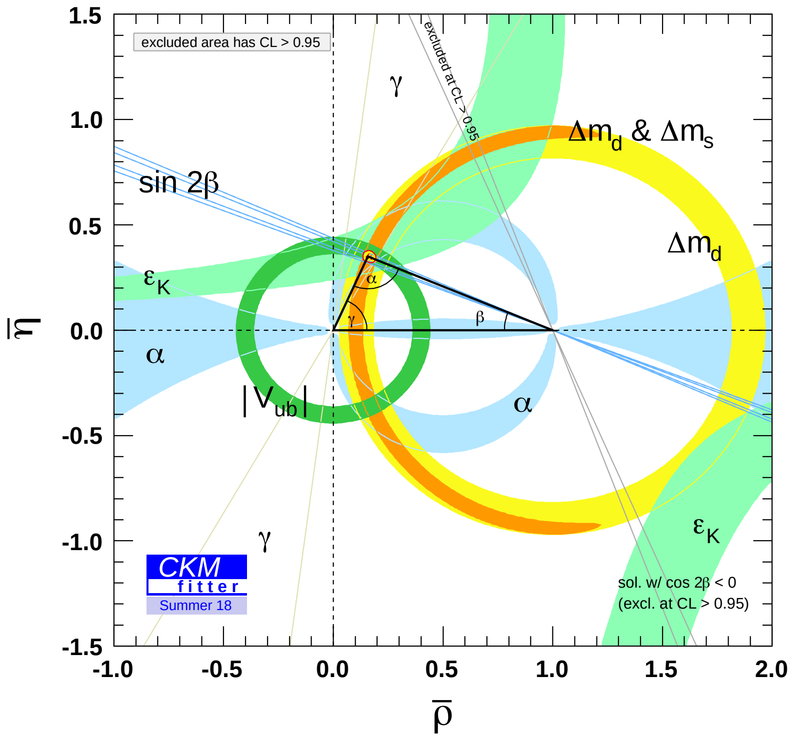
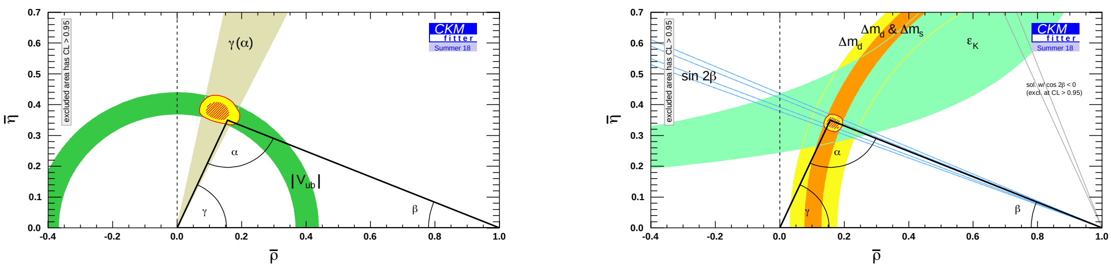


Figure 2.3: Experimental constraints on the unitarity triangle [54].



(a) Tree-level **(b)** Loop-level

Figure 2.4: Experimental constraints on the unitarity triangle from tree and loop-level processes [54].

2.3 Measuring γ

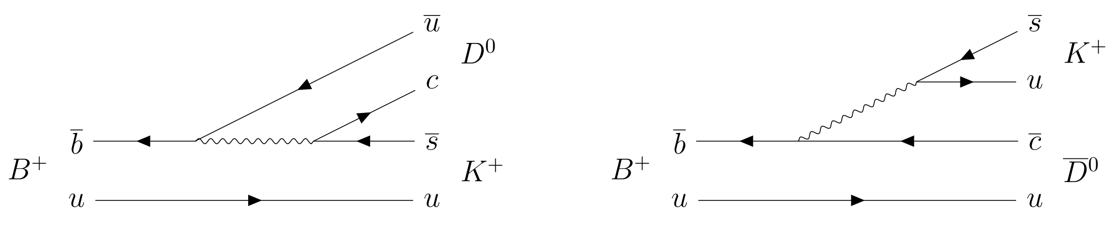
As can be seen from equation 2.36, γ is the weak phase difference between $V_{ud}V_{ub}^*$ and $V_{cd}V_{cb}^*$, and is therefore accessed in the interference of $b\to u$ and $b\to c$ decays. Traditionally, the most precise measurements of γ have been made with $B^+\to DK^+$ decays, where D represents a neutral charm meson that is a mixture of the D^0 and $\overline{D}^0$ flavour eigenstates. Measurements of γ have also been made with B^0 and B_s^0 decays, but not yet with B_c^+ decays [57]. The analogous γ-sensitive decays in the B_c^+ system are $B_c^+\to D_s^{(*)+}D^{(*)}$. In this section, the method to measure γ is described using $B^+\to DK^+$ as a textbook case (section 2.3.1), followed by the extension of this method to $B_c^+\to D_{(s)}^{(*)+}D^{(*)}$ decays (section 2.3.2).

2.3.1 $B^+\to DK^+$

The tree-level Feynman diagrams for $B^+\to DK^+$ decays are shown in figure 2.5. Diagrams where a W^+ boson produces a quark-antiquark pair that end up in the same meson are called "colour allowed" as the pair is produced in the required colour singlet state. However, this is not true for the diagrams where the quark-antiquark pair end up in different mesons, which results in a suppression factor of $1/N_c$, where N_c is the number of colours. The Cabibbo suppressed $b\to u$ diagram (figure 2.5a) is paired with an additional colour suppression, while the Cabibbo favoured $b\to c$ diagram (figure 2.5b) is not. The ratio of the suppressed to the favoured decay is

$$r_B = \left|\frac{A(B^+\to D^0K^+)}{A(B^+\to \overline{D}^0K^+)}\right| = \left|\frac{A(B^-\to \overline{D}^0K^-)}{A(B^-\to D^0K^-)}\right| = \mathcal{O}(0.1), \tag{2.41}$$

which limits the sensitivity to γ.



(a) $B^+\to D^0K^+$ **(b)** $B^+\to \overline{D}^0K^+$

Figure 2.5: Feynman diagrams for $B^+\to DK^+$ decays. $B^+\to D^0K^+$ is Cabibbo and colour suppressed, while $B^+\to \overline{D}^0K^+$ is Cabibbo and colour favoured.

It is possible for the neutral D meson produced in the $B^+ \to DK^+$ decay to oscillate into its antiparticle before decaying, which could result in a bias in the measurement of γ. Corrections due to D^0-$\overline{D}^0$ mixing are expected to be small ($\lesssim 1°$) [58] and for simplicity are ignored in the following discussion.

Interference occurs when D^0 and $\overline{D}^0$ decay to the same final state. Three different methods to measure γ have been developed, depending on which D decay is chosen. The Gronau-London-Wyler (GLW) method [59,60] uses the decay of the D into a *CP* eigenstate, such as K^+K^-. The Atwood-Dunietz-Soni (ADS) method [61–63] considers decays of the D with a pattern of Cabibbo dominance and suppression to counterbalance that of the B decays, such as $K^-\pi^+$. The Giri-Grossman-Soffer-Zupan (GGSZ) method [64] involves a Dalitz analysis of three-body D decays, such as $K^0_S\pi^+\pi^-$.

2.3.1.1 GLW method

The GLW method considers D decays to *CP* eigenstates, such that $f = \overline{f}$. The *CP* even ($+$) and odd ($-$) eigenstates of the neutral D system are made up of a superposition of D^0 and $\overline{D}^0$,

$$D^0_+ = \frac{1}{\sqrt{2}}(|D^0\rangle + |\overline{D}^0\rangle), \tag{2.42}$$

$$D^0_- = \frac{1}{\sqrt{2}}(|D^0\rangle - |\overline{D}^0\rangle). \tag{2.43}$$

The decay amplitudes can therefore be written as

$$A(B^+ \to D^0_\pm K^+) = \frac{1}{\sqrt{2}}(A(B^+ \to D^0K^+) \pm A(B^+ \to \overline{D}^0K^+)), \tag{2.44}$$

$$A(B^- \to D^0_\pm K^-) = \frac{1}{\sqrt{2}}(A(B^- \to D^0K^-) \pm A(B^- \to \overline{D}^0K^-)). \tag{2.45}$$

Since only $b \to u$ transitions involve a *CP*-violating weak phase in the Wolfenstein parameterisation, the flavour amplitudes follow,

$$A(B^+ \to \overline{D}^0K^+) = A(B^- \to D^0K^-), \tag{2.46}$$

$$A(B^+ \to D^0K^+) = e^{2i\gamma}A(B^- \to \overline{D}^0K^-), \tag{2.47}$$

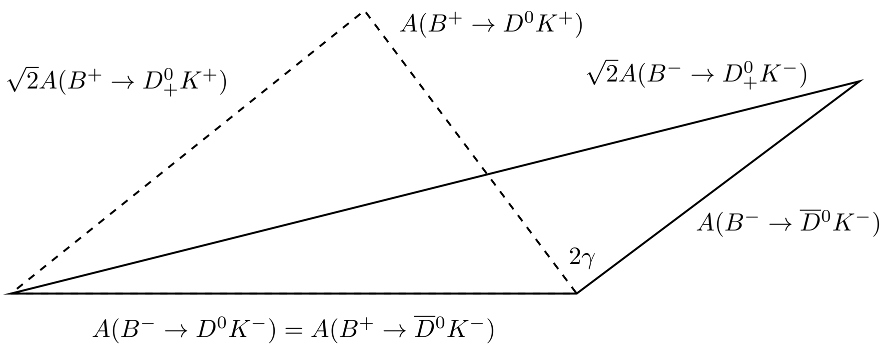


Figure 2.6: Illustration of the relationship between the $B^+ \to DK^+$ amplitudes and γ for the GLW method. The amplitude triangles are not drawn to scale and in reality are much more stretched since $A(B^+ \to D^0 K^+)$ is an order of magnitude smaller than $A(B^+ \to \overline{D}^0 K^+)$.

where γ is the weak phase difference between $B^+ \to D^0 K^+$ and $B^+ \to \overline{D}^0 K^+$ decays. The amplitude relations above can be represented as triangles in the complex plane, as shown in figure 2.6. Due to the small value of r_B, the triangles are highly stretched.

Making use of the formalism introduced in section 2.2.1, the physics parameters are related to the following experimental observables,

$$\mathcal{A}^{CP\pm} = \frac{\Gamma(B^- \to D^0_\pm K^-) - \Gamma(B^+ \to D^0_\pm K^+)}{\Gamma(B^- \to D^0_\pm K^-) + \Gamma(B^+ \to D^0_\pm K^+)} \tag{2.48}$$

$$= \frac{\pm 2r_B \sin(\delta_B)\sin(\gamma)}{1 + r_B^2 \pm 2r_B \cos(\delta_B)\cos(\gamma)}, \tag{2.49}$$

and

$$\mathcal{R}_{CP\pm} = 2\frac{\Gamma(B^- \to D^0_\pm K^-) + \Gamma(B^+ \to D^0_\pm K^+)}{\Gamma(B^- \to D^0 K^-) + \Gamma(B^+ \to \overline{D}^0 K^+)} \tag{2.50}$$

$$= 1 + r_B^2 \pm 2r_B \cos(\delta_B)\cos(\gamma). \tag{2.51}$$

where δ_B is the strong phase difference between $B^+ \to D^0 K^+$ and $B^+ \to \overline{D}^0 K^+$ decays. The hadronic parameters, r_B and δ_B, are different for each B decay, but γ is universal to all interfering $b \to u$ and $b \to c$ decays.

This method can be extended to quasi-*CP* eigenstates, such as $K^+K^-\pi^0$, if the fraction of *CP* even content F_{CP} is known. The observables then become,

$$\mathcal{A}^{CP\pm} = \frac{\pm 2r_B(2F_{CP}-1)\sin(\delta_B)\sin(\gamma)}{1+r_B^2 \pm 2r_B(2F_{CP}-1)\cos(\delta_B)\cos(\gamma)}, \tag{2.52}$$

$$\mathcal{R}_{CP\pm} = 1+r_B^2 \pm 2r_B(2F_{CP}-1)\cos(\delta_B)\cos(\gamma). \tag{2.53}$$

Measurements of F_{CP} can be made with dedicated charm experiments, such as CLEO-c [65] and BES III [66].

$\mathcal{A}^{CP\pm}$ and $\mathcal{R}_{CP\pm}$ are invariant under three symmetry operations: $\gamma \leftrightarrow \delta_B$, $(\gamma,\delta_B) \rightarrow (-\gamma,-\delta_B)$, and $(\gamma,\delta_B) \rightarrow (\gamma+\pi,\delta_B+\pi)$. The GLW method can therefore only determine γ up to an eight-fold ambiguity. Measurements involving different strong phases can help to resolve these ambiguities.

2.3.1.2 ADS method

The ADS method exploits D decays to non-*CP* eigenstates, namely the Cabibbo favoured $D^0 \rightarrow K^-\pi^+$ decay and the doubly Cabibbo suppressed $D^0 \rightarrow K^+\pi^-$ decay. The amplitude of the favoured B decay followed by the doubly Cabibbo suppressed D decay is similar to that of the suppressed B decay followed by the favoured D decay, enhancing the *CP* asymmetry. The downside to this approach is that the signal yields are smaller. The experimental observables in this case are

$$\mathcal{A}_{ADS} = \frac{\Gamma(B^- \rightarrow [K^+\pi^-]_D K^-) - \Gamma(B^+ \rightarrow [K^-\pi^+]_D K^+)}{\Gamma(B^- \rightarrow [K^+\pi^-]_D K^-) + \Gamma(B^+ \rightarrow [K^-\pi^+]_D K^+)} \tag{2.54}$$

$$= \frac{2r_B r_D \sin(\delta_B+\delta_D)\sin(\gamma)}{r_D^2 + r_B^2 + 2r_B r_D \cos(\delta_B+\delta_D)\cos(\gamma)}, \tag{2.55}$$

and

$$\mathcal{R}_{ADS} = \frac{\Gamma(B^- \rightarrow [K^+\pi^-]_D K^-) + \Gamma(B^+ \rightarrow [K^-\pi^+]_D K^+)}{\Gamma(B^- \rightarrow [K^-\pi^+]_D K^-) + \Gamma(B^+ \rightarrow [K^+\pi^-]_D K^+)} \tag{2.56}$$

$$= r_D^2 + r_B^2 + 2r_B r_D \cos(\delta_B+\delta_D)\cos(\gamma), \tag{2.57}$$

which introduces two new hadronic parameters, r_D and δ_D. δ_D is the strong phase difference between the two D decay amplitudes and r_D is the ratio of suppressed to favoured D decays.

This method can also be extended to multibody D final states, such as $K^-\pi^+\pi^-\pi^+$, if the coherence factor κ_D that describes the dilution due to interference between resonances is known. The observables become,

$$\mathcal{A}_{ADS} = \frac{2r_B r_D \kappa_D \sin(\delta_B+\delta_D)\sin(\gamma)}{r_D^2 + r_B^2 + 2r_B r_D \kappa_D \cos(\delta_B+\delta_D)\cos(\gamma)}, \tag{2.58}$$

$$\mathcal{R}_{ADS} = r_D^2 + r_B^2 + 2r_B r_D \kappa_D \cos(\delta_B+\delta_D)\cos(\gamma), \tag{2.59}$$

where external input from charm experiments is required for the values of κ_D, δ_D and r_D.

As there are only two observables $(\mathcal{A}_{ADS}, \mathcal{R}_{ADS})$ and three physics parameters (γ, r_B, δ_B), the ADS method can only be used to extract γ if two different final states are considered. Alternatively, a combined GLW/ADS analysis provides enough information to constrain γ.

2.3.1.3 GGSZ method

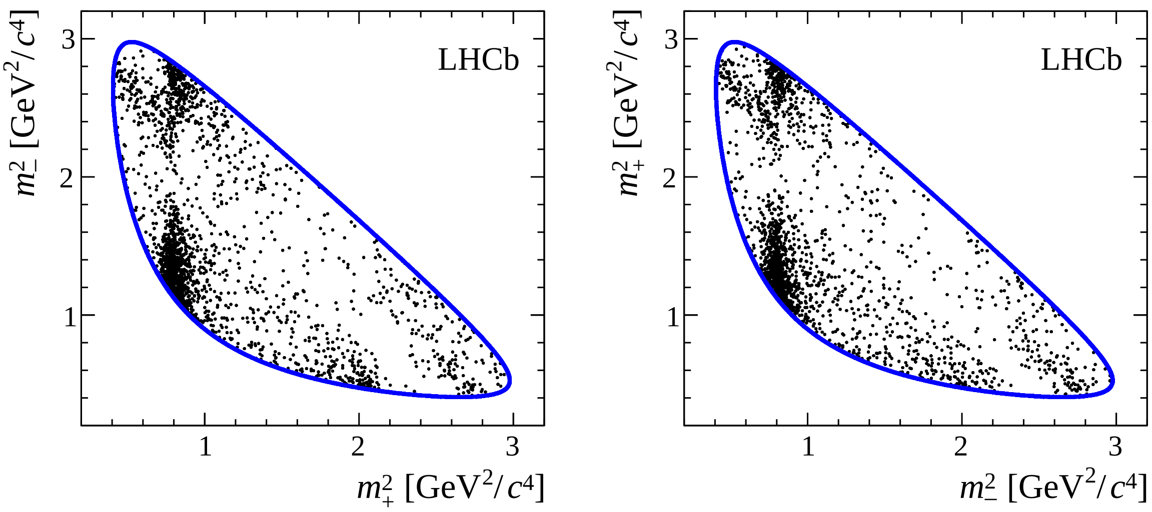
The GGSZ method considers three-body self-conjugate D decays. Decays such as $D^0 \to K^0_S\pi^+\pi^-$ are Cabibbo favoured and hence have larger branching fractions than those used in the other methods. The GGSZ method examines differences between B^+ and B^- across the phase space of the D decay, taking advantage of the various interfering resonances instead of averaging them out as is done with the coherence factor. This results in an enhanced sensitivity to γ, but is experimentally much more challenging since an amplitude analysis is required.

The Dalitz plots for $D \to K^0_S\pi^+\pi^-$ decays can be seen in figure 2.7. The amplitude of the D^0 decay at a particular point in Dalitz space is

$$A_D(m_+^2, m_-^2) = A(m_+^2, m_-^2)e^{i\delta_D(m_+^2, m_-^2)} \tag{2.60}$$

$$= A(D^0 \to K^0_S\pi^+\pi^-), \tag{2.61}$$

where $m_+^2(m_-^2)$ is the invariant mass of the $K^0_S\pi^+(K^0_S\pi^-)$ pair. Since CP violation in charm is expected to be small, it is assumed that $A(D^0 \to K^0_S\pi^+\pi^-) = A(\overline{D}^0 \to K^0_S\pi^-\pi^+)$.



(a) $B^+ \to DK^+$ **(b)** $B^- \to DK^-$

Figure 2.7: Dalitz plots for $D \to K^0_S\pi^+\pi^-$ decays from $B^\pm \to DK^\pm$ [67]. The blue line indicates the kinematic boundary.

The B partial decay rates can then be written as

$$\begin{aligned} d\Gamma(B^\pm \to [K^0_S\pi^+\pi^-]_D K^\pm) \propto\, & |A_D(m^2_\pm, m^2_\mp)|^2 + r_B^2|A_D(m^2_\mp, m^2_\pm)|^2 \\ & + 2|A_D(m^2_\pm, m^2_\mp)||A_D(m^2_\mp, m^2_\pm)|[r_B\cos(\delta_B \pm \gamma)\cos(\delta_D(m^2_\pm, m^2_\mp)) \\ & + r_B \sin(\delta_B \pm \gamma)\sin(\delta_D(m^2_\pm, m^2_\mp))]. \end{aligned} \tag{2.62}$$

Comparing the D Dalitz distributions for B^+ and B^- therefore allows r_B, δ_B and γ to be determined. A direct measurement can lead to biases for small values of r_B [68], so instead, the Cartesian *CP* violation observables are used,

$$x_\pm = r_B \cos(\delta_B \pm \gamma), \tag{2.63}$$

$$y_\pm = r_B \sin(\delta_B \pm \gamma). \tag{2.64}$$

The Cartesian coordinates are invariant under just one symmetry operation: $(\gamma, \delta_B) \to (\gamma + \pi, \delta_B + \pi)$. Therefore, the GGSZ method can resolve γ up to a two-fold ambiguity.

2.3.2 $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$

Decays of B_c^+ to two charm mesons, $B_c^+ \to D^+_{(s)}D$, have been proposed to measure γ [12–17]. A measurement of γ with these decays would be significant for two reasons. First, it would help to further constrain the CKM picture of the SM and improve the precision of the least well-known angle of the unitarity triangle. Second, since no evidence of *CP* violation has been seen in B_c^+ decays before, it would shed light on the behaviour of *CP* violation in a previously unexplored system.

A measurement of γ is also possible with B_c^+ decays to excited charm mesons. Decays with one excited charm meson in the final state, $B_c^+ \to D_{(s)}^{*+} D$ or $B_c^+ \to D_{(s)}^{+} D^*$, can be used to measure γ in the same way as $B_c^+ \to D_{(s)}^+ D$ decays. In the case of B_c^+ decays with two excited charm mesons, $B_c^+ \to D_{(s)}^{*+} D^*$, γ can be extracted from an angular analysis [14].

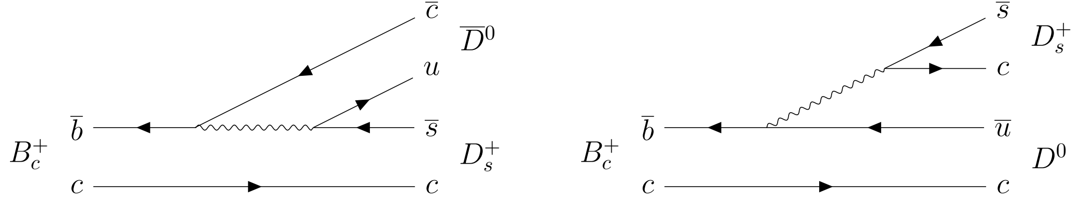
The Feynman diagrams for $B_c^+ \to D_s^+ D$ decays are shown in figure 2.8. This time, the Cabibbo favoured $b \to c$ decay, $B_c^+ \to D_s^+ \overline{D}^0$, is paired with the colour suppression and the Cabibbo suppressed $b \to u$ decay, $B_c^+ \to D_s^+ D^0$, is colour allowed. Therefore, crucially, the two amplitudes are similar in size,

$$r_{B_c^+} = \left|\frac{A(B_c^+ \to D_s^+ D^0)}{A(B_c^+ \to D_s^+ \overline{D}^0)}\right| = \left|\frac{A(B_c^+ \to D_s^- \overline{D}^0)}{A(B^- \to D_s^- D^0)}\right| = \mathcal{O}(1), \tag{2.65}$$

resulting in an excellent sensitivity to γ. To preserve the large interference, $B_c^+ \to D_s^+ D$ decays are ideally suited to a measurement of γ where the amplitude of the D^0 and $\overline{D}^0$ decays are the same, *i.e.* the GLW or GGSZ methods.

Decays of $B_c^+ \to D^+ D$ can be obtained from $B_c^+ \to D_s^+ D$ by interchanging the s and d quarks (known as U-spin symmetry). In this case, due to the different CKM elements involved,

$$r_{B_c^+} = \left|\frac{A(B_c^+ \to D^+ D^0)}{A(B_c^+ \to D^+ \overline{D}^0)}\right| = \left|\frac{A(B_c^+ \to D^- \overline{D}^0)}{A(B^- \to D^- D^0)}\right| = \mathcal{O}(0.1). \tag{2.66}$$



(a) $B_c^+ \to D_s^+ \overline{D}^0$ **(b)** $B_c^+ \to D_s^+ D^0$

Figure 2.8: Feynman diagrams for $B_c^+ \to D_s^+ D$ decays. $B_c^+ \to D_s^+ \overline{D}^0$ is Cabibbo favoured and colour suppressed, while $B_c^+ \to D_s^+ D^0$ is Cabibbo suppressed and colour favoured.

Although these decays are sensitive to γ, they suffer from the same experimental problems as $B^+ \to DK^+$ decays. Hence, the most promising approach is measuring γ with $B_c^+ \to D_s^+ D$ decays.

Although all three methods can be used to extract γ in a manner analogous to $B^+ \to DK^+$ decays, the discussion here focuses on the GLW method. This is the simplest method to extract γ that takes full advantage of the large interference between $B_c^+ \to D_s^+ D^0$ and $B_c^+ \to D_s^+ \overline{D}^0$. It is also favoured by the theory community and as such, has been developed much more extensively than any other method.

2.3.2.1 GLW method

The application of the GLW method to $B_c^+ \to D_s^+ D$ decays is very similar to that of $B^+ \to DK^+$ decays. The *CP* eigenstates defined in equation 2.44, allow the following amplitude relations to be written,

$$A(B_c^+ \to D_s^+ D_\pm^0) = \frac{1}{\sqrt{2}}(A(B_c^+ \to D_s^+ D^0) \pm A(B_c^+ \to D_s^+ \overline{D}^0)), \tag{2.67}$$

$$A(B_c^- \to D_s^- D_\pm^0) = \frac{1}{\sqrt{2}}(A(B_c^- \to D_s^- D^0) \pm A(B_c^- \to D_s^- \overline{D}^0)), \tag{2.68}$$

where the flavour amplitudes follow,

$$A(B_c^+ \to D_s^+ \overline{D}^0) = A(B_c^- \to D_s^- D^0), \tag{2.69}$$

$$A(B_c^+ \to D_s^+ D^0) = e^{2i\gamma} A(B_c^- \to D_s^- \overline{D}^0). \tag{2.70}$$

Figure 2.9 shows the representation of these amplitudes as triangles in the complex plane. In the $B_c^+ \to D_s^+ D$ system, the triangles have sides of similar length, resulting in a much more favourable scenario than that of the highly stretched triangles in the $B^+ \to DK^+$ system (figure 2.6).

The experimental observables are

$$\mathcal{A}^{CP\pm} = \frac{\Gamma(B_c^- \to D_s^- D_\pm^0) - \Gamma(B_c^+ \to D_s^+ D_\pm^0)}{\Gamma(B_c^- \to D_s^- D_\pm^0) + \Gamma(B_c^+ \to D_s^+ D_\pm^0)} \tag{2.71}$$

$$= \frac{\pm 2 r_{B_c^+}(2F_{CP} - 1)\sin(\delta_{B_c^+})\sin(\gamma)}{1 + r_{B_c^+}^2 \pm 2 r_{B_c^+}(2F_{CP} - 1)\cos(\delta_{B_c^+})\cos(\gamma)}, \tag{2.72}$$

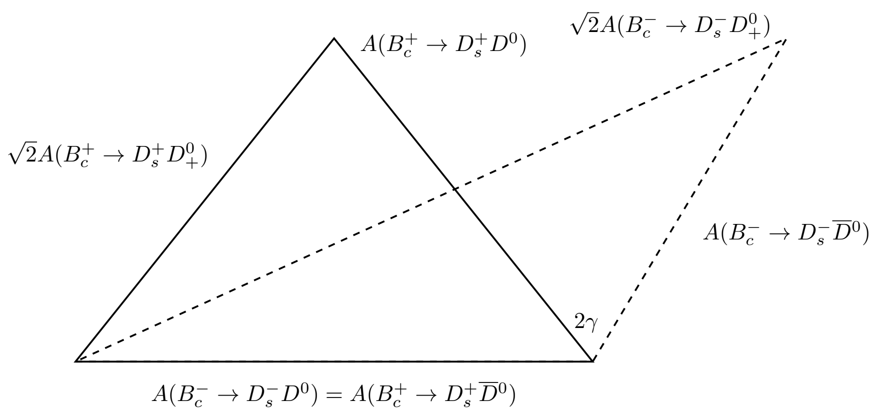


Figure 2.9: Illustration of the relationship between the $B_c^+ \to D_s^+ D$ amplitudes and γ for the GLW method. The triangles have interior angles of $\mathcal{O}(1)$ since $A(B_c^+ \to D_s^+ D^0)$ and $A(B_c^+ \to D_s^+ \overline{D}^0)$ are expected to be similar in size.

and

$$\mathcal{R}_{CP\pm} = 2\frac{\Gamma(B_c^- \to D_s^- D_\pm^0) + \Gamma(B_c^+ \to D_s^+ D_\pm^0)}{\Gamma(B_c^- \to D_s^- D^0) + \Gamma(B_c^+ \to D_s^+ \overline{D}^0)} \tag{2.73}$$

$$= 1 + r_{B_c^+}^2 \pm 2r_{B_c^+}(2F_{CP} - 1)\cos(\delta_{B_c^+})\cos(\gamma). \tag{2.74}$$

where $\delta_{B_c^+}$ is the strong phase difference between $B_c^+ \to D_s^+ D^0$ and $B_c^+ \to D_s^+ \overline{D}^0$. Since both final states are isospin 1/2 and the strong interaction conserves isospin, $\delta_{B_c^+}$ is expected to vanish [14]. This eliminates uncertainties related to $\delta_{B_c^+}$ and enables a clean determination of γ. However, in practice there may be some residual strong phase difference due to resonance effects [14].

2.3.2.2 Experimental prospects

The flavour specific $B_c^+ \to D_{(s)}^+ D$ decays are related to the high-yield $B^+ \to D_{(s)}^+ \overline{D}^0$ channel in the following way,

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_{(s)}^+ D)}{\mathcal{B}(B^+ \to D_{(s)}^+ \overline{D}^0)} = \frac{N(B_c^+ \to D_{(s)}^+ D)}{N(B^+ \to D_{(s)}^+ \overline{D}^0)}\frac{\varepsilon(B^+ \to D_{(s)}^+ \overline{D}^0)}{\varepsilon(B_c^+ \to D_{(s)}^+ D)}, \tag{2.75}$$

where f_c/f_u is the ratio of the B_c^+ to B^+ production cross-sections, N stands for the signal yields and ε the total efficiencies. With this equation, it is possible to predict the B_c^+ signal yields given estimations for the other parameters.

	Prediction for the branching fraction $[10^{-6}]$			
Channel	Ref. [20]	Ref. [17]	Ref. [16]	Ref. [21]
$B_c^+ \to D_s^+ \overline{D}^0$	2.3 ± 0.5	4.8	1.7	2.1
$B_c^+ \to D_s^+ D^0$	3.0 ± 0.5	6.6	2.5	7.4
$B_c^+ \to D^+ \overline{D}^0$	32 ± 7	53	32	33
$B_c^+ \to D^+ D^0$	0.10 ± 0.02	0.32	0.11	0.32
$B_c^+ \to D_s^{*+} \overline{D}^0$	0.7 ± 0.2	4.5	1.0	0.7
$B_c^+ \to D_s^{*+} D^0$	2.5 ± 0.4	8.5	6.9	9.3
$B_c^+ \to D^{*+} \overline{D}^0$	12 ± 3	49	17	9
$B_c^+ \to D^{*+} D^0$	0.09 ± 0.02	0.40	0.38	0.44
$B_c^+ \to D_s^+ \overline{D}^{*0}$	2.6 ± 0.6	7.1	4.3	2.4
$B_c^+ \to D_s^+ D^{*0}$	1.9 ± 0.4	6.3	0.6	1.3
$B_c^+ \to D^+ \overline{D}^{*0}$	34 ± 7	75	83	38
$B_c^+ \to D^+ D^{*0}$	0.07 ± 0.02	0.28	0.03	0.05
$B_c^+ \to D_s^{*+} \overline{D}^{*0}$	2.8 ± 0.7	26	4.7	1.6
$B_c^+ \to D_s^{*+} D^{*0}$	2.4 ± 0.4	40	5.4	4.5
$B_c^+ \to D^{*+} \overline{D}^{*0}$	34 ± 9	330	84	21
$B_c^+ \to D^{*+} D^{*0}$	0.08 ± 0.03	1.59	0.28	0.20

Table 2.1: Estimates of the branching fractions of $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ decays in units of 10^{-6}.

The $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ decays have never before been observed. Some theoretical predictions for the branching fractions are listed in table 2.1. As expected, the $B_c^+ \to D_s^+ D^0$ and $B_c^+ \to D_s^+ \overline{D}^0$ decays have branching fractions of comparable size. The branching fraction of $B_c^+ \to D^+ \overline{D}^0$ is an order of magnitude larger, so it is expected that this decay would be observed first. The predictions for $\mathcal{B}(B_c^+)$ show some variation between the different calculations, reflecting the relatively limited theoretical understanding.

Production of a B_c^+ meson requires the simultaneous production of $b\bar{b}$ and $c\bar{c}$ pairs, hence it is rarer than the other B mesons. Only the product of $f_c \times \mathcal{B}(B_c^+)$ is accessible experimentally and since neither term is *a priori* known, theory input for $\mathcal{B}(B_c^+)$ is required to estimate f_c. The LHCb experiment has measured $R = \frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to J/\psi \pi^+)}{\mathcal{B}(B^+ \to J/\psi K^+)} = (0.683 \pm 0.018 \pm 0.009)\%$ [69] at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ in the region $0 < p_T < 20\,\text{GeV}/c^2$ and $2.0 < y < 4.5$, where p_T and y refer to the transverse momentum and rapidity of the B meson, respectively. The measurement of R at

$\sqrt{s} = 7\,\text{TeV}$ is consistent [70]. Predictions for $\mathcal{B}(B_c^+ \to J/\psi\,\pi^+)$ range from 6.1×10^{-4} [71] to 2.9×10^{-3} [72] and $\mathcal{B}(B^+ \to J/\psi\,K^+)$ has been measured to be $(1.010 \pm 0.029) \times 10^{-3}$ [3]. Therefore, f_c/f_u can be estimated to be between $0.24 - 1.1\%$.

The relative efficiency, $\varepsilon(B^+ \to D_{(s)}^+ \overline{D}^0)/\varepsilon(B_c^+ \to D_{(s)}^+ D)$, is assumed to be 2 due to the difference in lifetime $\tau(B^+)/\tau(B_c^+) \sim 3$ (see section 5.3).

A yield of $33,734 \pm 187$ was measured for $B^+ \to D_s^+ \overline{D}^0$ decays and $1,866 \pm 46$ for $B^+ \to D^+ \overline{D}^0$ decays using Run I LHCb data, corresponding to an integrated luminosity of $\mathcal{L} = 3.0\,\text{fb}^{-1}$ collected in 2011 and 2012 at centre-of-mass energies of $\sqrt{s} = 7$ and $8\,\text{TeV}$, respectively [1]. The B^+ yields increase as more data is collected at higher centre-of-mass energies and with improvements to the detector. Assuming the B^+ production cross-section scales linearly with centre-of-mass energy, the increase in yield relative to the Run I value is

$$\frac{N}{N_I} = \frac{\mathcal{L}}{\mathcal{L}_I} \frac{\sqrt{s}}{\sqrt{s}_I} \frac{\varepsilon}{\varepsilon_I}. \tag{2.76}$$

The LHCb detector is currently undergoing an upgrade and will begin its third operational run in 2021. One of the defining features of the upgrade is that a full event reconstruction will be applied to every event as part of a fully software trigger (see section 3.3.3 for an explanation of the current trigger). This is expected to greatly improve the trigger efficiency, particularly for fully hadronic B decays. For the similar $B^+ \to K^+\pi^-\pi^+(\overline{D}^0 \to K^+\pi^-)$ decay, the trigger efficiency will increase from 26% in Run I to 83% in Run III and beyond [73]. The expected increase in B^+ yields over time is shown in table 2.2.

The final missing ingredient is the yields of the $B_c^+ \to D_s^+ D_\pm^0$ decays, where $D_\pm^0$ decays to CP eigenstates, K^+K^- or $\pi^+\pi^-$. This requires estimates to be made for the physics parameters: $\gamma = 70°$, $\delta_{B_c^+} = 0°$ and $r_{B_c^+} = 3V_{ub}V_{cs}/V_{cb}V_{us} = 1.2$. Making use of equations 2.71 and 2.73, the yields are then given by

$$N(B_c^+ \to D_s^+ D_\pm^0) = \frac{1}{2} N(B_c^+ \to D_s^+ \overline{D}^0) \mathcal{R}_{CP\,\pm} (1 - \mathcal{A}^{CP\,\pm})\zeta, \tag{2.77}$$

$$N(B_c^- \to D_s^- D_\pm^0) = \frac{1}{2} N(B_c^- \to D_s^- D^0) \mathcal{R}_{CP\,\pm} (1 + \mathcal{A}^{CP\,\pm})\zeta, \tag{2.78}$$

	Estimation of the B^+ yield increase				
	Run I	Run II	Run III	Run IV	Run V
Input	2012	2018	2023	2029	2033
$\mathcal{L}$ (fb^{-1})	3	6	14	28	250
$\sqrt{s}$ (TeV)	7, 8	13	14	14	14
$\varepsilon_{\text{trigger}}$ (%)	26	26	83	83	83
N/N_I	1	3.4	27	54	490
$(N/N_I)_{\text{cumulative}}$	1	4.4	32	86	570

Table 2.2: The integrated luminosity, centre-of-mass energy, trigger efficiency, B^+ yield relative to Run I values and the cumulative increase in B^+ yields relative to Run I values for Runs I-V of the LHCb experiment.

where

$$\zeta = \frac{\varepsilon(D^0_+ \to K^+K^-)\mathcal{B}(D^0_+ \to K^+K^-) + \varepsilon(D^0_+ \to \pi^+\pi^-)\mathcal{B}(D^0_+ \to \pi^+\pi^-)}{\varepsilon(D^0 \to K^-\pi^+)\mathcal{B}(D^0 \to K^-\pi^+) + \varepsilon(D^0 \to K^-\pi^+\pi^-\pi^+)\mathcal{B}(D^0 \to K^-\pi^+\pi^-\pi^+)} \sim 0.1, \tag{2.79}$$

accounts for the differences in efficiencies and branching fractions of the D decays. It is assumed that extra kaons improve the efficiency by 15% due to their more unique particle identification signature and each additional track reduces the efficiency by a factor two,

$$\frac{\varepsilon(D^0_+ \to K^+K^-)}{\varepsilon(D^0 \to K^-\pi^+)} = 1.15,\ \frac{\varepsilon(D^0_+ \to \pi^+\pi^-)}{\varepsilon(D^0 \to K^-\pi^+)} = 0.85,\ \text{and}\ \frac{\varepsilon(D^0 \to K^-\pi^+\pi^-\pi^+)}{\varepsilon(D^0 \to K^-\pi^+)} = 0.25. \tag{2.80}$$

Table 2.3 lists the predictions for the B_c^+ yields over time with the most optimistic inputs ($f_c/f_u = 1.1\%$ and $\mathcal{B}(B_c^+)$ from reference [17]). With the most pessimistic inputs ($f_c/f_u = 0.24\%$ and $\mathcal{B}(B_c^+)$ from reference [16]), the yields are approximately an order of magnitude smaller. From these estimates, it is clear that a measurement of γ with B_c^+ decays is extremely experimentally challenging and likely not possible without the full LHCb data set up to Run V. Further improvements to the detector [74] and the inclusion of extra decay modes could improve the prospects. In the short term, observation of $B_c^+ \to D^+\overline{D}^0$ seems likely and may even be possible with the existing Run I and II data sets if there is an upward fluctuation. Observation of $B_c^+ \to D^+\overline{D}^0$ would help to improve the theoretical understanding of B_c^+ decays and constrain the predictions for $B_c^+ \to D_s^+ D$ decays.

	Estimation of the B_c^+ yield				
	Run I	Run II	Run III	Run IV	Run V
Channel	2012	2018	2023	2029	2033
$B_c^+ \to D_s^+ \overline{D}^0$	0.10	0.45	3.2	8.8	58
$B_c^+ \to D_s^+ D^0$	0.14	0.61	4.4	12	80
$B_c^+ \to D^+ \overline{D}^0$	1.5	6.5	47	130	840
$B_c^+ \to D^+ D^0$	0.01	0.04	0.28	0.76	5.1
$B_c^+ \to D_s^+ D_+^0$	0.02	0.07	0.53	1.5	9.6
$B_c^- \to D_s^- D_+^0$	0.01	0.04	0.27	0.72	4.8

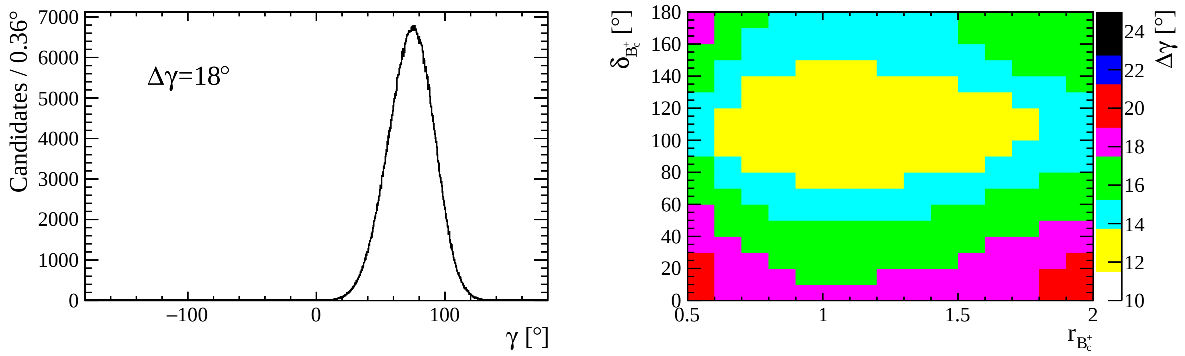
Table 2.3: Optimistic estimates of the yields of $B_c^+ \to D_{(s)}^+ D$ decays for Runs I-V of the LHCb experiment. The yields listed are cumulative, *i.e.* the number listed for Run II refers to the sum of Run I and II.

For a given set of input values for $N(B_c^+ \to D_s^+ D^0) + N(B_c^+ \to D_s^+ \overline{D}^0)$ and the physics parameters ($r_{B_c^+}$, $\delta_{B_c^+}$ and γ), the yields of the flavour and *CP* even modes can be determined from equations 2.65, 2.77 and 2.78. The uncertainty on γ can then be estimated using "toys". Each toy draws a new value for the B_c^+ yield from a Poisson distribution with mean equal to the input yield. The new set of yields is then used to re-calculate the physics parameters. The distribution of γ for $1,000,000$ toys, generated with the estimation of the Run V yields from table 2.3, $N(B_c^+ \to D_s^+ D^0) + N(B_c^+ \to D_s^+ \overline{D}^0) = 138$, $r_{B_c^+} = 1.2$, $\delta_{B_c^+} = 0°$ and $\gamma = 70°$, is shown in figure 2.10a. The spread in γ indicates an uncertainty of $\sim 18°$. The dependence of the uncertainty on γ on the choice of $r_{B_c^+}$ and $\delta_{B_c^+}$ is shown in figure 2.10b. As these toys are signal only, in reality the uncertainties are expected to be slightly larger due to the presence of background. Nonetheless, it is evident that a measurement of γ is possible with small yields of $\mathcal{O}(10)$ in the *CP* even mode $B_c^+ \to D_s^+ D_+^0$. For comparison, current γ measurements using $B^+ \to D_+^0 K^+$ decays require large $\mathcal{O}(10,000)$ yields [75].

2.4 The *CP* asymmetry in $B^+ \to D_{(s)}^+ \overline{D}^0$ decays

The *CP* asymmetry in $B^+ \to D_{(s)}^+ \overline{D}^0$ decays is defined as

$$\mathcal{A}^{CP}(B^+ \to D_{(s)}^+ \overline{D}^0) \equiv \frac{\Gamma(B^- \to D_{(s)}^- D^0) - \Gamma(B^+ \to D_{(s)}^+ \overline{D}^0)}{\Gamma(B^- \to D_{(s)}^- D^0) + \Gamma(B^+ \to D_{(s)}^+ \overline{D}^0)}. \quad (2.81)$$



(a) Distribution of γ for $1,000,000$ toys generated with $r_{B_c^+} = 1.2$ and $\delta_{B_c^+} = 0°$.

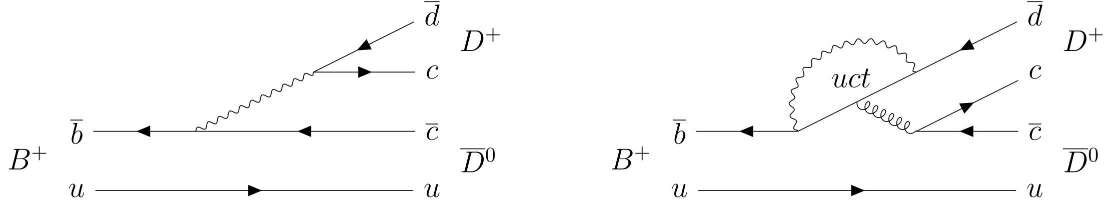
(b) The uncertainty on γ as a function of $r_{B_c^+}$ and $\delta_{B_c^+}$.

Figure 2.10: Toy studies to estimate the sensitivity to γ for $N(B_c^+ \to D_s^+ D^0) + N(B_c^+ \to D_s^+ \overline{D}^0) = 138$ and $\gamma = 70°$.

The decays of $B^0 \to D^+ D^-$, $B^0 \to D^0 \overline{D}^0$ and $B^+ \to D^+ \overline{D}^0$ are related by isospin symmetry. Their branching fractions and *CP* asymmetries are therefore expected to be consistent with one another within the isospin symmetry framework [76, 77].

As discussed in section 2.2.1, a non-zero *CP* asymmetry requires two interfering contributions with different weak and strong phases. Decays of $B^+ \to D_{(s)}^+ \overline{D}^0$ are dominated by tree-level $b \to c$ transitions, but also receive contributions from loop-level transitions and to a lesser extent, annihilation diagrams. The Feynman diagrams for the tree and loop-level transitions are shown in figure 2.11. The more similar in size the amplitudes are, the greater the interference. Hence the Cabibbo suppressed $B^+ \to D^+ \overline{D}^0$ decay is expected to have a larger *CP* asymmetry than the Cabibbo favoured $B^+ \to D_s^+ \overline{D}^0$ decay.

Table 2.4 lists some of the predictions for the *CP* asymmetries in the SM. The calculation of *CP* asymmetries is difficult to perform since it involves non-perturbative



(a) Tree-level

(b) Loop-level

Figure 2.11: Feynman diagrams for $B^+ \to D^+ \overline{D}^0$ decays. The diagrams for $B^+ \to D_s^+ \overline{D}^0$ decays are obtained by replacing the d quark with an s quark.

	Prediction for $\mathcal{A}^{CP}$ (%)				
Channel	Ref. [78]	Ref. [79]	Ref. [80]	Ref. [18]	Ref. [19]
$B^+ \to D_s^+ \overline{D}^0$	~ -0.1	0.3	$-0.26^{+0.05}_{-0.04}$	—	$[-0.39, -0.29]$
$B^+ \to D^+ \overline{D}^0$	$0.6^{+0.6}_{-0.1}$	-4.8	$4.4^{+1.0}_{-0.4}$	$[3.87, 6.03]$	—

Table 2.4: Predictions for $\mathcal{A}^{CP}(B^+ \to D_{(s)}^+ \overline{D}^0)$ in the SM.

hadronic parameters so often, it is non-trivial to compare an experimental measurement with the SM prediction. The presence of *CP* violation (whether SM or BSM) is much easier to determine and only requires a non-zero measurement for the *CP* asymmetry.

New physics contributions can enhance the *CP* asymmetries in these decays [18, 19, 80, 81]. Two models that allow particularly large enhancements are R-parity violating supersymmetry [18] and the fourth generation model [19]. All SM particles have R-parity of $+1$ while supersymmetric particles have R-parity of -1. Depending on the choice of couplings, R-parity violating supersymmetry predicts $\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0) = [-2.59, 11.05]\%$ [18]. In the fourth generation model, there are additional up-type (t') and down-type quarks (b'). Like u, c and t quarks, heavy t' quarks are expected to contribute to loop-level transitions and hence can significantly affect the branching fractions and *CP* asymmetries of doubly charmed B^+ decays. The fourth generation model predicts $\mathcal{A}^{CP}(B^+ \to D_s^+ \overline{D}^0) = [-12.02, -0.32]\%$ [19].

The most precise measurements of the *CP* asymmetry in $B^+ \to D^+ \overline{D}^0$ decays before those presented in this thesis are from the Belle [82] and BaBar [83] experiments. They measured $\mathcal{A}^{CP} = (0 \pm 8 \pm 2)\%$ [84] and $\mathcal{A}^{CP} = (-13 \pm 14 \pm 2)\%$ [85], respectively, where the first uncertainties are statistical and the second systematic. The combination of the two gives $\mathcal{A}^{CP} = (-3 \pm 7)\%$ [3]. The *CP* asymmetry in $B^+ \to D_s^+ \overline{D}^0$ decays has never previously been measured.

Chapter 3

The LHCb experiment

"It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong."
— Richard P. Feynman

The Large Hadron Collider beauty (LHCb) experiment [86] is a specialised heavy quark flavour detector collecting data at the LHC at CERN. Its primary physics goals are the precise measurement of *CP* violation parameters and indirect searches for new physics. This chapter describes CERN (section 3.1), the LHC (section 3.2) and the LHCb detector (section 3.3). An overview of the subdetectors dedicated to tracking (section 3.3.1) and particle identification (section 3.3.2) is given, followed by a discussion of the trigger (section 3.3.3) and the LHCb software (section 3.3.4).

3.1 CERN

In the aftermath of the second world war, a number of eminent scientists identified the need for Europe to have a world-class atomic physics research laboratory. Such a laboratory would provide a force for unity in post-war Europe and allow scientists to share the cost of increasingly expensive research facilities. In 1952, a provisional council (in French, the "Conseil Européen pour la Recherche Nucléaire" or "CERN") was set up to oversee the creation of the laboratory. The council selected Geneva as the site for the laboratory due to its central European location, Swiss neutrality during the war and the presence of a number of existing international organisations. The "Organisation

Européenne pour la Recherche Nucléaire" (European Organisation for Nuclear Research) was officially founded two years later [87], but the CERN acronym stuck and is still in use today.

CERN was originally founded with 12 member states, including the United Kingdom. It has since expanded to 22 member states and 8 associate member states, 3 of which are in the process of upgrading to full membership. In 2017, CERN employed 3,400 personnel (scientists, engineers, technicians, and administrators), with a further 14,000 associated with CERN but employed elsewhere [88]. CERN is the largest particle physics research laboratory in the world.

Since its establishment, CERN has been at the forefront of particle physics research. Some notable discoveries include:

- **Neutral currents** with the Gargamelle bubble chamber in 1973 [89, 90];
- **The W and Z bosons** with the UA1 and UA2 experiments in 1983 [91–94];
- **The number of light neutrino species** with the ALEPH, DELPHI, L3 and OPAL experiments in 1989 [95–98];
- **Direct CP violation** with the NA31 and NA48 experiments in 1988 and 1999, respectively [99, 100];
- **The Higgs boson** with the ATLAS and CMS experiments in 2012 [4, 5];
- **Pentaquarks** with the LHCb experiment in 2015 [101];
- **CP violation in charm hadrons** with the LHCb experiment in 2019 [34].

3.2 The LHC

The Large Hadron Collider [102] is the most powerful particle accelerator ever built. It was designed to accelerate protons to an energy of 7 TeV and collide them with a combined centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$.[1] The LHC was built by CERN over a period of 10 years, with construction starting in 1998 and finishing in 2008. The design, construction and subsequent operation of the LHC involved the collaboration of over

[1]For comparison, the second highest energy particle collider, the Tevatron, collided protons and antiprotons at $\sqrt{s} = 1.8\,\text{TeV}$ [103].

10,000 scientists based in institutions in over 100 countries and is one of the world's most ambitious scientific and engineering projects.

The LHC is 27 km in circumference and situated roughly 100 m underground in a circular tunnel that originally housed the Large Electron-Positron collider. It straddles the French-Swiss border near the city of Geneva, with the majority of its length on the French side. The depth varies from 175 m under the Jura mountains to 50 m towards Geneva with a slight gradient of 1.4°.

Although the LHC is primarily a *pp* collider, around one month of each data-taking year is dedicated to heavy-ion collisions. To date, there have been Pb-Pb, *p*-Pb and Xe-Xe collisions. Colliding heavy-ion beams are expected to be of a high enough temperature and energy density to produce a quark-gluon plasma, replicating the conditions that are thought to have existed a fraction of a second after the Big Bang.

The LHC is comprised of two beams that circulate protons (or ions) in opposite directions. The protons are brought to collision simultaneously at four points around the ring. Large caverns have been excavated around these collision points to house purpose-built detectors. Figure 3.1 shows a schematic of the LHC and the positions of the four detectors. The detectors are:

- **ATLAS** (A Toroidal LHC ApparatuS) [104]: the first of two general-purpose detectors. ATLAS was designed to search for the Higgs Boson and new particles predicted by beyond the Standard Model theories such as supersymmetry. ATLAS is the largest detector ever built for a particle collider at 46 m long, 25 m high, and 25 m wide. It weighs 7,000 tonnes.
- **CMS** (Compact Muon Solenoid) [105]: the second general-purpose detector. Although ATLAS and CMS share the same physics goals, different choices were made in the design of the two detectors. CMS is 21 m long, 15 m wide and 15 m high. It takes the crown for the heaviest detector ever built, weighing in at 14,000 tonnes (twice that of ATLAS).
- **ALICE** (A Large Ion Collider Experiment) [106]: a specialised heavy-ion detector. ALICE aims to study the properties of quark-gluon plasma in order to gain insight into quantum chromodynamics (QCD, the theory of the strong interaction). ALICE is 26 m long, 16 m high, 16 m wide and weighs 10,000 tonnes.
- **LHCb** (Large Hadron Collider beauty) [86]: a specialised heavy quark flavour detector. The primary physics goals of LHCb are the precise determination of *CP*

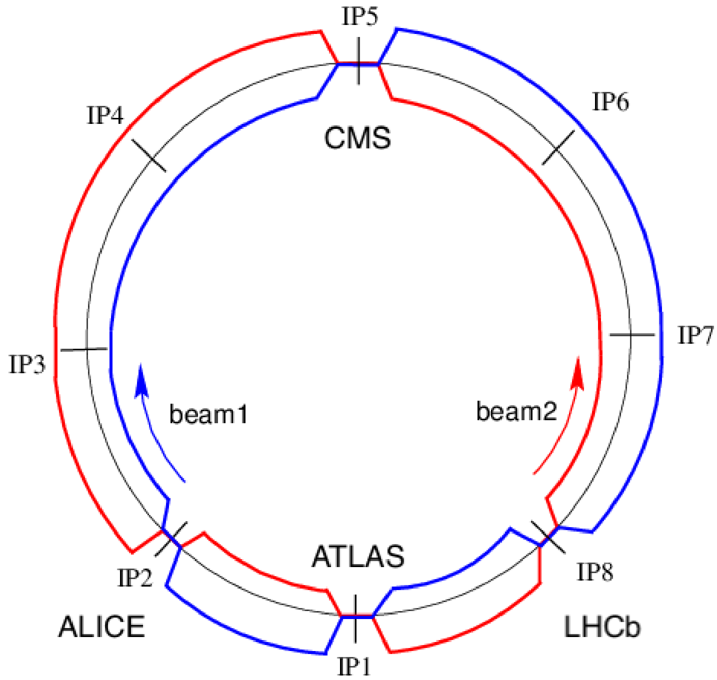


Figure 3.1: A schematic of the LHC, edited from [107]. LHCb is situated at IP8.

violation parameters and indirect searches for new physics. LHCb is 21 m long, 10 m high, 13 m wide and weighs 5,600 tonnes. The LHCb detector is described in more detail in section 3.3.

3.2.1 The accelerator complex

Before protons reach the LHC, they are first passed through a complex chain of accelerators, shown in figure 3.2, each of which successively boosts the speed of the protons in the beam. Protons begin life as hydrogen gas, which is passed through an electric field in order to strip off the electrons. The surviving protons are then fed into a linear accelerator, LINAC 2, where they reach an energy of 50 MeV. Next, the protons are passed into the Proton Synchrotron Booster, which brings them to an energy of 1.4 GeV. Following this, they are injected into the Proton Synchrotron (PS), CERN's first synchrotron and once the world's highest energy particle accelerator, where they reach an energy of 25 GeV. The final pre-accelerator is the Super Proton Synchrotron (SPS), where protons reach an energy of 450 GeV. Protons travel from the SPS to the LHC along two different paths (TI 2 and TI 8), which allows the beam to be split into opposite directions.

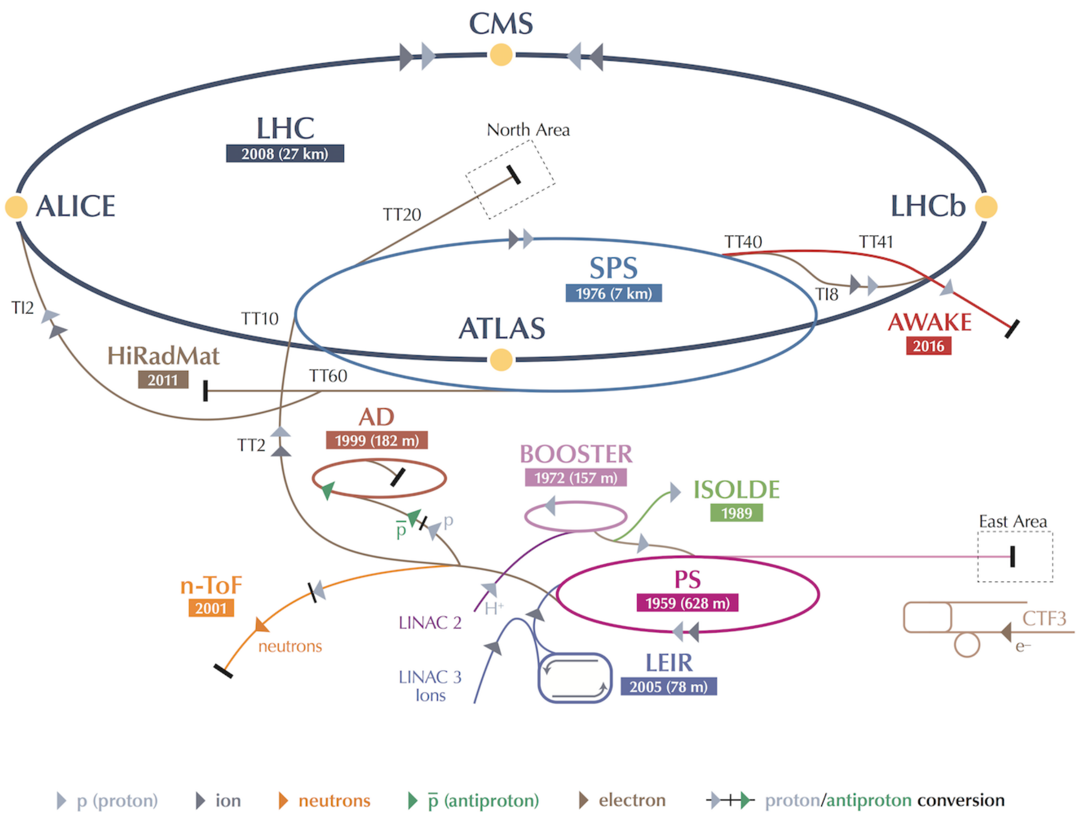


Figure 3.2: The CERN accelerator complex [108]. Protons are accelerated in LINAC 2 (up to 50 MeV), the Proton Synchrotron Booster (1.4 GeV), the Proton Synchrotron (25 GeV) and the Super Proton Synchrotron (450 GeV) before being injected into the LHC.

Once in the LHC, protons are accelerated up to their final collision energy using radio-frequency (RF) cavities [109]. RF cavities are metallic chambers that contain oscillating electromagnetic fields. In the LHC, each RF cavity is tuned to oscillate at 400 MHz. The amount of energy a proton gains in the cavity depends on its time of arrival, resulting in the beam being sorted into discrete groupings called “bunches”. Each beam is designed to contain 2808 bunches, each with around 10^{11} protons. The bunches are separated by 25 ns, giving a maximum collision rate of 40 MHz. It takes around 20 minutes for protons in the LHC to reach the design energy of 7 TeV, at which point they are travelling at a speed of $0.999999991c$ ($3\ \text{ms}^{-1}$ less than c).

In order to prevent collisions between the beams and gas molecules, the LHC beam pipes are kept in an ultra-high vacuum, with a pressure of around 10^{-10} mbar. The vacuum is maintained in two different ways, depending on whether the pipes are in the 24 km of arcing sections or 3 km of straight sections of the ring. In the straight sections, a "getter coating" that absorbs residual molecules when heated is used. The coating is applied to the walls of the beam pipes then heated to 300° at regular intervals in order to maintain a low pressure. This method can't be used for the arcs since the beam pipes are in contact with heat-sensitive magnetic coils so instead, cryogenic pumping is used. As the temperature drops to just 1.9 K, the gases condense and stick to the walls of the beam pipe.

Protons are steered around the LHC ring using 1,232 dipole magnets, which generate a uniform magnetic field of up to 8.3 T. The magnetic field is gradually ramped up as the protons accelerate to maintain the same beam trajectory. To increase the likelihood of a collision, the beams are focused using 392 quadrupole magnets. These consist of four magnetic poles positioned symmetrically around the beam pipe to focus horizontally or vertically in turn.

Once the beams are at the desired energy, magnets are used to redirect them such that they cross over and to further squeeze them at the point of collision. The four detectors then start collecting data, which typically lasts for around 12 hours until the beam is "dumped" (redirected into a graphite block designed to absorb the beam energy). A beam dump is initialised when the number of protons in the beam gets too low or to ensure the safety of the LHC.

The accelerators in the chain leading up to the LHC were used to make many of CERN's important discoveries in the latter half of the 20th century. For example, the SPS was instrumental in the discovery of the W and Z bosons [91–94]. The accelerators have since been re-purposed and improved to serve as the LHC injector chain. The LHC may one day act as a pre-accelerator for the next generation of particle accelerator, as has been proposed for the Future Circular Collider [110].

3.2.2 Running conditions

The two most important numbers for an accelerator are the collision energy and the luminosity. The energy determines what particles can be produced in the collisions and the luminosity how often. Although the LHC was completed a decade ago, the design energy of $\sqrt{s}$ =14 TeV has not yet been achieved. This is due to an incident in 2008, when a faulty electrical connection resulted in significant damage to the LHC [111], taking around a year to repair. As a result, early operation of the LHC has been more cautious. To date, there have been two operational runs. Run I took place during 2011 and 2012

Run	Year	$\sqrt{s}$ (TeV)	Integrated Luminosity (fb^{-1})		
			ATLAS	CMS	LHCb
I	2011	7	5.38	5.87	1.14
I	2012	8	23.2	23.2	2.19
II	2015	13	4.21	4.22	0.36
II	2016	13	38.5	41.0	1.88
II	2017	13	50.3	50.2	1.87
II	2018	13	64.9	66.8	2.46

Table 3.1: The centre-of-mass energy and integrated luminosity delivered by the LHC to the ATLAS, CMS and LHCb experiments in *pp* collisions for each data-taking year [114]. The luminosity collected by each experiment depends on the data-taking efficiency, which for LHCb is $\sim 90\%$ [115].

at a centre-of-mass energy of 7 and 8 TeV, respectively. After Run I, there was a long shutdown period to prepare the LHC to operate at higher energies. Run II took place from 2015 to 2018 at a centre-of-mass energy of 13 TeV. The running conditions of the LHC are summarised in table 3.1.

Instantaneous luminosity, $\mathcal{L}$, is a measure of the number of collisions that take place in a detector per cm^2 and per second [112]. It's determined by the number of protons in each bunch, N, the width of the bunch at the interaction point, σ, and the bunch crossing frequency, f,

$$\mathcal{L} = fN^2/(4\pi\sigma^2). \tag{3.1}$$

The LHC was designed to operate at an instantaneous luminosity of $10^{34}\,\text{cm}^{-2}s^{-1}$ for the ATLAS and CMS experiments and a reduced value of $10^{32}\,\text{cm}^{-2}s^{-1}$ for the LHCb experiment. ATLAS and CMS see an average of $\mathcal{O}(10)$ *pp* interactions per bunch crossing, while LHCb only sees $\mathcal{O}(1)$. This reduction is essential to reduce radiation damage, bring detector occupancies down to acceptable levels and remove ambiguities in determining which *pp* vertex a B meson originated from. The LHCb luminosity is kept constant by adjusting the overlap of the colliding beams throughout the data-collection period [113]. Table 3.1 shows the integration of the instantaneous luminosity over time for each of the detectors.

3.3 The LHCb detector

The LHCb detector [86,116–118] is designed for the study of flavour physics using hadrons containing a b or a c quark. Its primary physics goals are laid out in a roadmap document that identifies 6 key measurements, including the tree-level determination of the CKM angle γ [119]. Over the years, the LHCb physics programme has grown to be much broader than originally foreseen.

A diagram of the layout of the LHCb detector can be seen in figure 3.3. LHCb is a forward detector with an angular acceptance of 10-300 mrad in the horizontal plane and 10-250 mrad in the vertical plane. Acceptances in particle physics are often given in terms of pseudorapidity, defined as

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \tag{3.2}$$

where θ is the angle between a particle's momentum and the positive direction of the beam axis. In these coordinates, the LHCb acceptance is $2 < \eta < 5$. The minimum acceptance is dictated by the beam pipe, while the maximum acceptance was decided based on budgetary constraints and the limited space available in the cavern.

The shape of the LHCb detector is motivated by the predicted angular distribution of b quark production at the LHC. The dominant $b\bar{b}$ production mechanisms are gluon-gluon fusion, quark-antiquark annihilation and gluon splitting. The resulting $b\bar{b}$ pairs are predominantly produced at small angles with respect to the beam pipe and hence a significant fraction fall within the LHCb acceptance, as shown in figure 3.4. This is due to the fact that the interacting partons have highly asymmetric momenta in the laboratory frame. The angular distributions of the quarks from the $b\bar{b}$ pair are also highly correlated, so if one falls within the LHCb acceptance, the other is very likely to do so also. This is important for the study of neutral B hadrons, since the flavour at production (before mixing) can be determined from the other B hadron in the event.

The fragmentation fraction, f_q, is the fraction of b quarks that hadronise into a B_q hadron. The possibilities are: B^+ ($f_u \sim 40\%$), B^0 ($f_d \sim 40\%$), B_s^0 ($f_s \sim 10\%$), B_c^+ ($f_c \sim 1\%$, see section 2.3.2.2) or b-baryons ($f_{baryon} \sim 10\%$). The sum of all possibilities must equal 1,

$$f_u + f_d + f_s + f_c + f_{baryon} = 1. \tag{3.3}$$

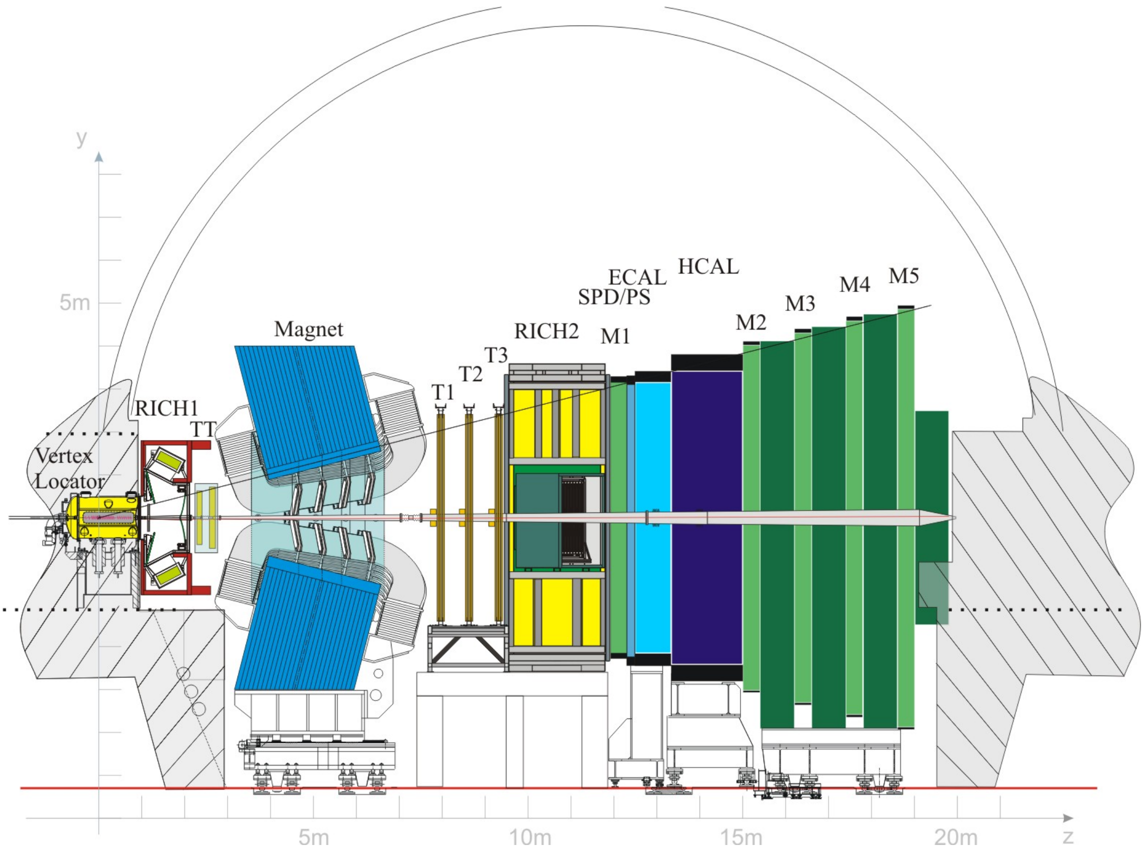


Figure 3.3: The LHCb detector shown side on [86]. Under the LHCb coordinate system, the z axis is along the beam pipe with the pp interaction point on the left-hand side. The y axis points from the floor of the cavern up to ground level and the x axis is perpendicular to the plane of the figure.

The fragmentation fractions are not constants, but depend on the B production spectrum: transverse momentum, $p_T = \sqrt{p_x^2 + p_y^2}$, and η [3].

The LHCb detector is made up of several different subdetectors, each of which is designed to measure different particle properties. The information from each of these subdetectors is combined in order to make a measurement of the charge, momenta and species of all the particles involved in the event. Section 3.3.1 describes the subdetectors dedicated to tracking and section 3.3.2 those dedicated to particle identification. The trigger that is used to reduce the collision rate of $\mathcal{O}(\text{ MHz})$ down to $\mathcal{O}(\text{ kHz})$ with as little loss of signal as possible is described in section 3.3.3. Section 3.3.4 gives an overview of LHCb software.

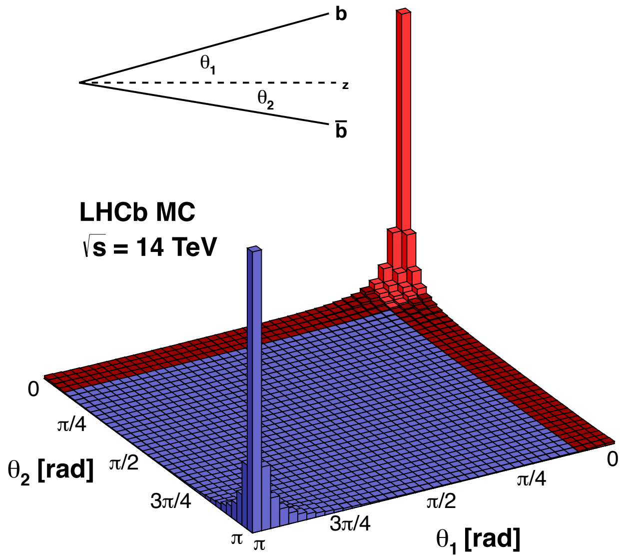


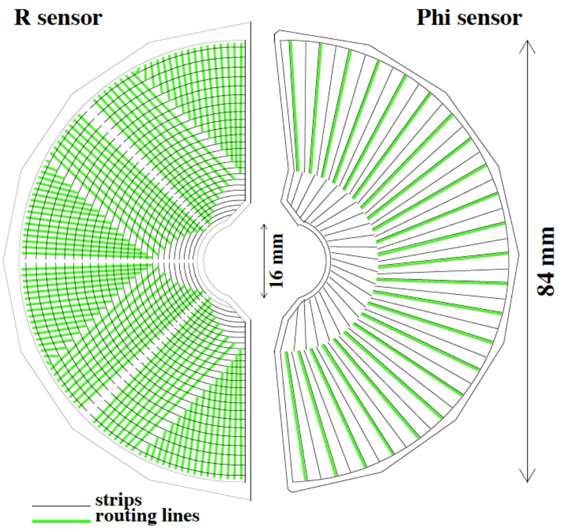
Figure 3.4: The angular distribution of $b\bar{b}$ production in simulated pp collisions at 14 TeV [120]. The LHCb acceptance is shown in red.

3.3.1 Tracking

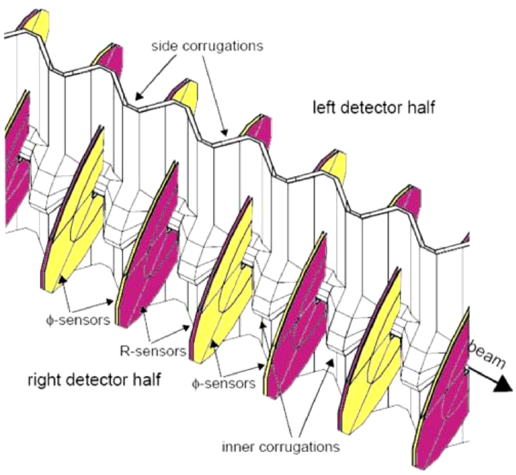
The LHCb tracking system is designed to provide efficient reconstruction of the paths of charged particles through the detector. When combined with a magnetic field, this provides important information on charge and momentum. Precise tracking information is also essential to determine how far a particle travels before decaying (giving its lifetime) and which tracks originate from the same decay vertex (giving its decay mode). The LHCb tracking system consists of the vertex locator (section 3.3.1.1), the dipole magnet (section 3.3.1.2) and tracking stations just before and after the magnet (section 3.3.1.3). Track reconstruction and the performance of the tracking system is discussed in section 3.3.1.4.

3.3.1.1 Vertex locator

The VErtex LOcator (VELO) [121] surrounds the pp interaction point and is used to make precise measurements of the tracks originating from the collisions. Decays of B mesons are typically characterised by their long lifetimes, which for highly boosted particles at the LHC, corresponds to large flight distances of $\sim 1\,\text{cm}$. The VELO is



(a) The R-sensors consist of azimuthal strips, whereas the Φ-sensors consist of radial strips.



(b) The R- and Φ-sensors are placed on the front and back of the modules in an alternating pattern. The corrugated RF-foil separates the VELO vacuum from the LHC vacuum.

Figure 3.5: Diagram of R- and Φ-sensors in the VELO subdetector [121].

therefore important in distinguishing the pp interaction point (the primary vertex, PV) from the displaced B decay vertex (the secondary vertex, SV) and identifying a key signature of B decays.

The VELO is composed of a series of 21 stations, each with two silicon modules arranged approximately 1 m along the beam direction. To achieve the best resolution on the position of each decay vertex, the sensors need to be positioned as close to the interaction point as possible. This was accomplished by designing the VELO in two halves that can be drawn in or retracted on demand. This enables the VELO to keep a safe distance of 29 mm during the injection phase when the beam is defocused, and once stable beams are established, to be moved to just 8 mm from the interaction point. These positions are referred to as “open” and “closed”, respectively.

Each silicon module is made up of two different types of sensor positioned on opposing sides, shown in figure 3.5. The R-sensors measure the radial distance from the beam, while the Φ-sensors measure the azimuthal angle. There is a small overlap between the two modules when the VELO is closed to ensure the full angular coverage is maintained. The gap in the middle allows the beam to pass through.

The displacement of the sensors along the beam axis, shown in figure 3.6, enables the z coordinate to be determined. The spacing of the sensors is chosen such that all

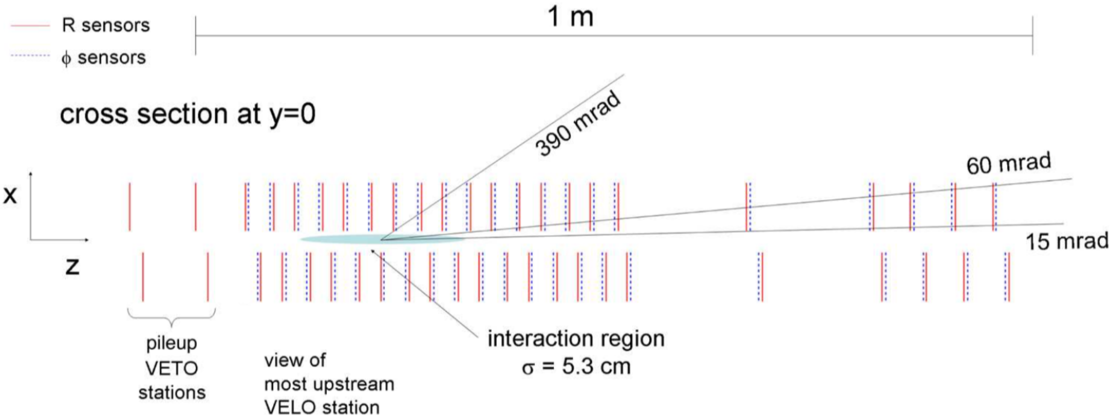


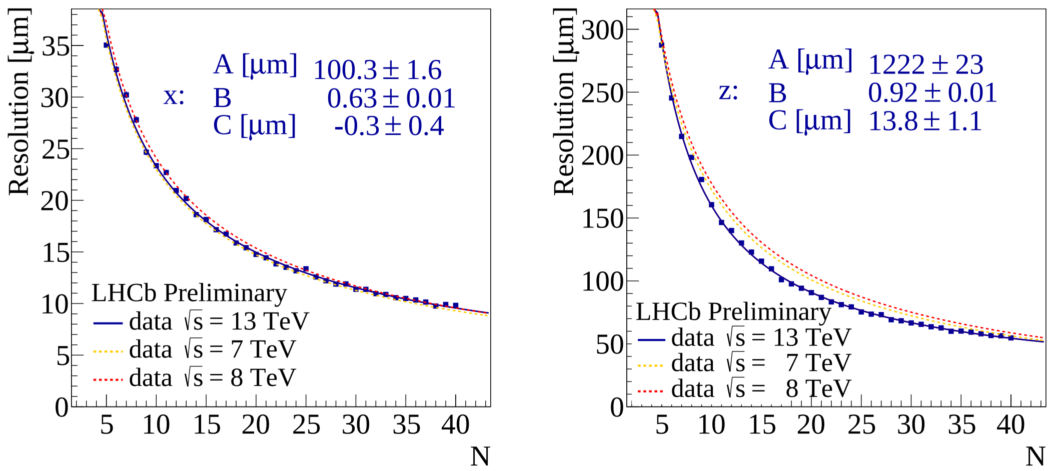
Figure 3.6: The distribution of sensor modules along the beam direction in the VELO subdetector [121].

tracks inside the LHCb acceptance cross at least 3 VELO stations. There are 2 sensors located in the backwards direction (at negative z), known as the "pileup VETO stations". These stations contain R-sensors only and were designed to estimate the number of pp interactions in each bunch crossing, but ended up primarily being used for luminosity measurements.

The VELO is inside an aluminium enclosure, known as the "RF-box". It serves three purposes:

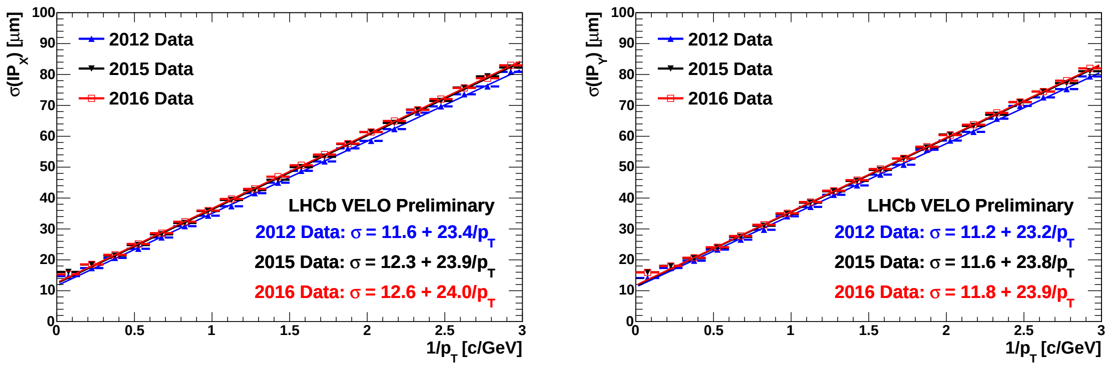
- It separates the vacuum in the VELO from the LHC vacuum. The outgassing of the VELO sensors would otherwise pollute the LHC vacuum.
- It allows a mirror current to travel parallel with the beam. This prevents impedance disruptions to the LHC beam.
- It protects the VELO electronics from RF interference caused by the bunch structure of the beam.

The surface of the box facing the beam is a thick 0.3 mm corrugated sheet, known as the "RF foil" and is shown in figure 3.5b. The corrugations in the foil allow for an overlap between the two detector halves. Particles originating from the pp interaction point travel through the RF-foil and undergo multiple scattering before encountering the VELO. This limits the accuracy with which the position of the PV can be determined. Figure 3.7 shows the PV resolution measured by the VELO for different centre-of-mass energies. The resolution improved from Run I to Run II due to the implementation of real-time alignment and calibration.



(a) PV resolution in the x direction. **(b)** PV resolution in the z direction.

Figure 3.7: PV resolution as a function of the number of associated tracks for $\sqrt{s}$ =7, 8 and 13 TeV, measured using 2011, 2012 and 2015 data, respectively [122].



(a) IP resolution in the x direction. **(b)** IP resolution in the y direction.

Figure 3.8: IP resolution as a function of $1/p_T$ for 2012, 2015 and 2016 data [122].

Another important measure of the VELO performance is the impact parameter (IP) resolution. The IP is defined as the distance of closest approach between the PV and the extrapolated particle trajectory. Particles, such as B mesons, that originate from the PV are expected to have a small IP, however, tracks that originate from displaced vertices, such as those from the SV, are expected to have large values. The IP is therefore a useful discriminating variable in the selection of B decays. Figure 3.8 shows the IP resolution as a function of $1/p_T$ for 2012, 2015 and 2016 data. The linear relationship is due to multiple scattering.

3.3.1.2 Magnet

In the presence of a magnetic field, the paths of charged particles become curved. The direction of this curvature reveals whether the particle is positively or negatively charged and the radius of curvature determines the particle's momentum. A high momentum particle will take a relatively straight path through the detector, whereas a low momentum particle will move in tight spirals.

The magnetic field in the LHCb detector is supplied by a dipole magnet [123], located between the TT and T1-T3 tracking stations. The region before the magnet is referred to as "upstream" and the region after is referred to as "downstream". The magnet has a bending power of 4 T m in the y direction for a 10 m track length. It consists of two saddle-shaped aluminium coils, arranged inside an iron yoke. The magnetic flux generated by the two coils is shaped and guided by the yoke. The maximum magnetic field strength in the centre of the magnet is 1.1 T.

The LHCb dipole magnet is a "warm" magnet that doesn't require cryogenic cooling. A warm magnet was chosen since it allows a rapid ramp-up of the field, enabling the polarity to be reversed on a roughly biweekly basis such that similar amounts of data are collected in the *MagUp* (positive y direction) and *MagDown* (negative y direction) configurations. Collecting equal amounts of data in each polarity is necessary in order to control detector asymmetries at the 10^{-3} level in studies of CP violation [124].

3.3.1.3 Tracking stations

The remainder of the tracking system consists of 4 stations: the Tracker Turicensis (TT) [116] and the T-stations (T1-T3), located just before and after the magnet, respectively. Each T-station is divided into two regions: the inner tracker (IT) [125], which surrounds the beam pipe, and the outer tracker (OT) [126–128], which covers the rest of the LHCb acceptance. The TT and IT make use of silicon technology, since regions of the detector close to the beam pipe experience high particle occupancies and therefore require fine spatial resolutions. However, silicon is expensive and so different detector technology is used in the OT where occupancies are lower.

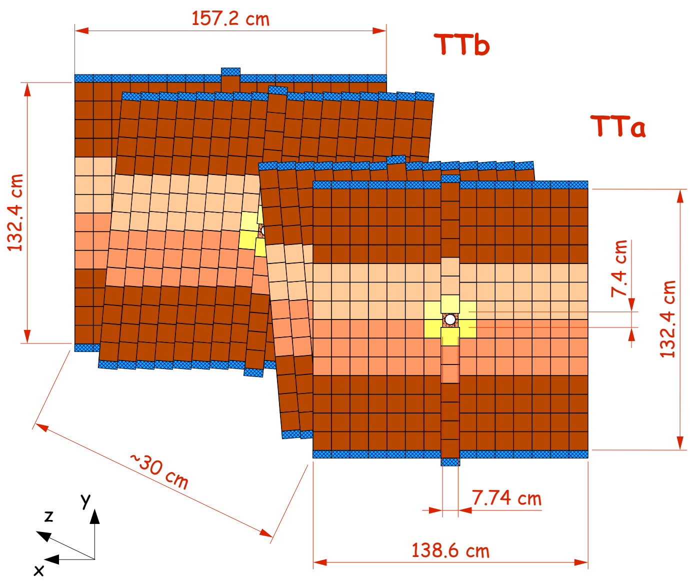


Figure 3.9: Schematic of the TT subdetector [129]. The first and last layers are vertical, while the second and third are rotated by $-5°$ and $+5°$, respectively.

Tracker Turicensis

The Tracker Turicensis (TT) is a crucial element in track reconstruction. Long-lived particles, such as K^0_S or Λ, often decay after the VELO and are therefore reconstructed using hits in the TT and T-stations. The TT is also useful in reducing the background from "ghost" tracks. Ghost tracks arise when tracks in the VELO are matched with the wrong tracks in the T-stations and can be mitigated by requiring the extrapolated track corresponds to hits in the TT. Additionally, the TT is useful in the reconstruction of low momentum particles that are swept out of the LHCb acceptance by the magnet. Finally, the TT experiences a small fringe field from the magnet, which can be used to make relatively imprecise measurements of track p_T for use in the trigger.

The TT is 150 cm wide and 130 cm high, covering the entire LHCb acceptance. It consists of 4 detector layers made up of silicon microstrip sensors. The 4 layers, shown in figure 3.9, are arranged in 2 pairs along the beam axis, with a gap of around 30 cm between them. The first and last layers are vertical and designed to measure the track x coordinate, while the second and third layers are rotated by $-5°$ and $+5°$, respectively, to allow a 3D track reconstruction to be performed. The TT has a resolution of 50 μm for a single hit.

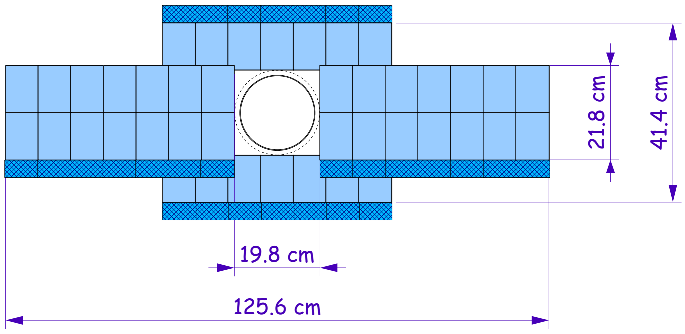


Figure 3.10: Schematic of the IT subdetector [129]. The IT is arranged in a cross-shape around the beam pipe.

Inner tracker

The inner tracker (IT) uses the same silicon microstrip technology as the TT and has the same 50 μm hit resolution. The IT consists of 4 detector layers each with 4 boxes in a cross-shaped pattern around the beam pipe, as shown in figure 3.10. Like the TT, the layers are positioned at slight angles of (0°, −5°, +5°, 0°) to aid track reconstruction. The IT is 120 cm wide and 40 cm high and although it only covers 1.3% of the geometrical acceptance, 20% of the tracks pass through it. The layout and dimensions of the IT were determined from the requirement that the OT occupancy is below 10%.

Outer tracker

Unlike the other components of the LHCb tracking system, the outer tracker (OT) is not made out of silicon, but is instead a gaseous straw-tube detector. The OT extends from the IT out to the edge of the LHCb acceptance, an area of approximately $5 \times 6\,\mathrm{m}^2$. It consists of 12 double layers of straw-tubes, as shown in figure 3.11. Each of the 3 T-stations contains 4 layers, which are arranged with the same geometry as for the TT and IT layers. Each straw-tube is 2.4 m long with a 4.9 mm inner diameter and filled with a mixture of argon, carbon dioxide and oxygen, which results in a fast drift time of less than 50 ns. The OT has a hit resolution of 200 μm.

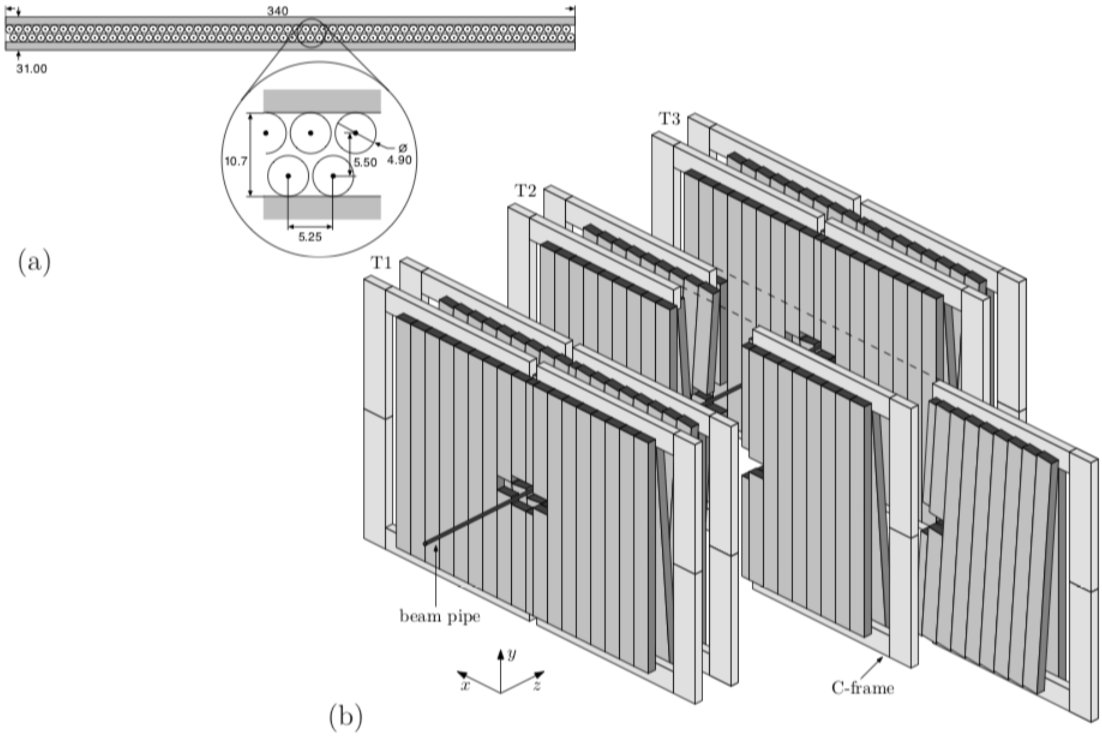
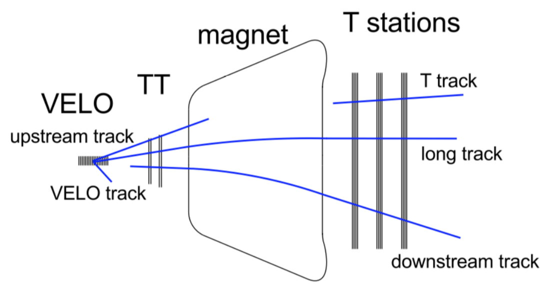


Figure 3.11: Schematic of the OT subdetector [127]. **(a)** The double layer structure. **(b)** Arrangement of the straw-tubes in the T-stations.

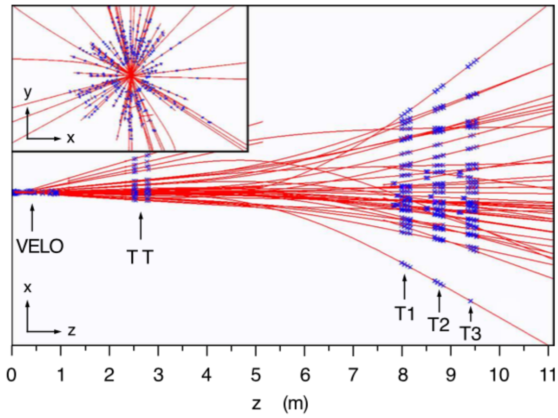
3.3.1.4 Track reconstruction and performance

Hits in the VELO, TT, IT and OT are combined using track reconstruction algorithms to identify the path a charged particle took through the LHCb detector. The first step in track reconstruction is to identify a track seed, which is an initial track candidate made from hits in the VELO or the T-stations where the magnetic field strength is low. The seeds are then passed to a Kalman filter fit [130], which iteratively adds hits from other subdetectors whilst simulataneously updating the predictions for the track trajectories. A Kalman filter fit was chosen since it accounts for multiple scattering and corrects for the energy lost as particles move through detector material. The reconstruction algorithms aim to find all tracks in the event, not just those from a decay of a B hadron. Tracks are divided into types, depending on which subdetectors they passed through:

- **Long tracks**: these pass through the full tracking system, from the VELO to the T-stations. As this is the maximum possible information available for reconstruction, these tracks have the best momentum resolution and are therefore the most useful in identifying B hadron decays.



(a) The subdetectors traversed by different track types [131].



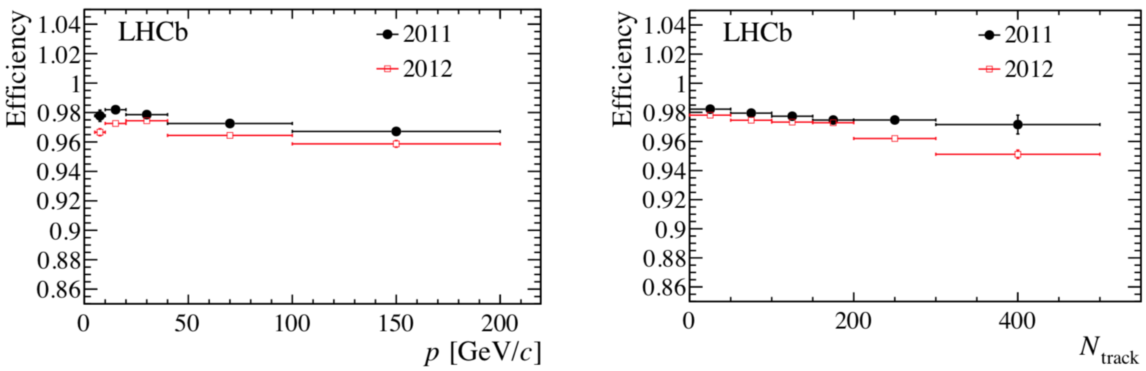
(b) The reconstructed tracks with their assigned hits [86].

Figure 3.12: A schematic of the different track types: long, upstream, downstream, VELO and T-tracks.

- **Upstream tracks**: these pass through the VELO and the TT only. In general, these are low momentum tracks that are swept out of the LHCb acceptance by the magnetic field. They may also be used for B hadron reconstruction, although their momentum resolution is poor.
- **Downstream tracks**: these pass through the TT and T-stations only. These are typically formed by long-lived particles, such as K^0_S or Λ, that decay outside the VELO.
- **VELO tracks**: these are measured in the VELO only. They are typically large angle tracks that fall outwith the LHCb acceptance or backwards tracks. These tracks are useful in the reconstruction of the PV.
- **T-tracks**: these are measured in the T-stations only. They are likely to have been produced in secondary interactions with the detector material and are of fairly limited use.

The different track types can be seen in figure 3.12a. The hits that are used to reconstruct the tracks are illustrated in figure 3.12b.

The performance of the LHCb tracking system is quantified by the tracking efficiency and the momentum resolution. The tracking efficiency is defined as the probability that the trajectory of a charged particle that has passed through the full tracking system is reconstructed. It has been measured in Run I data using $J/\psi \to \mu^+\mu^-$ decays and is



(a) Tracking efficiency as a function of momentum. **(b)** Tracking efficiency as a function of the number of tracks in the event.

Figure 3.13: The tracking efficiency for 2011 and 2012 data [118].

shown in figure 3.13. The efficiency is above 95% [118]. The momentum resolution, $\delta p/p$, varies from 0.5% for low momentum particles (less than 20 GeV/c), to 1.0% for particles around 200 GeV/c [118].

3.3.2 Particle identification

The ability to distinguish between different particle species is an important tool in the selection of B hadron decays at LHCb. The separation of kaons and pions is particularly useful in the analyses considered in this thesis since the final states considered are composed entirely of these two types of particles. The subdetectors involved in particle identification (PID) are two ring imaging Cherenkov detectors (section 3.3.2.1), electromagnetic and hadronic calorimeters (section 3.3.2.2) and the muon system (section 3.3.2.3). Information from each of these subdetectors is used to calculate the likelihood that each track is a kaon, pion, proton, muon and electron.

3.3.2.1 Ring imaging Cherenkov detectors

The Ring Imaging CHerenkov (RICH) detectors [132–134] are primarily used to distinguish between different charged hadrons, *i.e.* kaons, pions and protons, over a momentum range of 1-100 GeV/c. There are two RICH detectors in LHCb. RICH1 is located between the VELO and the TT and RICH2 is located between the T-stations and the first muon station.

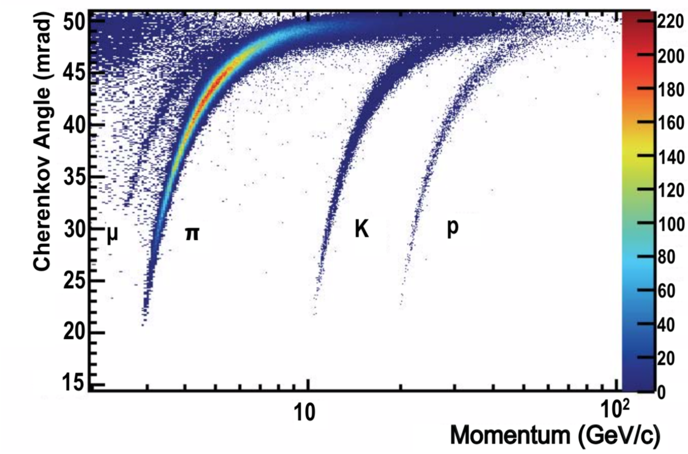


Figure 3.14: The Cherenkov angle as a function of momentum in C_4F_{10} gas in RICH1 [133]. The Cherenkov bands for muons, pions, kaons and protons are clearly visible.

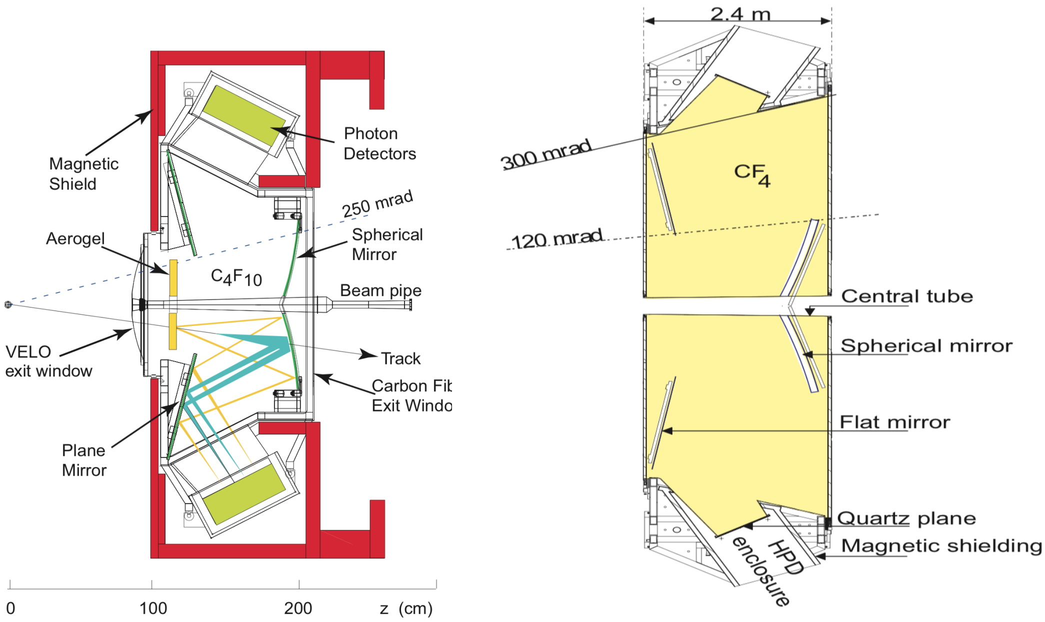
The physics principle underlying the RICH detectors is known as the Cherenkov effect [135]. When a charged particle travels faster than the local speed of light in a medium, it emits electromagnetic radiation known as Cherenkov light. The angle at which the light is emitted, θ_C, is related to the speed at which the particle is travelling relative to the speed of light according to

$$\cos\theta_C = \frac{1}{n\beta}, \tag{3.4}$$

where $\beta = v/c$ and n is the refractive index of the material. By combining a measurement of θ_C (giving v) with a measurement of momentum from the tracking system, the rest mass and hence identity of the particle can be determined. Figure 3.14 illustrates the distinguishing power Cherenkov light has for different particle species. Heavier particles require greater momenta in order to be above the speed threshold to emit Cherenkov light, which results in discrete bands separated in momentum for each particle type.

RICH1

The RICH1 detector provides PID information for low momentum particles (up to 60 GeV/c) and covers the full LHCb acceptance. It is positioned upstream so that it can measure low momentum tracks before they are swept out of the LHCb acceptance by the magnet. The layout of RICH1 can be seen in figure 3.15a. RICH1 uses two mediums: aerogel and "fluorobutane" (C_4F_{10}) gas, chosen because their refractive indices



(a) RICH1 **(b)** RICH2

Figure 3.15: Schematic of the two RICH detectors [86].

are well suited to the momentum range of interest. The cone of Cherenkov light emitted by a particle moving through the detector is focused and redirected out of the LHCb acceptance by a series of spherical and flat mirrors. Photon detectors are then used to measure the radius of the ring, from which θ_C can be determined. Many particles travel through the RICH detectors at once, creating a series of overlapping rings.

In Run II, the aerogel was removed. Although the aerogel provided PID information for kaons below the C_4F_{10} Cherenkov threshold, it compromised the overall PID performance by producing too many photons in the high multiplicity environment of RICH1. Removing the aerogel also greatly improved the speed of the RICH reconstruction since it reduced the number of photon candidates by more than a factor 2.

RICH2

The RICH2 detector provides PID information for high momentum particles (above 15 GeV/c). RICH2 has a reduced acceptance of ± 120 mrad horizontally and ± 100 mrad vertically, which is well suited for high momentum particles since these are predominantly produced close to the beam pipe. The layout of RICH2 can be seen in figure 3.15b. The medium in RICH2 is CF_4 gas, which has a lower refractive index in order to reduce the number of tracks above the Cherenkov threshold and hence the number of photons produced. As in RICH1, RICH2 has a series of mirrors to focus and redirect the light onto photon detectors.

Performance

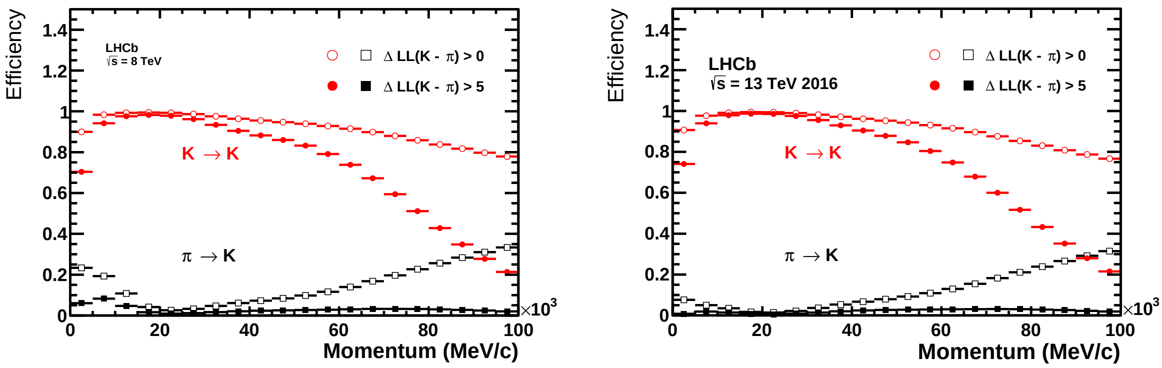
The trajectory of a particle through the RICH detectors is known from the tracking system. Therefore, it is possible to predict the expected distribution of Cherenkov photons for a given mass hypothesis. This prediction is compared to the observed distribution in order to calculate the likelihood of the hypothesis. The mass hypotheses considered are: pion, kaon, proton, muon and electron. As pions are the most abundant track type in the LHCb detector, all tracks are initially assumed to be pions and the global likelihood is calculated. Then for each track in turn, the likelihood is recalculated with each of the different mass hypotheses. The track is assigned the hypothesis that maximises the likelihood. The algorithm is described in detail in [136].

A common selection variable is the difference in log likelihood between the X (K, p, μ, e) hypothesis and the pion hypothesis,

$$\mathrm{DLL}_{X\pi} = \Delta \log \mathcal{L}(X - \pi) = \log \mathcal{L}(X) - \log \mathcal{L}(\pi). \quad (3.5)$$

If the particle is travelling below the Cherenkov threshold then the RICH detectors can't distinguish between different particle types and $\mathrm{DLL}_{X\pi}$ is defined to be equal to 0. In Run II, there are more cases like this due to the removal of the aerogel.

The performance of the RICH detectors is quantified in terms of the kaon identification efficiency and the pion misidentification rate, shown in figure 3.16. In Run I, with a loose requirement of $\mathrm{DLL}_{K\pi} > 0$, the kaon identification efficiency is $\sim 95\%$ with a pion misidentification rate of $\sim 10\%$ when averaged over momentum [133]. Tightening the cut to $\mathrm{DLL}_{K\pi} > 5$ results in a kaon identification efficiency of $\sim 85\%$ and a pion misidentification rate of $\sim 3\%$. The PID performance is not just dependent on momentum, but also the number of tracks in the event and the length of the track in the RICH detectors (greater at larger values of η).



(a) 2012 (b) 2016

Figure 3.16: Performance of the RICH detectors as a function of momentum for different $\mathrm{DLL}_{K\pi}$ requirements [137].

3.3.2.2 Calorimeters

The calorimeter system [138], located just after the first muon station, is designed to measure the positions and energies of hadrons, electrons and photons. It is composed of four components: a scintillating pad detector (SPD), a preshower detector (PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The interactions of the different particle types with the components of the detector are illustrated in figure 3.17. The calorimeter system performs several functions: the measurement of the transverse energy is fed into the trigger; it enables the reconstruction of π^0 and γ; and provides the identification of electrons. The calorimeter system is particularly important for neutral particles since these don't interact with the tracking system.

The SPD and the PS are in place to provide information for background rejection in the ECAL. They are scintillating pad detectors, separated by a 15 mm sheet of lead. The SPD identifies charged particles and is used to improve the separation of electrons and photons. The lead layer in the middle of the two detectors is used to initiate electromagnetic showers from electrons and photons via bremsstrahlung radiation and pair production. The charged particles created in the showers are then measured in the PS, which enables the separation of electrons and charged pions.

The ECAL is designed to measure the energy of electromagnetic showers. It makes use of the "shashlik" layout, in which layers of an absorber (in this case lead) are alternated with scintillating materials. Charged particles from the showers initiated in the lead create light as they pass through the scintillating layers. The light is transmitted to a

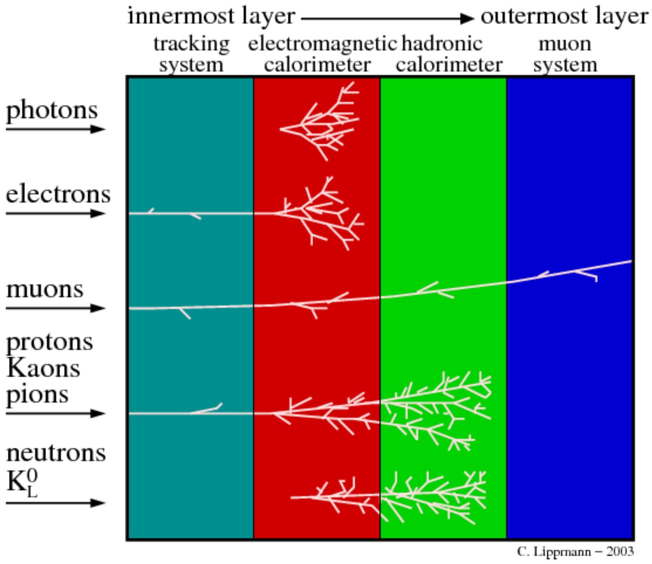


Figure 3.17: The interactions of photons, electrons, charged hadrons and neutral hadrons in different parts of the detector (the tracking system, ECAL, HCAL and muon system) [139].

photomultiplier by wavelength-shifting fibres. The 66 lead and scintillating layers have a thickness of 42 cm (or 25 radiation lengths), chosen such that the showers from high energy photons are fully contained. The clusters of energy deposited in the ECAL are summed to calculate the initial particle energy. The energy resolution is given by [140],

$$\frac{\sigma_E}{E} = \frac{8.5\% - 9.5\%}{E} \oplus 0.8\%, \tag{3.6}$$

where E is measured in GeV.

The HCAL measures the energy deposited by strongly interacting particles (pions, kaons, protons and neutrons). Its main purpose is to detect the presence of high p_T hadrons, which is a signature of B decays, for use in the trigger. The HCAL consists of plates of iron and scintillating material. Unlike the ECAL, the HCAL doesn't have a sufficient depth to fully contain energetic hadronic showers, so the energy measured is of limited use beyond the trigger. The energy resolution of the HCAL is given by [141],

$$\frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{E} \oplus (9 \pm 2)\%, \tag{3.7}$$

where E is measured in GeV.

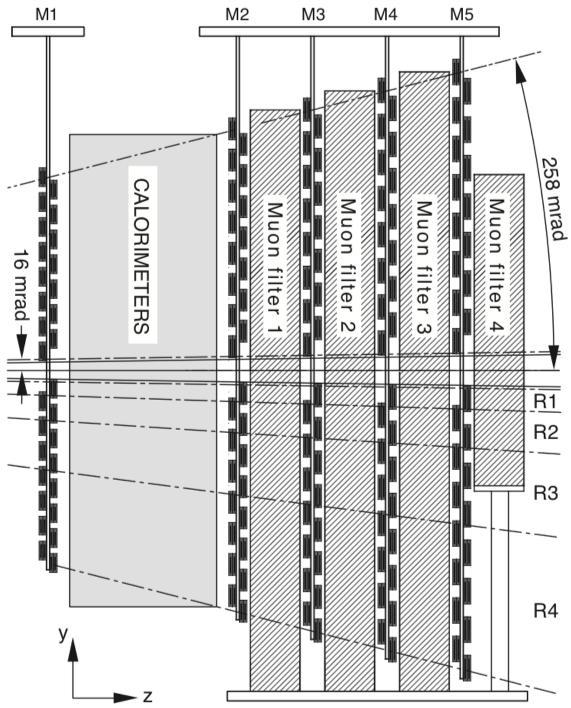


Figure 3.18: Schematic of the muon system [86]. M1 is located before the calorimeter system, M2-M5 are located after.

3.3.2.3 Muon system

Muons have a high penetrating power due to their relatively large mass and the fact that they don't interact via the strong force. Unlike most other particles, they are not stopped by the calorimeter system. Therefore, the muon system [142] is positioned at the very end of the detector where muons are the only particle likely to register a signal. The muon system has two main functions: it feeds a measurement of muon p_T into the trigger and provides muon identification information for use in the trigger and offline analysis.

The muon system consists of 5 stations, M1-M5, as shown in figure 3.18. M1 is positioned before the calorimeter system to improve the muon p_T measurement. M2-M5 are positioned after the calorimeter system and interwoven with iron absorbers 80 cm thick to select high momentum muons. The minimum momentum required for a muon to pass through all 5 stations is $\sim 6\,\mathrm{GeV}/c$.

The full muon system comprises 1,380 chambers and covers a total area of $435\,\mathrm{m}^2$. Of these, 1,368 are multiwire proportional chambers filled with a mixture of argon, carbon dioxide, and tetrafluoromethane (CF_4). The inner region of M1 contains 12 triple-Gas-Electron-Multiplier chambers to cope with the high particle density before the calorimeter system.

The M1-M3 stations have a high spatial resolution in the bending plane of the magnet (the x direction). These define the track direction and achieve a p_T resolution of 20%. The M4 and M5 stations have a limited spatial resolution and are primarily to identify highly penetrating muons. The dimensions of the muon stations scale with the distance from the interaction point to ensure the full LHCb acceptance is covered.

3.3.3 Trigger

Only around 1% of pp interactions are expected to produce a $b\bar{b}$ pair. Of these, only 15% of events will contain a B decay with all final state particles in the LHCb acceptance. Furthermore, the branching fractions of B decays used in the study of *CP* violation are typically $\mathcal{O}(10^{-3})$.[2] The bunch crossing rate of 40 MHz is far too high to be read out by the LHCb detector, therefore it is essential to have a trigger system [143–145] that can identify potentially interesting events to save for offline analysis. The trigger reduces the rate by 4 orders of magnitude, writing out a few kHz to storage.

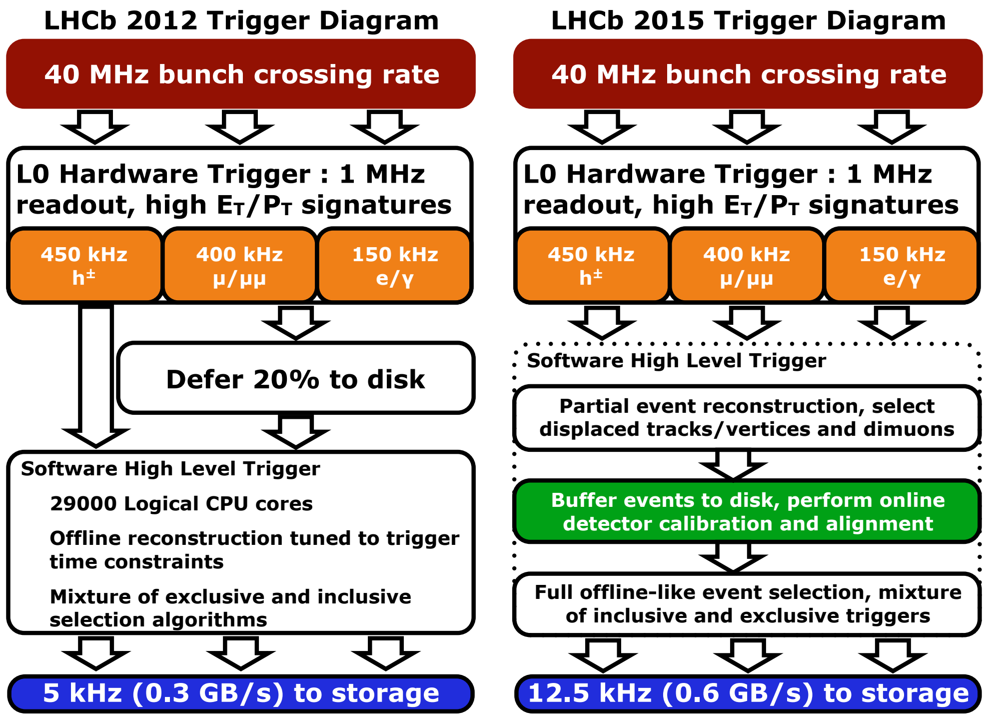
Decays of B hadrons have a number of characteristic signatures that are exploited in the trigger: long lifetimes result in B decay vertices displaced by $\sim 1\,\text{cm}$ from the pp interaction point; relatively large masses of over $1\,\text{GeV}/c^2$ result in high p_T daughters; and several key channels contain muons in their final state, which can cleanly be identified in the muon system. These properties enable B hadron decays to be efficiently selected while suppressing the otherwise overwhelming QCD background.

The trigger is made up of two stages: the Level 0 (section 3.3.3.1) trigger implemented in hardware and the high level trigger (section 3.3.3.2) implemented in software. Diagrams of the trigger structure in Run I and II can be seen in figure 3.19.

3.3.3.1 Level 0

The Level 0 (L0) trigger is used to reduce the 40 MHz bunch crossing rate down to 1 MHz, at which point the full LHCb detector can be read out. This requires a decision to be

[2] The rarest decay studied in this thesis is $B_c^+ \to D^+ D^0$, which is expected to have a branching fraction of $\mathcal{O}(10^{-7})$.



(a) Run I (b) Run II

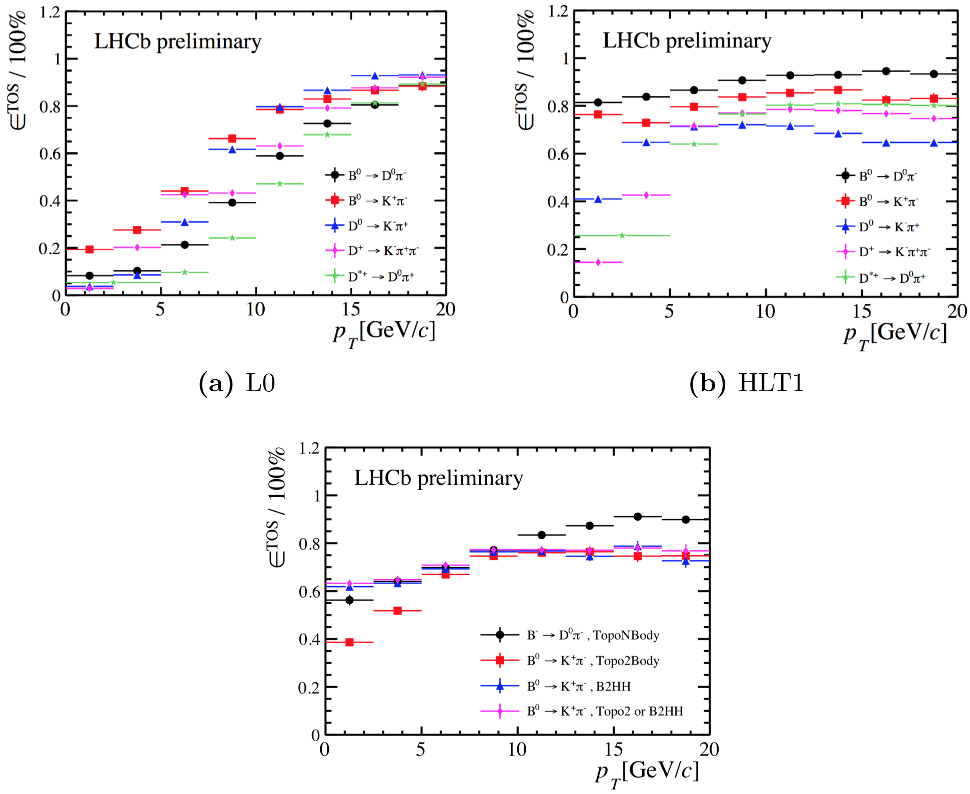
Figure 3.19: Diagram of the trigger structure [146].

made in under 4 μs. The only subdetectors capable of supplying information fast enough are the muon systems and the calorimeters. L0 therefore makes its decision based on these two subdetectors alone, selecting muons with high p_T or high E_T deposits in the calorimeters.

The L0 trigger requirements vary depending on the type of particle. The hadron trigger selects events in which $E_T > 3.68\,\text{GeV}$ is deposited in the HCAL, resulting in an output rate of 450 kHz (45% of the total L0 rate). The efficiency of the L0 hadron trigger for several hadronic decays is shown in figure 3.20a. Final states with soft (low momentum) particles, such as the pion from the $D^{*+} \to D^0\pi^+$ decay, have a reduced efficiency. The largest inefficiencies in the trigger chain occur at L0, particularly for hadronic decays.

3.3.3.2 High level trigger

The high level trigger (HLT) is implemented in C++ and runs on a dedicated computer farm. It is divided into two stages: HLT1 and HLT2. HLT1 performs a partial event



(a) L0 **(b)** HLT1

(c) HLT2

Figure 3.20: Trigger efficiencies for several hadronic decays as a function of mother p_T in 2012 data [144]. Triggered On Signal (TOS) indicates the trigger fired due to the signal candidate and not some other part of the event.

reconstruction and reduces the rate from 1 MHz to $\sim$70 kHz. Although the full detector is able to be read out at 1 MHz, only a partial event reconstruction can be performed due to the limited computing resources available. HLT1 reconstructs charged particle trajectories using information from the VELO and tracking stations.

HLT1 uses an "inclusive" trigger strategy [147, 148], whereby the trigger lines only require part of the B decay to be reconstructed. These lines can therefore select efficiently across the full range of B topologies. HLT1 requires there to be a single high p_T track with a good track fit quality that is well displaced from all pp interaction points. The HLT1 efficiency for several hadronic decays is shown in figure 3.20b.

The rate of events passed to HLT2 is low enough that a full event reconstruction can be performed. This means that trigger lines can be developed to select specific decays (known as "exclusive" lines). In addition to these, HLT2 has a set of inclusive lines that are designed to select decays of a similar topology [149]. These topological lines are designed to trigger efficiently on any B decay with at least 2 charged daughters. The efficiency of the 2-body topological trigger is shown in figure 3.20c.

HLT2 reduced the rate to 3 kHz in 2011, 5 kHz in 2012 and 12.5 kHz in Run II, which was then written to storage. The increase from 2011 to 2012 is largely due to writing 20% of the L0 output directly to disk, in a process known as "deferred triggering". These events were then processed between fills when there was less demand on the computing resources. An increase in computing resources for Run II allowed a higher rate to be written to storage.

Between Run I and II, there were also changes in the reconstruction used in the trigger. In Run I, the reconstruction was a simplified version of the offline reconstruction. The detector calibration and alignment parameters were calculated offline from already triggered data. In Run II, instead of deferring 20% of the L0 output, the output from HLT1 was completely deferred. This meant that the alignment and calibration could be performed between HLT1 and HLT2, enabling an offline quality reconstruction to be used upfront in HLT2.

3.3.4 Software

Software in LHCb [150] is based on an object-oriented C++ framework known as Gaudi [151]. Gaudi provides a shared infrastructure upon which "applications" to handle the specific needs of the experiment are developed. This enables experiment-wide changes to be easily implemented and prevents a duplication of effort by providing a number of features common to all applications, such as histogram creation.

To help with the processing of the data collected, LHCb member institutes pledge computing resources. During operation of the LHC, the computing demands are so great that no single institute can cope. Therefore, a distributed computing network, known as the Worldwide LHC Computing Grid (WLCG) [152], is used to pool all the resources together. The WLCG combines the resources of the four main LHC experiments (ATLAS, CMS, ALICE and LHCb) to make the world's largest computing grid.

Data recorded by LHCb is processed on the grid by a series of Gaudi applications, each building on the last, as illustrated in figure 3.21. A description of the applications responsible for simulation (section 3.3.4.1), digitisation (section 3.3.4.2), triggering (section 3.3.4.3), reconstruction (section 3.3.4.4), and analysis (section 3.3.4.5) is given.

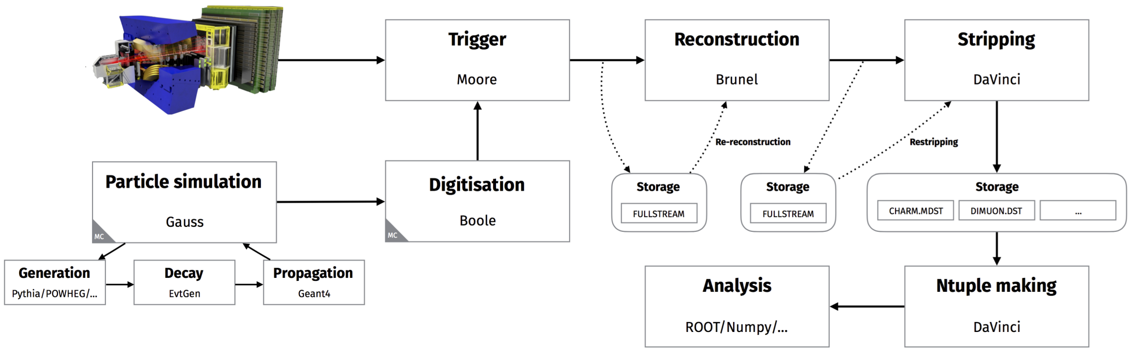


Figure 3.21: The LHCb data flow [153].

3.3.4.1 Simulation

Simulated events, known as "Monte Carlo" (MC), mimic the presence of specified decays in the LHCb detector in order to gain a better understanding of its response and performance. Simulation is generated by the GAUSS application [154]. It consists of two independent phases:

- **A generator phase**: the generation of the pp interaction, the particles produced and their decays. For B^+ meson production, the PYTHIA [155] generator is used. Due to the small value of f_c, production of B_c^+ mesons is too infrequent in PYTHIA to be computationally practical. Therefore, B_c^+ mesons are produced with a dedicated generator, BCVEGPY [156]. B mesons are decayed using EVTGEN [157] and final state radiation is generated with PHOTOS [158].

- **A simulation phase**: the simulation of the interaction of the particles with the detector material. This is carried out using the GEANT4 toolkit [159, 160].

3.3.4.2 Digitisation

The simulated hits in the detector are "digitised" by the BOOLE application. This means that the simulated hits are converted into the format provided by the detector electronics. After digitisation, simulation and real data are in the same format and are treated identically in the trigger, reconstruction and analysis steps.

3.3.4.3 Trigger

The application responsible for the software trigger (see section 3.3.3.2) is Moore. Moore is configured using a unique key, known as a Trigger Configuration Key (TCK). The TCK specifies which trigger lines to include and their cuts. This allows the same trigger requirements as used in data to be applied to simulation.

3.3.4.4 Reconstruction

The digitised data is reconstructed by the Brunel application to give physically intuitive information. Brunel performs a fit to the track hits to produce track objects with well defined trajectories and momenta. It also calculates energies from the clusters deposited in the calorimeters and PID information for each track from the outputs of the RICH, calorimeter and muon systems. The reconstructed data is saved as a “Data Summary Tape” (DST) file. In the case of simulation, Brunel can separately process additional information, known as “MC truth”, that specifies the conditions under which the particle was generated.

3.3.4.5 Analysis

The application that prepares the data for physics analysis is DaVinci. DaVinci takes the track objects from Brunel and performs fits to produce primary and decay vertices. It assigns a particle hypothesis to each track depending on the decay being considered and can calculate a number of kinematic variables, such as the B mass and lifetime. The stripping, described in section 4.4, also uses DaVinci. The output of DaVinci is either a Root ntuple, on which analysis can be performed, or another DST file, which can be further processed.

Chapter 4

Event selection

"Data, I think, is one of the most powerful mechanisms for telling stories. I take a huge pile of data and I try to get it to tell stories."
— Steven Levitt

The $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ decays are rare with predicted branching fractions of $\mathcal{O}(10^{-5} - 10^{-7})$. In addition, the small B_c^+ production cross-section, short lifetime and complex final state make observation of these decays experimentally extremely challenging. Therefore, it is essential to have an effective procedure in place to reduce the levels of background and identify the most signal-like events. Large $\mathcal{O}(10^3)$ yields are expected for the $B^+ \to D_{(s)}^+ \overline{D}^0$ decays. In this case, an effective selection can be used to minimise the statistical uncertainty on the *CP* asymmetry measurement. In this chapter, an overview of the data and simulation samples is given (section 4.1), followed by the definitions of a number of common selection variables (section 4.2). The first two stages of the selection, the trigger selection (section 4.3) and the centralised preselection (section 4.4), are then described. The tuning of the data and simulation is then detailed (section 4.5), followed by a description of the final two stages of the selection, the preselection (section 4.6) and the multivariate analysis (MVA) (section 4.7). The selection of $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ and $B^+ \to D_{(s)}^+ \overline{D}^0$ decays is identical up to the multivariate analysis.

4.1 Data and simulation samples

In this section, the data (section 4.1.1) and simulation (section 4.1.2) samples used in the analyses in this thesis are described. The decays of interest are $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ and $B^+ \to D_{(s)}^+ \overline{D}^0$, with the following subdecays: $D_s^+ \to K^+K^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$, $\overline{D}^0 \to K^+\pi^-$ or $K^+\pi^-\pi^+\pi^-$, and $D^* \to D\pi^0/\gamma$, where the soft neutral particles are not reconstructed.

4.1.1 Data

The data sample is the full Run I data set collected by the LHCb experiment. This corresponds to an integrated luminosity of $1.0\,\text{fb}^{-1}$ collected in 2011 at a centre-of-mass energy of $7\,\text{TeV}$ and $2.0\,\text{fb}^{-1}$ collected in 2012 at a centre-of-mass energy of $8\,\text{TeV}$. In 2011, approximately 43% of the data was collected in the *MagUp* polarity, whereas in 2012, the split is approximately even between *MagUp* and *MagDown*.

During the development of the selection, the data was "blinded" to prevent experimenter bias. This meant that data in the region $\pm 40\,\text{MeV}/c^2$ around the B_c^+ mass and the $B^+ \to D^+\overline{D}^0$ decay split by charge were not revealed until the analysis procedure was finalised. As a smaller *CP* asymmetry is expected in $B^+ \to D_s^+\overline{D}^0$, these decays were not blinded.

4.1.2 Simulation

The software used to produce simulated particle decays, otherwise known as Monte Carlo (MC), is described in section 3.3.4. Large samples of $B_c^+ \to D_{(s)}^{(*)+}\overline{D}^{(*)0}$ and $B^+ \to D_{(s)}^+\overline{D}^0$ simulation were produced using the 2012 beam settings with a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ and an average number of *pp* interactions per bunch crossing of $\nu = 2.5$. Decays of B_c^+ to a final state with a D^0 or $\overline{D}^0$ are expected to have identical kinematic distributions, so only the latter were generated.

Approximately 1 million events were simulated for each $B_{(c)}^+ \to D_{(s)}^{(*)+}\overline{D}^{(*)0}$ decay, with a roughly equal split between the *MagUp* and *MagDown* polarities. Events were generated with three different versions of the simulation software: `Sim08`, `Sim09a` and `Sim09b`. The kinematic distributions of decays generated with the same simulation version, *i.e.* `Sim09a` *vs.* `Sim09b`, are expected to be similar; however, there may be small differences between simulation versions, *i.e.* `Sim08` *vs.* `Sim09`. To prevent the waste of computing resources, only $B^+_{(c)}$ decays where all charged tracks fell within the LHCb acceptance ($10 \leq \theta \leq 400\,\mathrm{mrad}$) were simulated.

The B decay to two scalar partcles, $B^+_{(c)} \to D^+_{(s)} D$, was generated according to the decay kinematics, while the decay to a vector and a scalar particle, $B^+_c \to D^{*+}_{(s)} D$ or $B^+_c \to D^+_{(s)} D^*$, was generated taking into account the spin configuration. For the decay of the B^+_c to two vector particles, $B^+_c \to D^{*+}_{(s)} D^*$, the helicity amplitudes were chosen such that the decay was generated in either the longitudinal or transverse polarisation. The model used for the charged D decays reproduces the behaviour of the intermediate resonances, for example, $\phi(1020) \to K^+K^-$ and $K^*(892) \to K^-\pi^+$ for the decay of $D^+_s \to K^+K^-\pi^+$. The $\overline{D}^0$ decays were generated according to the decay kinematics. The D^* decays were generated according to the initial and final state spin.

4.2 Common selection variables

To make sense of the selection requirements, it is first helpful to define a number of common selection variables. Many different variables that exploit different signatures of B decays are utilised in the selection. These include:

- **Invariant mass ($\boldsymbol{m}$)**: the invariant mass of a system of N particles in natural units is

$$m(p_1, p_2, \dots p_N) = \sqrt{\left(\sum_i^N E_i\right)^2 - \left(\sum_i^N \mathbf{p}_i\right)^2}, \tag{4.1}$$

  where E is the energy and $\mathbf{p}$ is the momentum vector. If all the products of a decay are identified and assigned the correct mass hypothesis, then the invariant mass will peak at the mass of the decaying particle. Candidates assembled from random combinations of tracks (known as "combinatorial background") will have a smooth distribution.

- **Momentum ($\boldsymbol{p}$)**: the magnitude is given by

$$p = \sqrt{p_x^2 + p_y^2 + p_z^2}. \tag{4.2}$$

  Low momentum particles are swept out of the LHCb acceptance by the magnetic field.

- **Transverse momentum ($\boldsymbol{p_{\mathrm{T}}}$)**: the component of a particle's momentum perpendicular to the beam direction (the z axis),

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}. \tag{4.3}$$

  The relatively large B mass results in high p_T daughters.

- **Decay time ($\boldsymbol{t}$)**: the time between the production of a particle and its decay,

$$t = \frac{L}{\gamma v}, \tag{4.4}$$

  where L is the distance between the production and decay vertices, γ is the Lorentz factor and v is the velocity. Due to detector resolution effects, physical values of decay time can be measured to be slightly negative. B mesons have long lifetimes of $\mathcal{O}(1\,\mathrm{ps})$. As combinatorial background is made up of random combinations of tracks produced at the primary vertex (PV), the decay time is expected to peak at 0.

- **Decay time significance ($\boldsymbol{t/\Delta t}$)**: the decay time divided by its uncertainty. Placing a minimum requirement on the decay time significance removes particles with very large decay time uncertainties.

- **Distance of closest approach (DOCA)**: the shortest distance between a pair of tracks. Placing a maximum requirement on the DOCA of two tracks removes backgrounds that don't originate from the same decay vertex.

- **Impact parameter (IP)**: the distance of closest approach between the PV and the extrapolated particle trajectory. The trajectory is found by summing the momentum vectors of the decay products and extrapolating back. Particles, such as B mesons, that originate from the PV are expected to have a small IP, however, tracks that originate from displaced vertices, such as those from the secondary vertex (SV), are expected to have large values.

- **Impact parameter $\boldsymbol{\chi^2}$ ($\boldsymbol{\chi^2_{\mathrm{IP}}}$)**: the impact parameter χ^2 is the difference between the fit quality of the primary vertex (PV) reconstructed with and without the particle under consideration.

- **Vertex $\boldsymbol{\chi^2}$ ($\boldsymbol{\chi^2_{vtx}}$)**: the χ^2 of the vertex fit. Placing a maximum requirement on this removes poorly reconstructed vertices.

- **Vertex separation χ^2 (χ^2_{VS})**: the difference between the fit quality of an event where the particle is or is not constrained to have zero lifetime. Placing a minimum requirement on the vertex separation χ^2 removes backgrounds without any separation between the decay vertex and the PV.

- **Track χ^2/number of degrees of freedom (χ^2/ndf)**: the χ^2 per degree of freedom of the track fit. Placing a maximum requirement on this removes poor quality tracks.

- **DLL$_{K\pi}$**: the difference in log likelihood between the kaon and pion mass hypotheses (see equation 3.5). The larger the value of DLL$_{K\pi}$, the more likely the particle is to be a kaon.

- **cos(θ)**: the direction angle θ is the angle between the line drawn from the PV to the decay vertex of the particle and its momentum vector. Placing a minimum requirement on $\cos(\theta)$ removes backgrounds that don't originate from the PV.

- **Ghost probability**: the response of a neural network trained to discriminate between real and "ghost" tracks. Ghost tracks are formed when tracks in the VELO are matched with tracks in the T-stations that originate from different particles.

4.3 Trigger

The trigger, described in section 3.3.3, identifies potentially interesting events to save for offline analysis. An event needs to fire a trigger line at each stage of the trigger (L0, HLT1 and HLT2) to be saved, otherwise it is discarded. If the trigger fires due to particles originating from the signal B candidate, the event is described as triggered on signal (TOS). The trigger can fire due to the individual tracks that make up the candidate, or some combination of them, depending on the line being considered. If the trigger fires due to another part of the event, then it is described as triggered independently of signal (TIS). The two categories are not exclusive. For example, a track from the B candidate as well as from some other part of the event could simultaneously fire the trigger, making the event both TIS and TOS. It is also possible for an event to be neither TIS nor TOS. In this case neither the B candidate nor the rest of the event is sufficient alone to fire the trigger, but some combination of them is. This is known as triggered on both (TOB). Events in the TOB category only make up a small fraction of all events and are not included in the analyses in this thesis.

The B candidates are required to satisfy well-defined trigger requirements. Although the largest yields would come from analysing all triggered events, the output would be very difficult to interpret since different lines impose different requirements. The lines were chosen based on the particles in the final state and the decay topology.

L0

At L0, the B candidates are required to be TOS on the hadron trigger (`L0Hadron`) or TIS on the global trigger (`L0Global`). The hadron trigger requires a high E_T deposit in the HCAL. The global trigger takes the logical OR of all the triggers (single muon, dimuon, hadron, electron and photon). Events where the other B meson produced from the hadronisation of the $b\bar{b}$ pair caused the trigger to fire are TIS on `L0Global`. Including these helps to reduce the effect of the inefficiency of `L0Hadron`. The fractional yields selected by `L0Hadron` and `L0Global` are shown in table 4.1. There is a large overlap between the two lines, as is evident from the difference between the inclusive and exclusive yields.

Channel	$\overline{D}^0 \to K^+\pi^-$		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	
	incl.	excl.	incl.	excl.
`L0Hadron TOS`	0.75	0.40	0.67	0.43
`L0Global TIS`	0.50	0.25	0.57	0.33

Table 4.1: Fractional yield, after all selection requirements, of $B_c^+ \to D_s^+ \overline{D}^0$ simulation by L0 trigger line. Inclusive (incl.) yields refer to when the trigger line fired independent of the others (*e.g.* `TOS`), whereas exclusive (excl.) refers to when only that line fired (*e.g.* `TOS && !TIS`).

HLT1

At HLT1, the candidate is required to be TOS on `Hlt1TrackAllL0`. `Hlt1TrackAllL0` requires a single high p_T track with a good track fit quality that is well displaced from all pp interaction points. Since only one HLT1 line is considered, all the events in the final data set fired this line.

HLT2

At HLT2, the candidate is required to be TOS on `Hlt2Topo2BodyBBDT`, `Hlt2Topo3BodyBBDT`, or `Hlt2Topo4BodyBBDT`, or for the D_s^+ modes, `Hlt2IncPhi`. The topological triggers, `Hlt2Topo{2,3,4}BodyBBDT`, combine several different variables (scalar sum of the p_{T} of the tracks, the invariant mass, DOCA, *etc.*) into a boosted decision tree [148]. The `Hlt2IncPhi` line selects ϕ mesons built from two oppositely charged kaons by requiring several kinematic, particle identification and vertex constraints, including that the mass is within $20\,\mathrm{MeV}/c^2$ of the known ϕ mass. The $\phi \to K^+K^-$ decay is one of the main resonant contributions to the $D_s^+ \to K^+K^-\pi^+$ decay. The fractional yields selected by the HLT2 lines are shown in table 4.2. Again, there is a large overlap between the lines. For decays with $\overline{D}^0 \to K^+\pi^-$, the three-body topological trigger is the most important and for decays with $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$, the four-body topological trigger is the most important.

Channel	$\overline{D}^0 \to K^+\pi^-$		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	
	incl.	excl.	incl.	excl.
`Hlt2Topo2BodyBBDT`	0.69	0.02	0.58	0.01
`Hlt2Topo3BodyBBDT`	0.91	0.05	0.88	0.02
`Hlt2Topo4BodyBBDT`	0.81	0.02	0.93	0.05
`Hlt2IncPhi`	0.33	0.03	0.32	0.02

Table 4.2: Fractional yield, after all selection requirements, of $B_c^+ \to D_s^+ \overline{D}^0$ simulation by HLT2 trigger line. Inclusive (incl.) yields refer to when the trigger line fired independent of the others, whereas exclusive (excl.) refers to when only that line fired.

4.4 Stripping

Events that pass the trigger are saved to disk and a full event reconstruction is performed offline. In principle, the reconstructed files can be used for analysis. However, since only a very small fraction of the events are interesting for a given analysis and the limited number of copies would be accessed by multiple analysts at once, this would be very computationally inefficient. The solution is to apply a centralised preselection, known as the “stripping”.

Individual analysts write “stripping lines” that define a set of loose requirements to select a particular decay channel. Stripping lines are then grouped together into “streams”. The output of each stream is written to a separate file. If two stripping lines in the same stream select the same event, it is only written once, so lines with similar selections, such as all $B \to DX$ lines, are run together. The amount of disk space required by the output of the stripping would be minimised by only having one stream, however, increasing the number of streams makes accessing the output of a particular stripping line much more efficient for analysts.

The data collected in each year is stripped in its own “campaign”. The campaigns for 2011 and 2012 data are `Stripping21r1` and `Stripping21`, respectively. These both contain 13 streams. The output of a stream is either a DST or MicroDST (MDST) file, where MDST files store a subset of the information available in a DST file. The stream of interest for the analyses in this thesis is the `Bhadron` stream, which contains $B \to DX$ lines and is written in MDST format. There are approximately 850 lines in the `Bhadron` stream in `Stripping21(r1)`.

The `StrippingB2D0DBeauty2CharmLine` and `StrippingB2D0DD02K3PiBeauty2CharmLine` lines in the `Bhadron` stream are used to select $B^+_{(c)} \to D^{(*)+}_{(s)} D^{(*)}$ decays with $\overline{D}^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$, respectively. Table 4.3 lists the requirements these lines make on the B candidates. “B2CBBDT” refers to a boosted decision tree trained to select $B \to DX$ decays [161].

Figure 4.1 shows the invariant mass of the B candidates that pass the stripping requirements. There is a prominent $B^+ \to D^+_s \overline{D}^0$ peak at $\sim 5280\,\text{MeV}/c^2$, however, no $B^+ \to D^+ \overline{D}^0$ signal can be seen yet. Larger yields are expected in $B^+ \to D^+_s \overline{D}^0$ compared to $B^+ \to D^+ \overline{D}^0$ due to the factor ~ 25 larger branching fraction [3]. There is a small peak at the nominal B^+ mass for $B^+ \to D^+_s \overline{D}^0$ candidates using data from the D^+_s sidebands. This indicates the presence of the single charm background $B^+ \to K^+K^-\pi^+\overline{D}^0$. Below the B^+ mass, there are two satellite peaks characteristic of partially reconstructed $B^+ \to D^{*+}_s \overline{D}^0$, $B^+ \to D^+_s \overline{D}^{*0}$ and $B^+ \to D^{*+}_s \overline{D}^{*0}$ decays. The B^+_c mass is $\sim 6275\,\text{MeV}/c^2$. Clearly, the background levels need to be further reduced to search for the rare B^+_c signals.

Variable	Requirement
B2CBBDT	> 0.05
$m(B)$	$> 4750\,\mathrm{MeV}/c^2$, $< 7000\,\mathrm{MeV}/c^2$
$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}\ (B)$	< 10
$t(B)$	$> 0.2\,\mathrm{ps}$
$\chi^2_{\mathrm{IP}}(B)$	< 25
$\cos(\theta_B)$	> 0.999
$m(D^+_{(s)})$	$> 1769.62\,\mathrm{MeV}/c^2$, $< 2068.49\,\mathrm{MeV}/c^2$
$m(\overline{D}^0)$	$> 1764.84\,\mathrm{MeV}/c^2$, $< 1964.84\,\mathrm{MeV}/c^2$
$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}\ (D^+_{(s)},\overline{D}^0)$	< 10
$\chi^2_{\mathrm{VS}}\ (D^+_{(s)},\overline{D}^0)$	> 36
$\cos(\theta_{D^+_{(s)},\overline{D}^0})$	> 0
$\sum p_{\mathrm{T}}$ of $D^+_{(s)} \to hhh/\overline{D}^0 \to hh(hh)$	$> 1800\,\mathrm{MeV}/c$
$\min(\chi^2/\mathrm{ndf})$ of $D^+_{(s)} \to hhh$, $\overline{D}^0 \to hh(hh)$ tracks	< 2.5
$\max(p_{\mathrm{T}})$ of $D^+_{(s)} \to hhh$, $\overline{D}^0 \to hh(hh)$ tracks	$> 500\,\mathrm{MeV}/c$
$\max(p)$ of $D^+_{(s)} \to hhh$, $\overline{D}^0 \to hh(hh)$ tracks	$> 5000\,\mathrm{MeV}/c$
DOCA of $D^+_{(s)} \to hhh$, $\overline{D}^0 \to hh(hh)$ tracks	$< 0.5\,\mathrm{mm}$
$\mathrm{DLL}_{K\pi}$ for π (K)	$< 20 (> -10)$
χ^2/ndf of tracks	< 3
p_{T} of tracks	$> 100\,\mathrm{MeV}/c$
p of tracks	$> 1000\,\mathrm{MeV}/c$
χ^2_{IP} of tracks	> 4
ghost probability of tracks	< 0.4
p_{T} of at least one track	$> 1700\,\mathrm{MeV}/c$
p of at least one track	$> 10000\,\mathrm{MeV}/c$
χ^2_{IP} of at least one track	> 16
IP of at least one track	$> 0.1\,\mathrm{mm}$
$\sum p_{\mathrm{T}}$ of all tracks	$> 5000\,\mathrm{MeV}/c$
number of long tracks	< 500

Table 4.3: Requirements placed on $B^+_{(c)} \to D^+_{(s)} D$ candidates, where h refers to either a kaon or pion, in `Stripping21(r1)`.

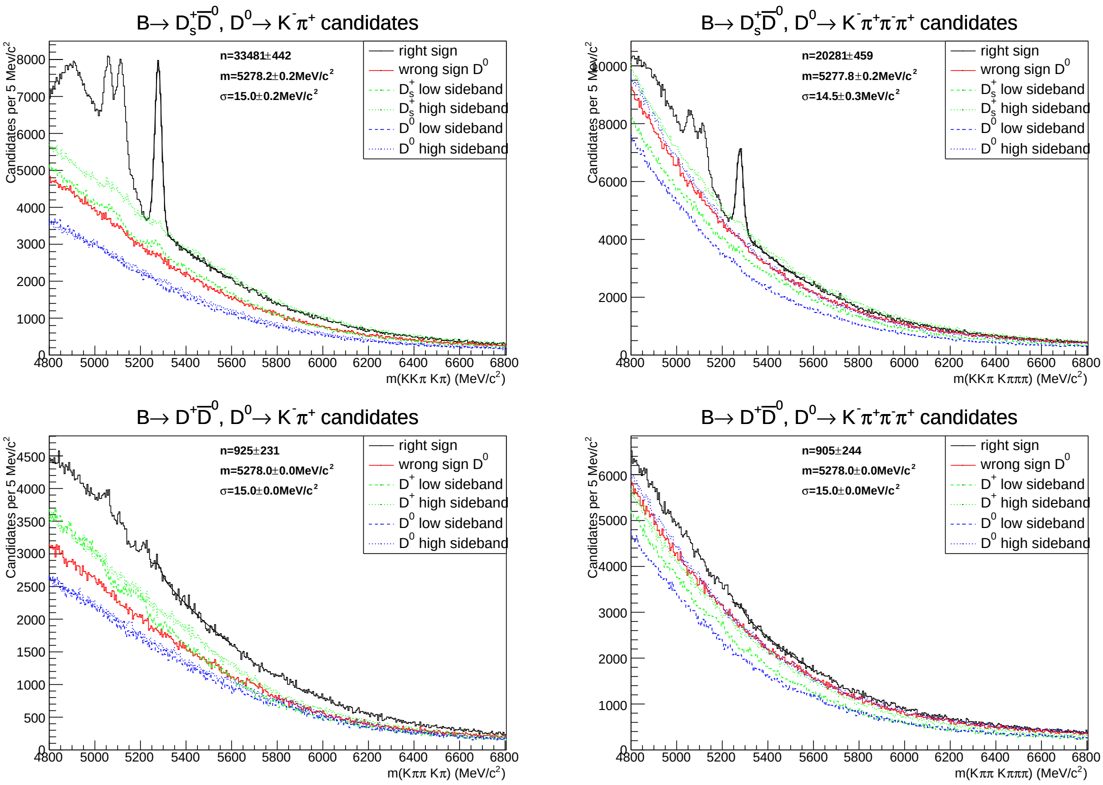


Figure 4.1: Distributions of the invariant mass of $B^+_{(c)} \to D^+_{(s)}\overline{D}^0$ ("right-sign") and $B^+_{(c)} \to D^+_{(s)}D^0$ ("wrong-sign") candidates after the stripping. The signal is defined as $|m(D) - m_D| < 25\,\text{MeV}/c^2$ and the sidebands are $25 < |m(D) - m_D| < 50\,\text{MeV}/c^2$. A fit to the data with the sum of an exponential function and a Gaussian function in the range $5240 - 5600\,\text{MeV}/c^2$ is overlaid. For the D_s^+ modes, the shape parameters in the fit are floating, while for the D^+ modes the displayed values have been fixed.

4.5 Calibrations, corrections, constraints and comparisons

The data is only available to analysts after the stripping. Before applying additional selection requirements, a number of refinements are made to the data and simulation. This section details the calibration of the tracker response in data (section 4.5.1) and the data-driven corrections applied to the simulation to account for known discrepancies (section 4.5.2). The kinematic fit incorporating information from the decay hypothesis as constraints to improve the B mass resolution is described (section 4.5.3), followed by verification that the data and simulation are in agreement (section 4.5.4).

4.5.1 Calibrations to data

The momentum measurement may be biased due to imperfect knowledge of the magnetic field and the tracker alignment. To account for this, a scale factor is applied to the data to calibrate the momentum of the tracks. The scale is derived in two stages: first $J/\psi \to \mu^+\mu^-$ decays are used to determine changes in the momentum scale between different data-taking periods and secondly, the absolute scale is determined from $B^+ \to J/\psi\, K^+$ decays, given the known B^+ mass, as a function of the K^+ track kinematics [162]. Applying the momentum scale calibration to the data ensures that particle masses are consistent with the world-average values.

4.5.2 Corrections to simulation

Imperfect modelling of the underlying physics and/or the detector can lead to disagreements between data and simulation. In order to match the simulation with the data as closely as possible, a number of corrections are applied:

- **Momentum smearing and scaling**: the simulation is known to overestimate the performance of the tracking system, resulting in mass resolutions that are too narrow compared to data. The momentum of the tracks are smeared to bring them into agreement with the data. A momentum scale of 0.9994 is also applied to improve the agreement of data and simulation in the invariant mass of the D mesons (see figures A.2 and A.3).

- **$\boldsymbol{B_c^+}$ lifetime**: the `Sim08` simulation was generated with a B_c^+ lifetime of $\tau_{\text{gen}}(B_c^+) = 453.9\,\text{fs}$, which is significantly less than the current world-average $\tau_{\text{WA}}(B_c^+) = (507 \pm 9)\,\text{fs}$ [3]. As the relative efficiency of B_c^+ to B^+ is sensitive to the lifetime, a weight is applied to correct the decay time distribution,

$$w = \frac{\tau_{\text{gen}}}{\tau_{\text{WA}}} e^{t(1/\tau_{\text{gen}} - 1/\tau_{\text{WA}})}, \tag{4.5}$$

  where t is the true decay time. Weighted in this way, the overall normalisation is unaffected. No such weight is necessary for the `Sim09` simulation since the lifetime used in the generation is up-to-date.

- **Particle identification**: the performance of the RICH subdetectors is also overestimated in simulation, resulting in better separating power between kaons and pions than is observed in data. A transformation [163] is applied to the track $\mathrm{DLL}_{K\pi}$ distributions to bring them into agreement with the data (see figure A.5).

4.5.3 Constrained *B* mass

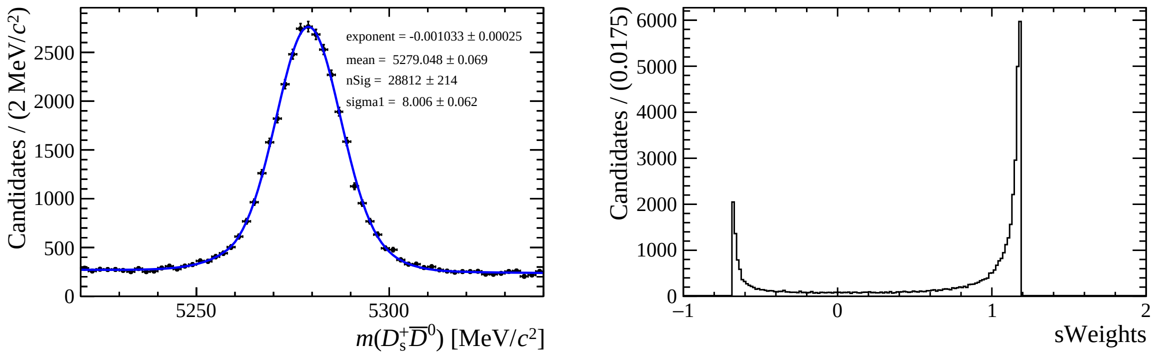
The invariant mass of the B candidate is calculated from the four-momenta of the final state tracks. However, for a real $B^+_{(c)} \to D^{(*)+}_{(s)} D^{(*)}$ decay, the invariant masses of the charm mesons should correspond to their known values and the three-momenta of the tracks should point back to the PV. This additional information can be incorporated as constraints in a kinematic fit [130] to improve the B mass resolution. In the kinematic fit, the charm masses are fixed to their known values, all particles from the $D^+_{(s)}$, $\overline{D}^0$ or $B^+_{(c)}$ decays are constrained to originate from their decay vertex and the $B^+_{(c)}$ is constrained to originate from the PV. Applying these constraints results in a $\sim 40\%$ improvement in the B mass resolution. The constrained B mass is depicted with $m(D^+_{(s)}D)$ and is used throughout the rest of this thesis unless otherwise stated.

4.5.4 Data/simulation comparison

The simulation samples play a number of important roles in the analyses. They are used to train and optimise the multivariate classifiers, determine the relative efficiency between B_c^+ and B^+ decays and to constrain the signal shapes in the mass fit. Therefore it is essential to verify that the simulation reproduces the behaviour of the data. Although the expected B_c^+ signals are too small to be able to compare with simulation, a comprehensive comparison can be done with the large B^+ samples.

The *sPlot* technique [164] is used to unfold the B^+ signal from the background. This technique makes use of *discriminating variables* (variables in which the distributions of the signal and background are known) to infer the distributions of the signal and background components in a set of *control variables* (the variables of interest). This means that the distribution of the signal can be determined in the control variables without any prior knowledge.

The discriminating variable used is the constrained B mass (with the preselection applied). The fit probability density function (PDF) is the sum of a B^+ signal model and a background model. The B^+ signal model is described in section 5.2.1 and the background model is an exponential function with a floating exponent. The fit is performed to the



(a) Constrained B mass with the fit PDF overlaid. **(b)** Distribution of sWeights.

Figure 4.2: Fit to the discriminating variable and the resulting sWeight distribution for $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ candidates.

constrained B mass $\pm 60\,\text{MeV}/c^2$ around the B^+ peak, as shown in figure 4.2a for the high-statistics mode, $B^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays. Figure 4.2b shows the sWeights derived from this fit. Signal-like events peak just above 1, whereas background-like events peak at negative values.

Appendix A shows the comparison of the sWeighted B^+ data with the simulation. All variables related to the analysis are shown for the high-statistics mode, $B^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays, as well as those specific to the D^+ for $B^+ \to D^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays and the $\overline{D}^0$ for $B^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ decays. In general, there is good agreement, both between data and simulation and between 2011 and 2012 data. The 2012 simulation therefore provides a good description of the full Run I data set. As shown in figure A.6, the only remaining discrepancy is due to the simplistic phase space model used for the $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ decay, which is addressed in section 5.6.2. No requirements are placed on the $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ Dalitz variables in the selection.

4.6 Preselection

Before the MVA, the preselection is applied to further reduce the background levels with little loss of signal. This ensures that the background samples used for training are those that most successfully mimic the signal. The preselection requirements are:

- $p_{\mathrm{T}}(B) > 4\,\text{GeV}/c$,
- $t(D_s^+, \overline{D}^0)/\Delta t > -3$, $t(D^+)/\Delta t > 3$,

- $\chi^2_{\text{IP}}(B) < 15$,
- $\text{DLL}_{K\pi}(\pi) < 10$, $\text{DLL}_{K\pi}(K) > -5$,
- $|m(D) - m_D| < 25\,\text{MeV}/c^2$.

The decay time significance requirement is tighter for D^+ to reject the single charm background $B^+ \to \pi^-\pi^+\pi^+\overline{D}^0$, where the π^- is misidentified as a K^-. The branching fraction of $B^+ \to \pi^-\pi^+\pi^+\overline{D}^0$ decays is two orders of magnitude larger than decays of $B^+ \to D^+\overline{D}^0$ with $D^+ \to K^-\pi^+\pi^+$ [3]. The misidentification of a pion as a kaon shifts the B mass upwards, causing this background to pollute the B upper mass sideband $5350 < m(D^+_{(s)}D) < 6200\,\text{MeV}/c^2$. Using data from the upper mass sideband as a background sample to train a MVA would therefore preferentially remove $B^+ \to \pi^-\pi^+\pi^+\overline{D}^0$ decays, as opposed to the combinatorial background that appears at the B_c^+ mass. Introducing the tighter lifetime significance requirement to remove this background results in a relatively small inefficiency of $\sim 15\%$ on $B^+ \to D^+\overline{D}^0$ decays due to the long D^+ lifetime.

The charged D meson decays are $D_s^+ \to K^+K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. There is cross-feed between the two modes if the K^- from the D_s^+ decay is misidentified as a π^- or if either of the pions from the D^+ decay are misidentified as a kaon. This results in two shifted mass peaks approximately $60\,\text{MeV}/c^2$ above and below the B^+ mass peak, as shown in figure 4.3. Instead of applying tighter $\text{DLL}_{K\pi}$ requirements on the relevant tracks, more sophisticated vetoes that exploit the resonance structure of the decays are employed [165]:

- D_s^+ candidates are rejected if, under the $K^-\pi^+\pi^+$ hypothesis, the invariant mass corresponds to the D^+ mass ($|m(K^-\pi^+\pi^+) - m_{D^+}| < 25\,\text{MeV}/c^2$), the $\phi(1020) \to K^+K^-$ resonance is not present ($|m(K^+K^-) - m_\phi| > 10\,\text{MeV}/c^2$) and the K^+ is not likely to be a kaon ($\text{DLL}_{K\pi}(K^+) < 10$).
- D^+ candidates are rejected if, under the $K^+K^-\pi^+$ hypothesis, the invariant mass corresponds to the D_s^+ mass ($|m(K^+K^-\pi^+) - m_{D_s^+}| < 25\,\text{MeV}/c^2$) and either the $\phi(1020) \to K^+K^-$ resonance is present ($|m(K^+K^-) - m_\phi| < 10\,\text{MeV}/c^2$) or the π^+ is likely to be a kaon ($\text{DLL}_{K\pi}(\pi^+) > -10$). The cross-feed vetoes for D^+ are applied twice for each π^+.

The vetoes reduce the cross-feed down to negligible levels with little loss of signal.

The efficiency of the preselection on signal and background is shown in table 4.4. The background is strongly suppressed while most of the signal survives. The efficiencies for

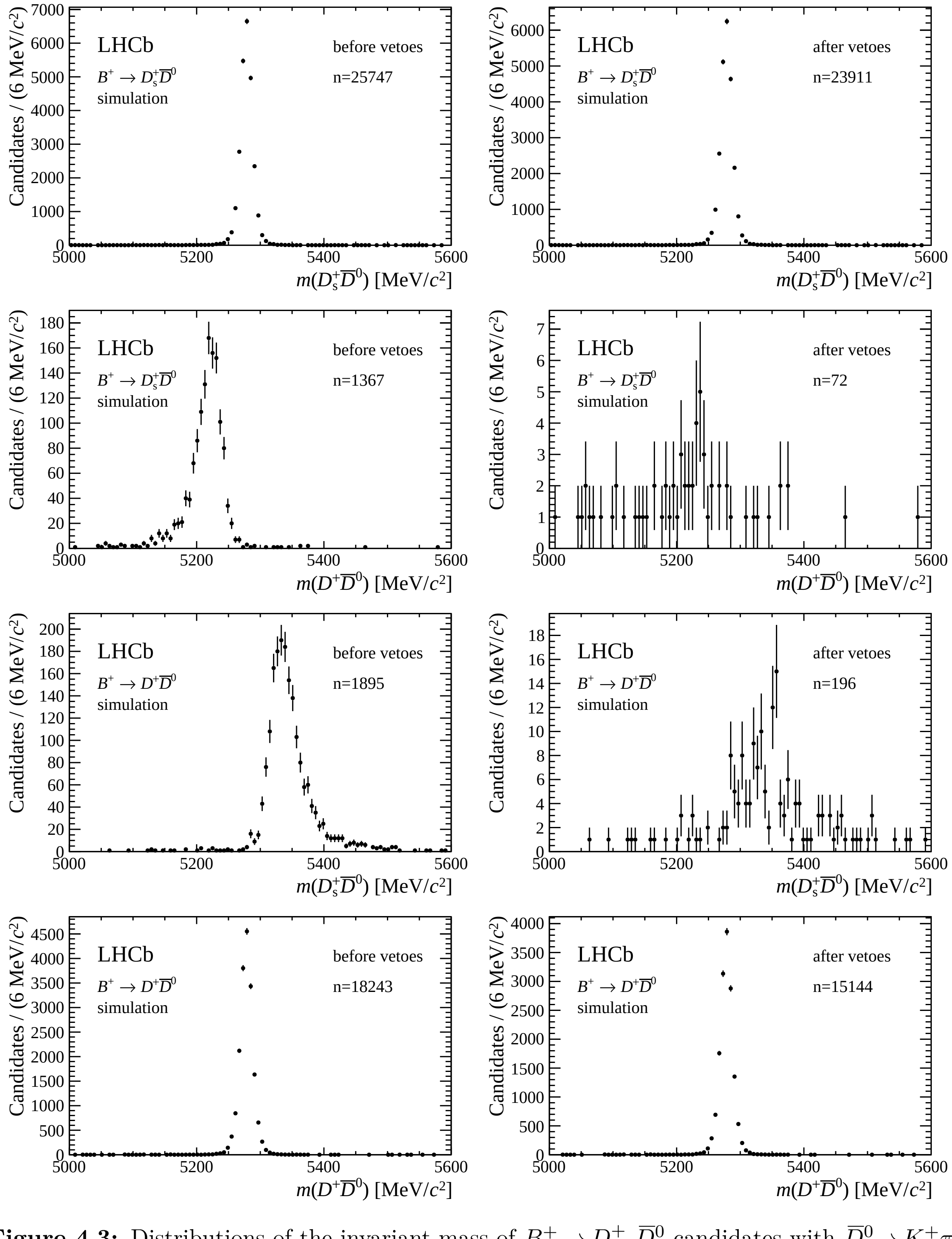


Figure 4.3: Distributions of the invariant mass of $B^+ \to D^+_{(s)}\overline{D}^0$ candidates with $\overline{D}^0 \to K^+\pi^-$ decays from $B^+ \to D^+_{(s)}\overline{D}^0$ simulation before and after the cross-feed vetoes.

Channel	$\varepsilon_{B^+ \text{ simulation}}$	$\varepsilon_{B_c^+ \text{ simulation}}$	$\varepsilon_{\text{data background}}$
$D_s^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	80.0 ± 0.3	77.6 ± 0.2	18.8 ± 0.1
$D_s^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	71.3 ± 0.7	69.9 ± 0.4	15.0 ± 0.1
$D^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	61.7 ± 0.4	60.6 ± 0.3	7.8 ± 0.1
$D^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	56.9 ± 0.7	53.7 ± 0.5	7.4 ± 0.1

Table 4.4: The efficiency of the preselection, in %, evaluated for signal using $B^+_{(c)} \to D^+_{(s)}\overline{D}^0$ simulation. The number of B candidates is evaluated by fitting the data with a Gaussian function $\pm 60\,\text{MeV}/c^2$ around the nominal B mass. The efficiency on background is evaluated from the number of candidates in the $5350 < m(D^+_{(s)}D) < 6200\,\text{MeV}/c^2$ range.

the D^+ mode are lower than for the D_s^+ mode due to the tighter decay time significance requirement and cross-feed vetoes. The $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ modes have lower efficiencies than $\overline{D}^0 \to K^+\pi^-$ due to the requirements on the additional tracks.

Figure 4.4 shows the constrained mass of the B candidates that pass the preselection requirements (*cf.* the invariant B mass of the stripping output in figure 4.1). Due to the large reduction in combinatorial background, the D^+ modes are now clearly visible. In addition, the mass resolution of the D_s^+ modes has improved due to the use of the constrained B mass in place of the invariant B mass. The peak from the single charm $B^+ \to K^+K^-\pi^+\overline{D}^0$ background is no longer visible due to the constraint on the D_s^+ mass.

A small peak is visible for "wrong sign" $B^+ \to D_s^+ D^0$ candidates; however, due to the flavour structure, no $B^+ \to D^+_{(s)} D^0$ signal is expected. Instead, the peak is made up of two sources of background. First, the decay of $B^+ \to D^+_{(s)}\overline{D}^0$ followed by the doubly Cabibbo suppressed $\overline{D}^0 \to K^-\pi^+(\pi^-\pi^+)$ decay, which occurs $\sim 0.4\%$ as often as the Cabibbo favoured decay $\overline{D}^0 \to K^+\pi^-(\pi^+\pi^-)$ [3]. Second, the decay of $B^+ \to D^+_{(s)}\overline{D}^0$ followed by the favoured $\overline{D}^0 \to K^+\pi^-(\pi^+\pi^-)$ decay, where a pair of oppositely charged kaons and pions are misidentified as each other, which is estimated from simulation to be less than 1%.

4.7 Multivariate analysis

The MVA is the last stage of the event selection. MVA algorithms take as input a signal sample, background sample and set of discriminating variables. For n input

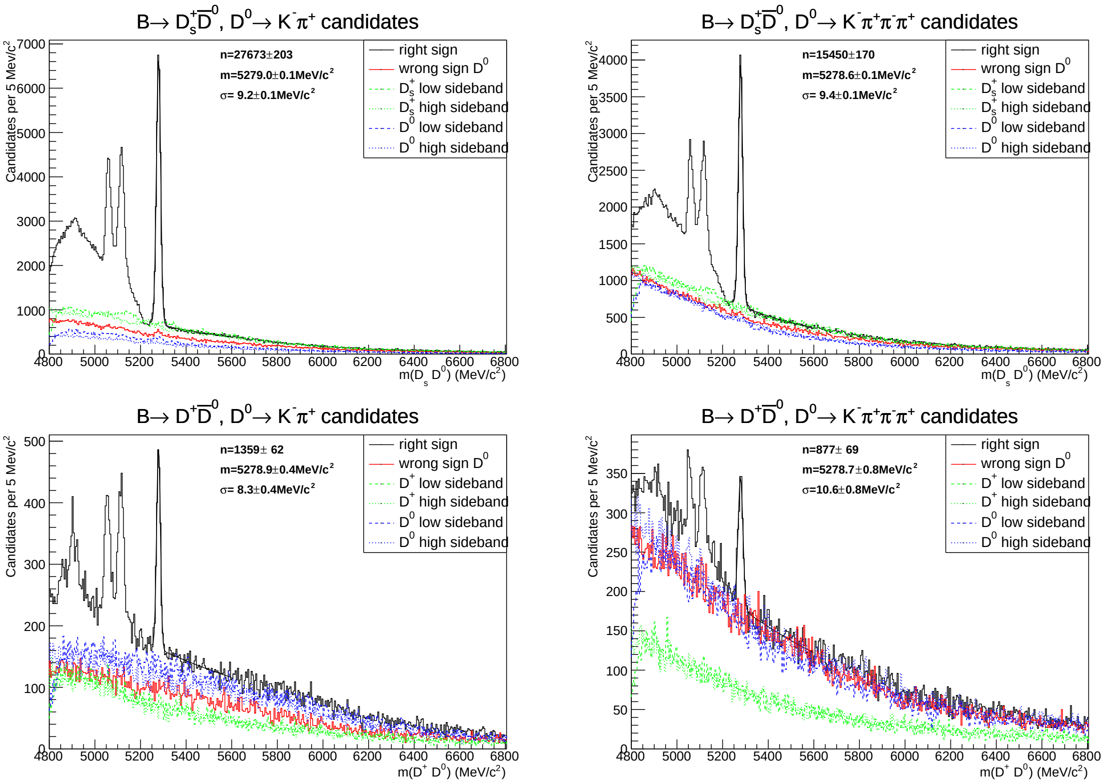


Figure 4.4: Distributions of the constrained mass of $B^+_{(c)} \to D^+_{(s)}\overline{D}^0$ ("right-sign") and $B^+_{(c)} \to D^+_{(s)} D^0$ ("wrong-sign") candidates after the preselection. The signal is defined as $|m(D) - m_D| < 25\,\text{MeV}/c^2$ and the sidebands are $25 < |m(D) - m_D| < 50\,\text{MeV}/c^2$. A fit to the data with the sum of an exponential function and a Gaussian function in the range $5240 - 5600\,\text{MeV}/c^2$ is overlaid.

variables, the MVA algorithm classifies the n dimensional variable space as signal or background, depending on the properties of the signal and background samples. As MVA algorithms can exploit correlations between input variables, they perform better than the traditional method of applying rectangular requirements. The output of the MVA is a single classifier that indicates how signal-like an event is. In this section, the properties of the background are investigated (section 4.7.1), the training of the MVA algorithm is described (section 4.7.2), and the procedure to optimise the requirement placed on the MVA response is detailed (section 4.7.3). Separate MVAs are trained for the $D^+_s D$ and $D^+ D$ final states with $\overline{D}^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ decays and for the B^+_c search and the CP asymmetry measurement, a total of 8.

4.7.1 Investigation of background properties

To develop an effective MVA algorithm to reduce the background, first the properties of the background need to be understood. The invariant masses of the charm mesons are a useful tool to gain insight into the composition of the background. Combinatorial background can broadly be split into three categories:

- **Fake charm**: random combinations of tracks. Fake charm background will appear relatively flat in both $m(D)$.
- **Single charm**: the combination of one fake charm meson with one real charm meson. The fake charm meson will be flat in $m(D)$, whereas the real charm meson will peak at the D mass.
- **Double charm**: the combination of two real charm mesons. This type of background will peak in both $m(D)$ and is the most difficult to remove since it is the most similar to the signal.

By definition, each of these background types is not a real B decay, therefore their proportions can be determined using data from the B upper mass sideband $5350 < m(D^+_{(s)}D) < 6200\,\mathrm{MeV}/c^2$. The upper mass sideband is above the B^+ kinematic threshold, so is only expected to contain combinatorial background. Figure 4.5 shows the invariant mass of the $D^+_{(s)}$ meson against the invariant mass of the D meson. The D^+_s modes show a clear vertical band indicating real D (single charm) background, whereas the D^+ modes show a clear horizontal band, indicating real D^+ (single charm) background.

To quantify the charm content of the signal window, the 2-dimensional ($m(D^+_{(s)})$, $m(D)$) space was split into nine regions of equal area, labelled $a, b \dots i$ from left-to-right and top-to-bottom. The corner regions only contain combinatorial background, whereas the cross-shaped pattern in the middle contains additional real charm components. The yield of each type of background can be estimated with

$$N_{hh} = (N_a + N_c + N_g + N_i)/4, \tag{4.6}$$

$$N_{D^+_{(s)}h} = (N_d + N_f)/2 - N_{hh}, \tag{4.7}$$

$$N_{hD} = (N_b + N_h)/2 - N_{hh}, \tag{4.8}$$

$$N_{D^+_{(s)}D} = N_e - N_{D^+_{(s)}h} - N_{hD} - N_{hh}, \tag{4.9}$$

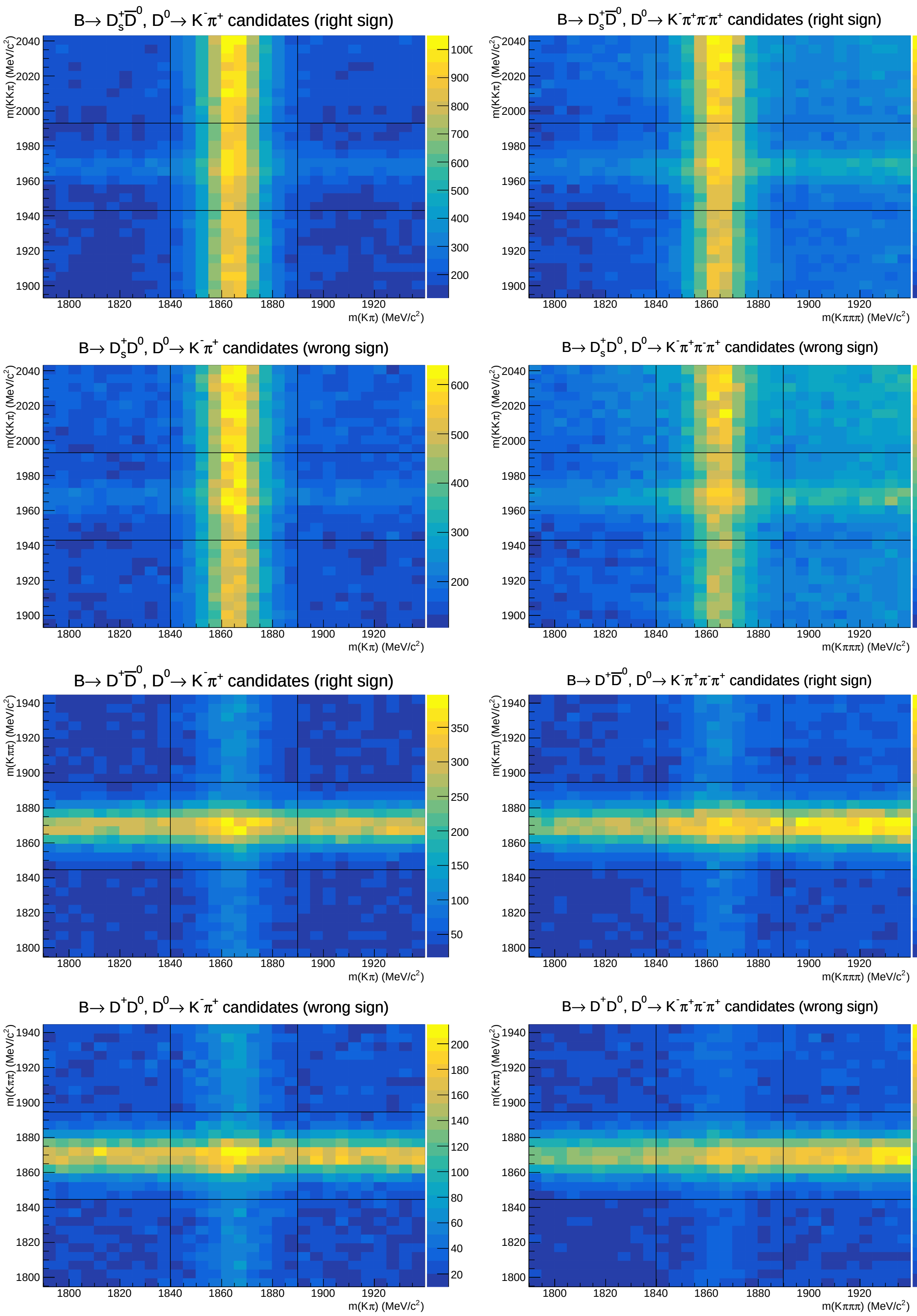


Figure 4.5: Distribution of $m(D_{(s)}^+)$ against $m(D)$ for $B^+ \to D_{(s)}^+ D$ candidates in the B upper mass sideband ($5350 < m(D_{(s)}^+ D) < 6200\,\mathrm{MeV}/c^2$) with the preselection applied.

Channel	f_{hh}	f_{hD}	$f_{D^+_{(s)}h}$	$f_{D^+_{(s)}D}$
$B_c^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	33.3 ± 0.1	66.8 ± 0.3	4.6 ± 0.2	-4.6 ± 0.6
$B_c^+ \to D_s^+ D^0$, $D^0 \to K^-\pi^+$	44.1 ± 0.2	54.5 ± 0.4	3.0 ± 0.3	-1.6 ± 0.7
$B_c^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	56.8 ± 0.2	42.7 ± 0.4	3.4 ± 0.3	-2.9 ± 0.6
$B_c^+ \to D_s^+ D^0$, $D^0 \to K^-\pi^+\pi^-\pi^+$	67.6 ± 0.2	29.6 ± 0.4	2.9 ± 0.4	-0.2 ± 0.7
$B_c^+ \to D^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	19.0 ± 0.2	19.7 ± 0.4	62.4 ± 0.5	-1.0 ± 1.0
$B_c^+ \to D^+ D^0$, $D^0 \to K^-\pi^+$	23.8 ± 0.2	18.6 ± 0.5	60.0 ± 0.7	-2.5 ± 1.3
$B_c^+ \to D^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	26.2 ± 0.2	11.8 ± 0.4	61.7 ± 0.5	0.2 ± 1.0
$B_c^+ \to D^+ D^0$, $D^0 \to K^-\pi^+\pi^-\pi^+$	27.6 ± 0.2	7.1 ± 0.4	66.0 ± 0.6	-0.8 ± 1.0

Table 4.5: Composition of the background in the region $5350 < m(D^+_{(s)}D) < 6200\,\text{MeV}/c^2$, expressed in %. The uncertainties are statistical only. The systematic uncertainty is $\sim 5\%$.

where the combinatorial background is assumed to scale linearly with mass. Table 4.5 lists the fractional yields calculated from these equations. The systematic uncertainty is estimated from the difference between the left and right regions as $\sim 5\%$. The double charm component is consistent with zero across all decay modes. There are significant contributions from single charm and pure combinatorial backgrounds.

4.7.2 Training

In this section, the signal and backgrounds samples (section 4.7.2.1), input variables (section 4.7.2.2), methods (section 4.7.2.3) and overtraining (section 4.7.2.4) are described. The MVA training is implemented using the `TMVA` package [166].

4.7.2.1 Signal and background samples

MVAs have the best performance when trained on large signal and background samples. This reduces the sensitivity of the classifier to statistical fluctuations and increases the amount of information available. The samples are:

- **Signal**: $B_c^+ \to D^+_{(s)}\overline{D}^0$ simulation for the B_c^+ search, and $B^+ \to D^+_{(s)}\overline{D}^0$ simulation for the *CP* asymmetry measurement. A requirement that the $B^+_{(c)}$ mass is within $\pm 40\,\text{MeV}/c^2$ of the nominal mass is applied to remove poorly reconstructed events. The size of the signal samples is $\mathcal{O}(10,000)$ or $\mathcal{O}(3,000)$ events for $B_c^+ \to D^+_{(s)}\overline{D}^0$ with

$\overline{D}^0 \to K^+\pi^-$ or $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$, respectively. The B^+ samples are approximately twice as large.

- **Background**: data from the B upper mass sideband $5350 < m(D^+_{(s)}D) < 6200\,\mathrm{MeV}/c^2$. To increase the size of the sample, both $B^+ \to D^+_{(s)}\overline{D}^0$ and $B^+ \to D^+_{(s)}D^0$ decays are used and the charm mass windows are increased from $\pm 25\,\mathrm{MeV}/c^2$ to $\pm 75\,\mathrm{MeV}/c^2$. Enlarging the charm mass window in this way introduces additional sources of combinatorial and single charm background. Since no significant double charm background is expected (see table 4.5), the background composition remains representative of the signal window. The size of the background samples is $\mathcal{O}(300,000)$ and $\mathcal{O}(2,000,000)$ events for the D^+ and D^+_s modes, respectively. The same background sample is used for the B^+_c search and the CP asymmetry measurement.

It is important to withhold a subset of the signal and background samples from training to validate the performance of the MVA algorithm. In an ideal comparison, the size of the training and testing samples would be the same. However, when the input samples are small, withholding half can have a dramatic effect on the performance. The solution is to use the k-folding technique [167]. Instead, k (= 5 in this case) MVAs are trained, each time withholding a different 20% of the sample for testing. When the response is evaluated, it is done so using the MVA that excluded that event from the training. This allows the majority of the signal and background samples to be used for training, whilst maintaining an unbiased evaluation of the response. The samples are divided into the five sub-samples based on their run number.

4.7.2.2 Input variables

The input variables used in the MVA training are listed in table 4.6 and the distributions of the signal and background samples in these variables are shown in appendix B. The logarithm is taken of the $\mathrm{DLL}_{K\pi}$, p_T and $t/\Delta t$ variables in order to better exploit the separating power in the long tails of these distributions. As the logarithm is only defined for positive numbers, the input variables are shifted and/or reflected in the y axis to make them positive definite. For example, since $\mathrm{DLL}_{K\pi}(\pi)$ is required to be less than 10 in the preselection, the transformation that is applied is $\ln(10 - \mathrm{DLL}_{K\pi}(\pi))$. As the invariant D masses are used in the training, a transformation is applied to prevent the MVA from excluding events in the enlarged background sample based solely on their $m(D)$. Events in the D sidebands ($25 < |m(D) - m_D| < 75\,\mathrm{MeV}/c^2$) are shifted by $+50(-50)\,\mathrm{MeV}/c^2$ if their mass is below (above) the signal window ($|m(D) - m_D| < 25\,\mathrm{MeV}/c^2$).

Particle	Input variable
$B^+_{(c)}$	$t/\Delta t$, χ^2_{vtx}, χ^2_{IP}
$D^+_{(s)}$, $\overline{D}^0$	m, $t/\Delta t$, χ^2_{vtx}
$D^+_{(s)}$ children	$m^2(\pi(K)^+K^-)$, $m^2(\pi^+K^-)$
Tracks	p_{T}, $\mathrm{DLL}_{K\pi}$

Table 4.6: The variables used in the MVA training.

Further preprocessing of the input variables to reduce correlations and transform their shapes into more appropriate forms is experimented with. The transformations applied in the training are:

- **Identity (I)**: no transformation.
- **Uniformisation (U)**: this transforms the signal and background such that they lie between 0 and 1 and their sum follows a uniform distribution.
- **Variable decorrelation (D)**: this applies a linear decorrelation based on the square-root of the covariance matrix.
- **Principal component decomposition (P)**: this transformation rotates the data into a set of linearly uncorrelated variables.

Applying two of these transformations, for example, uniformisation and variable decorrelation, is denoted with UD.

4.7.2.3 Methods

The methods considered are the boosted decision tree and the multilayer perceptron. These make use of different machine-learning techniques in order to achieve the best separation of signal and background.

Boosted decision tree

A boosted decision tree (BDT) combines the results of many "decision trees". A decision tree has a flowchart-like structure, as shown in figure 4.6. The first node (the "root node") takes as input the full training sample and applies a requirement on one of the input variables. The requirement is determined by scanning over each variable in a fixed number of steps and choosing the one that results in the best separation of signal

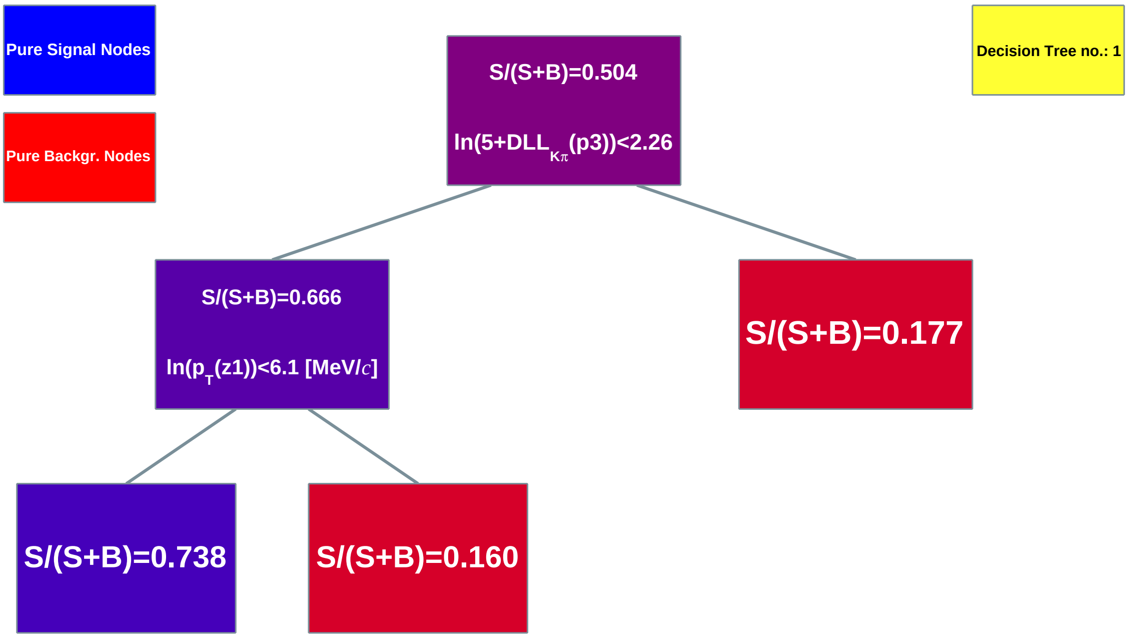


Figure 4.6: An example of a decision tree for the BDT trained to select $B_c^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays. The colour of the node is determined by its purity (number of signal events divided by the sum of signal and background), with blue indicating pure signal and red pure background.

and background. This creates two sub-samples that are then used as inputs to other nodes. The process continues until either the maximum number of requirements have been applied (known as the "maximum depth") or the number of events in the training sample is too low. The resulting "leaf nodes" are then classified as signal or background, depending on whether they are made up of primarily signal or background events.

If an event is misclassified by the decision tree, it is assigned a higher weight in the training of the next tree. By assigning these difficult events a higher weight, the next decision tree is more likely to get the classification right. This is known as "boosting" and was implemented using the AdaBoost (adaptive boost) algorithm [168]. Boosting improves the performance and statistical stability of the classifier. The process of training a decision tree and reweighting the training sample is repeated until a "forest" of decision trees is constructed. The forest is combined into a single classifier by taking the weighted average of the individual decision trees.

At their core, BDTs are based on a relatively simple one-dimensional optimisation at each node. This makes them insensitive to the inclusion of poorly discriminating input

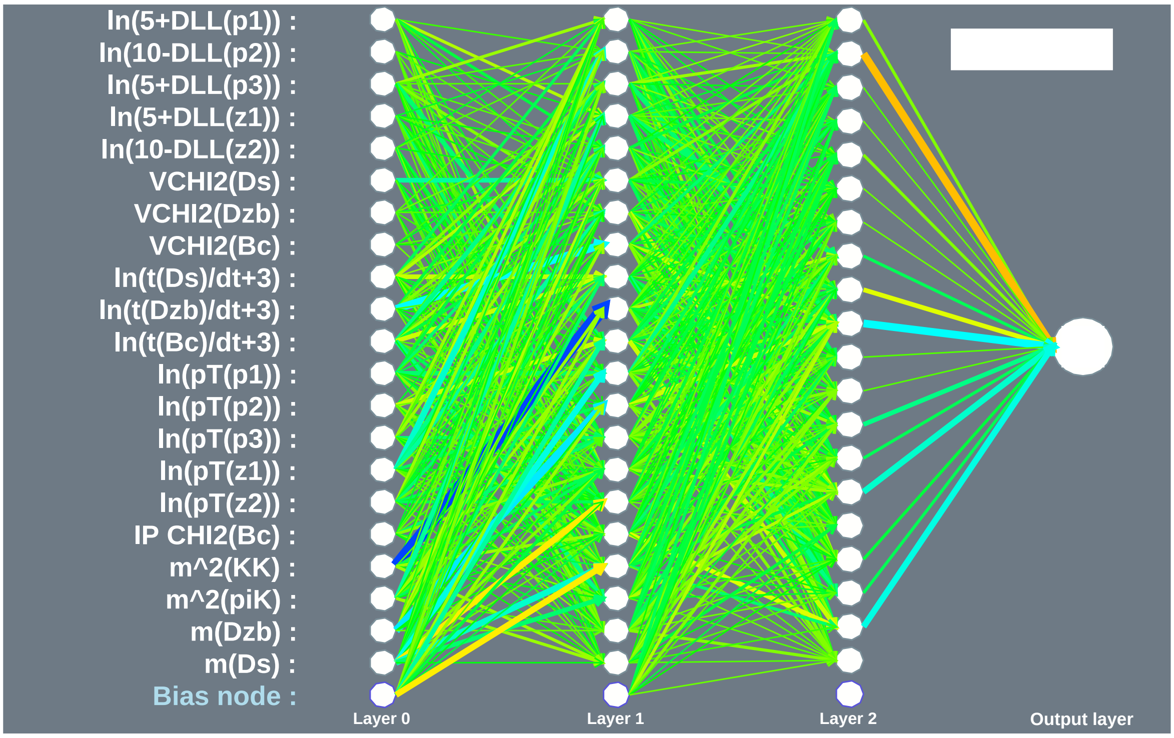


Figure 4.7: The architecture of the MLP trained to select $B_c^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays. Layer 0 is the input layer and layers 1 and 2 are hidden layers.

variables, as requirements will rarely be placed on them. However, this simplicity may limit their performance compared to other more sophisticated methods.

Multilayer perceptron

A multilayer perceptron (MLP) is a class of artificial neural network (ANN). ANNs are based loosely on the human brain, consisting of a collection of interconnected “neurons” (nodes). The complexity of an ANN is reduced by sorting the neurons into layers, where connections are only allowed between a given layer and the following layer. This is known as an MLP. The first layer of an MLP is the input layer, the last is the output layer, and all the layers inbetween are the *hidden* layers. This structure is shown in figure 4.7. If an MLP is trained on n input variables, the input layer consists of n neurons that hold the input values, and the outer layer contains just one number, the MLP response. The hidden layers can learn complex, non-linear features of the training sample to discriminate between signal and background.

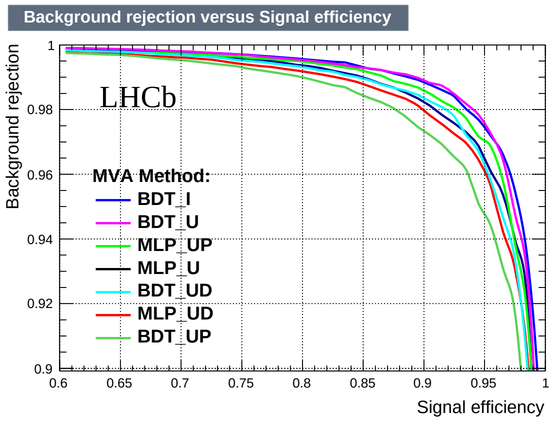


Figure 4.8: ROC curves for the classifiers trained to select $B_c^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays. The BDT and MLP methods were trained with various different transformations applied to the input variables.

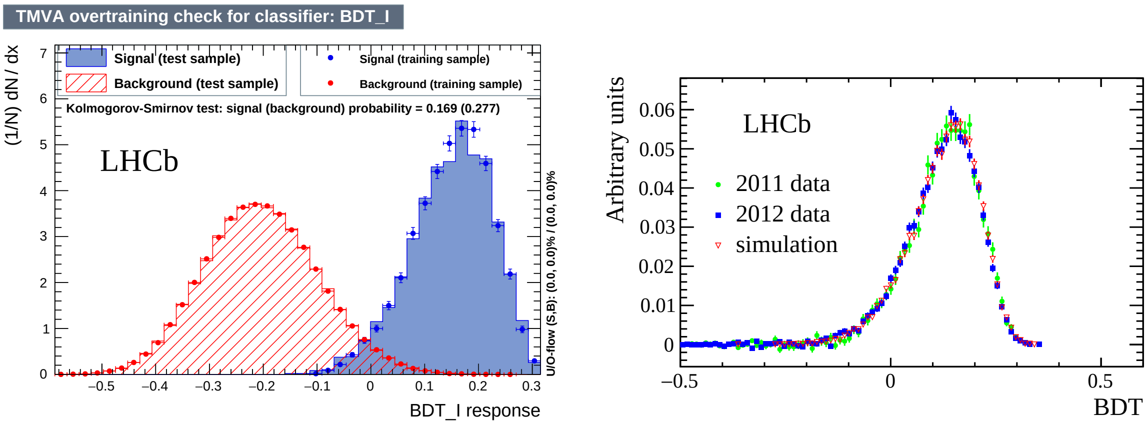
Performance

The performance of the two methods with the different input variable transformations is evaluated using the receiver operator characteristic (ROC) curves. As shown in figure 4.8, these display the signal efficiency against the background rejection. An ideal classifier would lie in the top-right corner, where there is 100% signal efficiency and 100% background rejection. In all channels, the classifiers that come closest to this are the BDT with the identity and the uniform transformations. As the ROC curves are very similar, the simplest of the two, the BDT with an identity transformation, is chosen as the MVA for the analyses. The MLP method and decorrelation transformations are seen to degrade the performance.

The distribution of the BDT response for the signal and background samples is shown in figure 4.9a. The two distributions are well separated, indicating that the BDT can successfully isolate signal. The BDT response is compared between sWeighted B^+ data and simulation in figure 4.9b. The agreement is excellent, which means that the simulation can be used to reliably calculate the efficiency of the requirement placed on the BDT response.

4.7.2.4 Overtraining

Overtraining results in the properties of a classifier being determined by a small number of candidates in the input samples, rather than general properties of the signal and



(a) Signal and background samples, split by training and testing.

(b) Comparison of sWeighted B^+ data with simulation.

Figure 4.9: The response of the MVA trained to select $B_c^+ \to D_s^+ \overline{D}^0$ with $\overline{D}^0 \to K^+\pi^-$ decays.

background candidates. This occurs when not enough data is used in the training to tune too many parameters of the MVA algorithm. Overtraining leads to an overestimation of the performance of the MVA in the training sample, which can be identified by comparing the performance between the training sample and an independent testing sample.

The Kolmogorov-Smirnov test [169] is used to test if two samples came from the same underlying distribution. The output is between 0 and 1, where 1 indicates the samples are identical and 0 that they are incompatible. The result of the Kolmogorov-Smirnov test for the training and testing samples is displayed in figure 4.9a. The maximum depth of the BDT is restricted to 2 and hence no significant overtraining is observed.

4.7.3 Optimisation

The output of the MVA is a single classifier, on which a requirement can be placed to reject background. The sensitivity to the rare B_c^+ signals and the statistical uncertainty of the *CP* asymmetry measurements depend on the value chosen for this requirement. A procedure was therefore put in place to determine the optimal value. The nature of the B_c^+ search and *CP* asymmetry measurements are very different, so they are optimised according to different criteria.

B_c^+ decays

The branching fractions of the B_c^+ decays are unknown, therefore the optimisation needs to be performed without any assumptions on the size or presence of a signal. This is achieved using the Punzi figure of merit [170],

$$\mathrm{FoM}_{\mathrm{Punzi}} = \frac{\varepsilon_S}{\sqrt{N_B} + a/2}, \tag{4.10}$$

where ε_S is the signal efficiency, N_B is the number of background events after a requirement has been placed on the BDT response and a is the target significance. A value of 5 is used for a, corresponding to the significance required for an observation. The signal efficiency is determined by counting the number of $B_c^+ \to D_{(s)}^+ D$ events in simulation before and after the BDT requirement.

The calculation of the number of background events in the signal region is complicated by two factors: first, the signal region was blinded and second, even if it was not, the signal content (and therefore background content) is unknown. The number of background events is therefore determined by extrapolating from the B upper mass sideband ($5350 < m(D_{(s)}^+ D) < 6200\,\mathrm{MeV}/c^2$).

Since the background is well-described by an exponential function, an analytical expression is derived to calculate the number of events in the signal window. For an exponential function, $f(m) = ae^{-bm}$, the number of events in the low-mass half N_l and high-mass half N_h of the B upper mass sideband, is given by

$$N_l = \int_{m_l}^{m_l+\Delta m/2} ae^{-bm} dm = \frac{a}{b} e^{-bm_l} \left[1 - e^{-b\Delta m/2}\right], \tag{4.11}$$

$$N_h = \int_{m_l+\Delta m/2}^{m_h} ae^{-bm} dm = \frac{a}{b} e^{-b(m_l+\Delta m/2)} \left[1 - e^{-b\Delta m/2}\right], \tag{4.12}$$

where $m_l = 5350\,\mathrm{MeV}/c^2$, $m_h = 6200\,\mathrm{MeV}/c^2$ and $\Delta m = m_h - m_l$. The exponential shape parameters can therefore be estimated as

$$b = \frac{2}{\Delta m} \ln\left(\frac{N_l}{N_h}\right), \tag{4.13}$$

$$a = \frac{bN_l}{1 - e^{-b\Delta m/2}} e^{bm_l}, \tag{4.14}$$

The number of background events in a $\pm 20\,\mathrm{MeV}/c^2$ window around the B mass is then given by

$$N_B = \int_{m_B-20}^{m_B+20} ae^{-bm}dm = \frac{a}{b}e^{-bm_B}(e^{20b} - e^{-20b}), \tag{4.15}$$

where $m_B = 6275\,\mathrm{MeV}/c^2$ for the B_c^+ search optimisation. This way, N_B can be determined for a large number of BDT requirements without performing non-linear fits to the background distribution.

As shown in figure 4.4, the background levels for $B_{(c)}^+ \to D_{(s)}^+ \overline{D}^0$ and $B_{(c)}^+ \to D_{(s)}^+ D^0$ candidates are similar, differing by less than a factor 2. The optimum requirement is therefore expected to be similar and these decays are optimised together. This is consistent with their treatment in the training of the MVA classifiers. The value of N_B used in the optimisation is the average of the background estimated with $B_{(c)}^+ \to D_{(s)}^+ \overline{D}^0$ and $B_{(c)}^+ \to D_{(s)}^+ D^0$ candidates.

The signal retention, background retention and Punzi figure of merit are shown as a function of the BDT requirement in figure 4.10a. The peak of the Punzi figure of merit corresponds to a region of phase space where the signal efficiency is strongly dependent on the BDT requirement. This region also shows some statistical fluctuations due to the extremely high background rejection ($\sim 99\%$). Choosing the peak of the Punzi figure of merit as the working point is therefore dangerous for two reasons. First, the efficiencies are very sensitive to data/simulation discrepancies. Second, the peak is sometimes poorly defined with multiple local maxima that have very similar values for the Punzi figure of merit, but strongly varying signal and background retentions. This can lead to exceedingly tight requirements with less than one background event expected in the signal region, resulting in problems with the stability of the mass fit. The solution is to choose the first BDT requirement with a Punzi figure of merit greater than 90% of the maximum as the working point.

The value of the optimal BDT requirement and its efficiency on the $B_{(c)}^+$ simulation and the B^+ data is shown in table 4.7. The signal efficiency for B_c^+ is higher than for B^+ since variables that differ between the two were included in the training, for example, $t/\Delta t$. The single charm $B^+ \to K^+K^-\pi^+\overline{D}^0$ background to $B^+ \to D_s^+\overline{D}^0$ is neglected, contributing an additional $\sim 1\%$ systematic uncertainty (see section 5.2.4). The efficiencies for data and simulation are in agreement within their uncertainties.

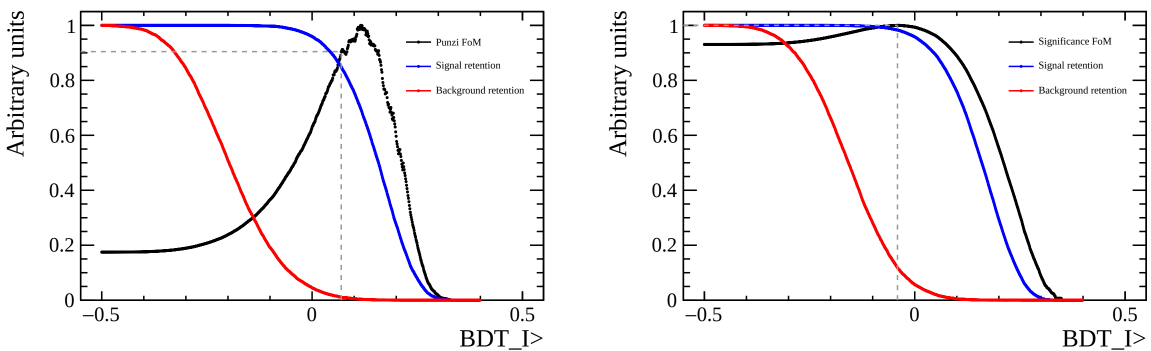
B^+ decays

Large $\mathcal{O}(10^3)$ yields are observed for the B^+ signal. In this case, the significance figure of merit was used,

$$\mathrm{FoM}_{\mathrm{significance}} = \frac{N_S}{\sqrt{N_S + N_B}}, \tag{4.16}$$

where N_S is the signal yield and as before, N_B is the background yield. Maximising the significance figure of merit maximises the statistical significance of the signal yields. This also minimises the statistical uncertainty on the CP asymmetry measurement. The background yield was estimated using equation 4.15 with $m_B = 5280\,\mathrm{MeV}/c^2$ using $B^+ \to D^+_{(s)}\overline{D}^0$ candidates only. The signal yield was determined by multiplying the B^+ yield in data without any BDT requirement with the signal efficiency calculated from simulation.

The signal retention, background retention and significance figure of merit are shown in figure 4.10b. The peak of the figure of merit is not strongly dependent on the BDT requirement and larger signal and background retentions mean the result is not susceptible to statistical fluctuations. Due to this better behaviour, the maximum of the significance figure of merit is chosen as the working point. The value of the optimal requirement and the efficiencies of B^+ data and simulation are shown in table 4.7. The efficiencies are higher than the optimisation for B_c^+ decays due to the use of a different figure of merit. The efficiencies measured using data and simulation are consistent within their uncertainties.



(a) Punzi figure of merit. **(b)** Significance figure of merit.

Figure 4.10: The figure of merit, signal retention and background retention as a function of the BDT requirement for $B^+_{(c)} \to D^+_s\overline{D}^0$ candidates with $\overline{D}^0 \to K^+\pi^-$ decays. The dashed grey lines indicate the optimisation points.

Training channel	MVA requirement	$\varepsilon_{B_c^+\,\text{simulation}}$	$\varepsilon_{B^+\,\text{simulation}}$	$\varepsilon_{B^+\,\text{data}}$
$B_c^+ \to D_s^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-$	0.069	84.3 ± 0.4	76.8 ± 0.3	74.9 ± 0.4
$B_c^+ \to D_s^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	0.048	84.8 ± 0.7	78.7 ± 0.5	77.6 ± 0.7
$B_c^+ \to D^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-$	0.032	86.8 ± 0.4	83.1 ± 0.3	82.4 ± 2.2
$B_c^+ \to D^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	0.033	85.1 ± 0.7	81.2 ± 0.5	72.6 ± 5.1
$B^+ \to D_s^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-$	-0.041	—	98.2 ± 0.1	97.2 ± 0.4
$B^+ \to D_s^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	-0.032	—	97.4 ± 0.2	97.6 ± 0.7
$B^+ \to D^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-$	0.005	—	93.9 ± 0.2	94.5 ± 3.3
$B^+ \to D^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	0.036	—	87.6 ± 0.4	79.9 ± 5.6

Table 4.7: Efficiency of the MVA requirement on $B_{(c)}^+ \to D_{(s)}^+ D$ decays, in %. The number of B candidates is evaluated by fitting the data $\pm 60\,\text{MeV}/c^2$ around the nominal B mass with the B^+ signal model (see section 5.2.1) for simulation or the sum of the B^+ signal model and an exponential function for data. The uncertainties are statistical only. The $B^+ \to D_s^+ \overline{D}^0$ modes have an additional systematic uncertainty of $\sim 1\%$ due to the single charm background $B^+ \to K^+K^-\pi^+\overline{D}^0$.

Chapter 5

Search for B_c^+ decays to two charm mesons

"There is nothing like looking, if you want to find something."
— J. R. R. Tolkien

Decays of B_c^+ to two charm mesons, $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$, are sensitive to the CKM angle γ. This chapter describes the first ever search performed for these decays using Run I data collected by the LHCb experiment and has been published in reference [1]. First, an overview of the analysis strategy is given (section 5.1), followed by discussion of the fit model (section 5.2), the efficiencies (section 5.3), the mass fits (section 5.4) and the fit validation (section 5.5). The systematic uncertainties are described (section 5.6) and lastly, the results for the B_c^+ branching fractions are presented (section 5.7).

5.1 Analysis strategy

The branching fractions of the B_c^+ decays are measured relative to the high-yield normalisation channel $B^+ \to D_{(s)}^+ \overline{D}^0$. No $B^+ \to D_{(s)}^+ D^0$ signal is expected due to the flavour structure, so $B^+ \to D_{(s)}^+ \overline{D}^0$ decays are used to normalise both $B_c^+ \to D_{(s)}^{(*)+} \overline{D}^{(*)0}$ and $B_c^+ \to D_{(s)}^{(*)+} D^{(*)0}$. The charm mesons are reconstructed in the following final states: $D_s^+ \to K^+ K^- \pi^+$, $D^+ \to K^- \pi^+ \pi^+$, $\overline{D}^0 \to K^+ \pi^-$ or $K^+ \pi^- \pi^+ \pi^-$. The soft neutral particle from the decays of the excited charm mesons is not reconstructed.

The branching fractions of B_c^+ decays to fully reconstructed final states are given by

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_{(s)}^+ D)}{\mathcal{B}(B^+\to D_{(s)}^+\overline{D}^0)} = \frac{N(B_c^+\to D_{(s)}^+ D)}{N(B^+\to D_{(s)}^+\overline{D}^0)}\frac{\varepsilon(B^+\to D_{(s)}^+\overline{D}^0)}{\varepsilon(B_c^+\to D_{(s)}^+ D)}, \tag{5.1}$$

where f_c/f_u is the ratio of the B_c^+ to B^+ production cross-sections, N stands for the signal yields and ε the total efficiencies. The invariant mass distributions of $B_c^+\to D_{(s)}^{*+} D$ and $B_c^+\to D_{(s)}^+ D^*$ decays are very similar, so the sum of their branching fractions, weighted by the branching fraction of the excited charm meson, is measured,

$$\begin{aligned}\frac{f_c}{f_u}&\frac{\mathcal{B}(B_c^+\to D_{(s)}^{*+} D)\mathcal{B}(D_{(s)}^{*+}\to D_{(s)}^+\pi^0/\gamma) + \mathcal{B}(B_c^+\to D_{(s)}^+ D^*)}{\mathcal{B}(B^+\to D_{(s)}^+\overline{D}^0)} \\ &= \frac{N(B_c^+\to D_{(s)}^{*+} D) + N(B_c^+\to D_{(s)}^+ D^*)}{N(B^+\to D_{(s)}^+\overline{D}^0)}\frac{\varepsilon(B^+\to D_{(s)}^+\overline{D}^0)}{\varepsilon(B_c^+\to D_{(s)}^{*+} D, D_{(s)}^+ D^*)},\end{aligned} \tag{5.2}$$

where $\varepsilon(B_c^+\to D_{(s)}^{*+} D, D_{(s)}^+ D^*)$ is the average efficiency of $B_c^+\to D_{(s)}^{*+} D$ and $B_c^+\to D_{(s)}^+ D^*$ decays. For $\overline{D}^{*0}$ and D_s^{*+} the branching fraction to a soft neutral particle and $\overline{D}^0$ and D_s^+, respectively, is 100%. For D^{*+}, some sensitivity is lost due to the decay of $D^{*+}\to D^0\pi^+$, which occurs approximately two-thirds of the time [3]. Branching fractions of $B_c^+\to D_{(s)}^{*+} D^*$ decays are corrected for $\mathcal{B}(D_{(s)}^{*+}\to D_{(s)}^+\pi^0/\gamma)$,

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_{(s)}^{*+} D^*)}{\mathcal{B}(B^+\to D_{(s)}^+\overline{D}^0)} = \frac{1}{\mathcal{B}(D_{(s)}^{*+}\to D_{(s)}^+\pi^0/\gamma)}\frac{N(B_c^+\to D_{(s)}^{*+} D^*)}{N(B^+\to D_{(s)}^+\overline{D}^0)}\frac{\varepsilon(B^+\to D_{(s)}^+\overline{D}^0)}{\varepsilon(B_c^+\to D_{(s)}^{*+} D^*)}. \tag{5.3}$$

The LHCb experiment has measured $R = \frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to J/\psi\pi^+)}{\mathcal{B}(B^+\to J/\psi K^+)}$ at centre-of-mass energies of 7 [70] and 8 TeV [69] in the region $0 < p_T < 20\,\mathrm{GeV}/c^2$ and $2.0 < y < 4.5$, where p_T and y refer to the transverse momentum and rapidity of the B meson. As branching fractions are independent of the centre-of-mass energy, comparing the two R measurements gives the dependence of f_c/f_u on energy. As the measurements of R are compatible within 1σ, the 2011 and 2012 data sets are analysed together.

5.2 Fit model

The signal yields are determined by fitting to the constrained mass distributions of the $B_{(c)}^+\to D_{(s)}^{(*)+} D^{(*)}$ candidates. The fit model is comprised of six components: the signals for the fully reconstructed B_c^+ and B^+ decays (section 5.2.1), the signal for B_c^+

decays with one excited charm meson in the final state (section 5.2.2), the signal for B_c^+ decays with two excited charm mesons in the final state (section 5.2.3), the single charm background (section 5.2.4) and the combinatorial background. The models for the signal shapes and single charm background are constrained from simulation, while the combinatorial background is found to be well described by the sum of an exponential function and a constant with all normalisation and shape parameters floating.

5.2.1 $B_{(c)}^+ \to D_{(s)}^+ D$ model

The constrained mass distributions of $B_{(c)}^+ \to D_{(s)}^+ D$ candidates are described by the sum of two Crystal Ball functions (a "double Crystal Ball", DCB) [171]. CB functions are commonly used to model processes with final state radiation. The choice of the DCB function is empirically motivated.

A Crystal Ball (CB) function consists of a Gaussian core with a power law tail and is given by

$$\mathrm{CB}(m|\alpha, n, \mu, \sigma) = N \begin{cases} \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) & \text{for } \frac{(m-\mu)}{\sigma} > -\alpha \\ A\left(B - \frac{(m-\mu)}{\sigma}\right)^{-n} & \text{for } \frac{(m-\mu)}{\sigma} \leq -\alpha, \end{cases} \tag{5.4}$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{|\alpha|^2}{2}\right), \tag{5.5}$$

$$B = \frac{n}{|\alpha|} - |\alpha|, \tag{5.6}$$

$$C = \frac{n}{|\alpha|}\frac{1}{n-1} \exp\left(-\frac{|\alpha|^2}{2}\right), \tag{5.7}$$

$$D = \sqrt{\frac{\pi}{2}}\left(1 + \mathrm{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right), \tag{5.8}$$

$$N = \frac{1}{\sigma(C+D)}. \tag{5.9}$$

The shape parameters are: the mean μ and width σ of the Gaussian core; the power of the tail n; and the number of sigma away from the mean that the distribution changes from Gaussian to a power law α. The sum of two Crystal Ball functions therefore has eight shape parameters, plus an additional parameter describing the relative normalisation f. To reduce the complexity, the mean was shared between the two Crystal Ball functions and the power of each tail was fixed to 2 since n and α are highly correlated. The power law tails were constrained to be on opposite sides. The parametrisation of the DCB was therefore

$$\mathrm{DCB}(m|\alpha_1, \alpha_2, \mu, \sigma, \sigma_{\text{ratio}}) = f\mathrm{CB}(m|\alpha_1, \mu, \sigma) + (1-f)\mathrm{CB}(m|\alpha_2, \mu, \sigma, \sigma_{\text{ratio}}). \quad (5.10)$$

The fits to the constrained mass distributions of B^+ simulation and data are shown in figure 5.1 and the fits to the B_c^+ simulation are shown in figure 5.2. The fits to simulation are performed with all six shape parameters floating, while for data, all shape parameters are fixed to the values obtained from simulation. The model derived from simulation provides an excellent description of the data.

As large B^+ signals are expected, some shape parameters are allowed to float in the final fit to account for any remaining data/simulation discrepancies. The tail parameters, ratio of Crystal Ball widths and relative normalisation are fixed from simulation, leaving the mean and width floating. As only very small B_c^+ signals are expected, all parameters apart from the mean are fixed from simulation. The mean of the B_c^+ model is fixed to the world-average value $m_{B_c^+} = 6274.9 \pm 0.8\,\mathrm{MeV}/c^2$ [3], which is around $1\,\mathrm{MeV}/c^2$ greater than the value used in the generation of the simulation.

5.2.2 $B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$ model

For B_c^+ decays to final states containing one excited charm meson, both $B_c^+ \to D_{(s)}^{*+} D$ and $B_c^+ \to D_{(s)}^+ D^*$ contribute. As their relative proportions are unknown, it is assumed that $\mathcal{B}(B_c^+ \to D_{(s)}^{*+} D) = \mathcal{B}(B_c^+ \to D_{(s)}^+ D^*)$. Since $\mathcal{B}(D^{*+} \to D^+ \pi^0/\gamma) \sim 32\%$, the decay of $B_c^+ \to D^+ D^*$ is given a larger weight than $B_c^+ \to D^{*+} D$.

Templates are derived by fitting the constrained mass distribution of $B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$ candidates with the sum of Gaussian functions, each with width $\sigma = 15\,\mathrm{MeV}/c^2$, and a spacing of $15\,\mathrm{MeV}/c^2$ between them,

$$f(m|a_i, \sigma, m_{min}, m_{max}) = \sum_{i=0}^{n} \frac{a_i}{\sqrt{2\pi}\sigma} \exp\left(\frac{(m - m_{min} - i\sigma)^2}{2\sigma^2}\right), \quad (5.11)$$

where $m_{min} = 5400\,\mathrm{MeV}/c^2$ is the low-mass edge of the histogram, $m_{max} = 6400\,\mathrm{MeV}/c^2$ is the high-mass edge of the histogram and $n = (m_{max} - m_{min})/\sigma$ is the number of

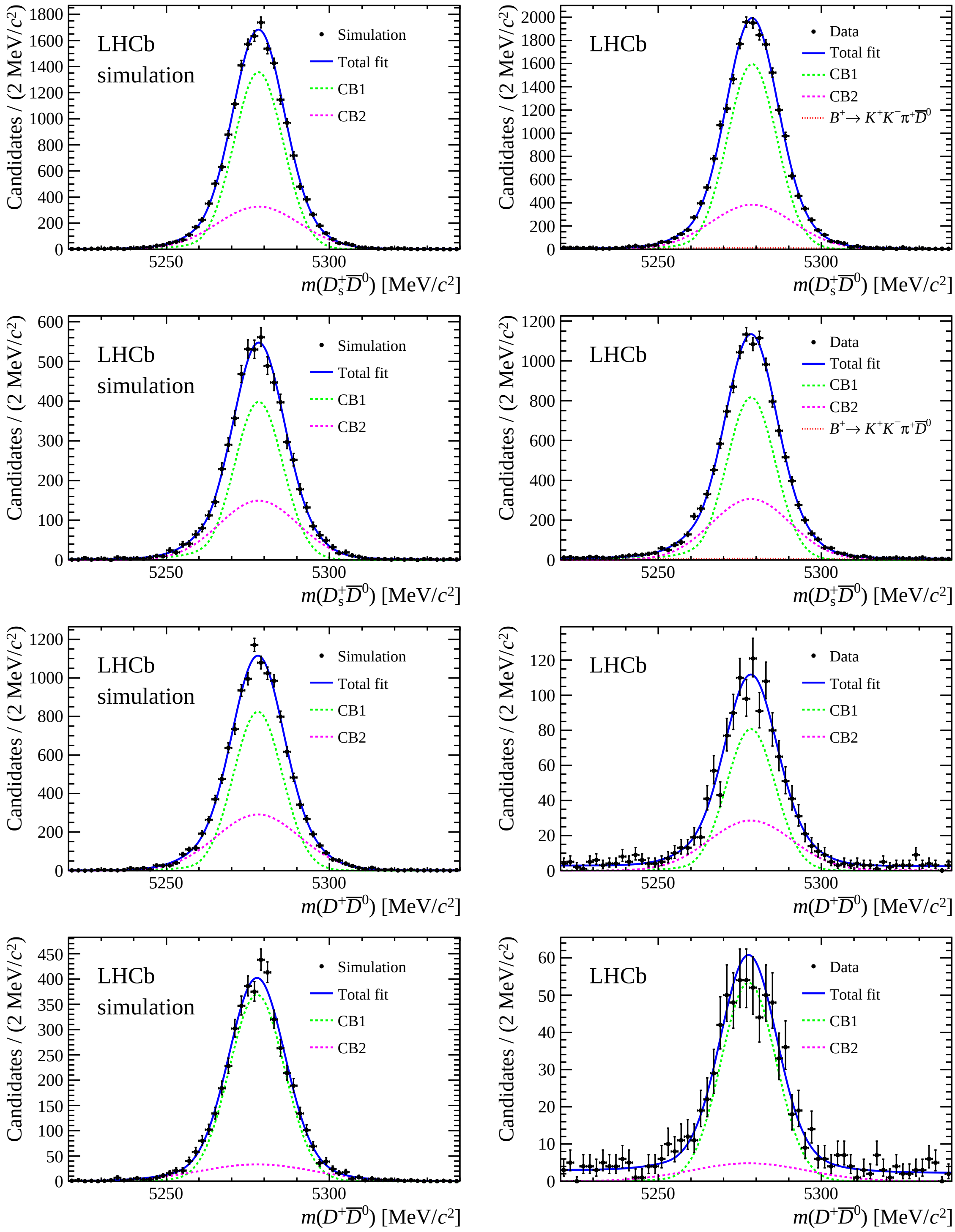


Figure 5.1: Constrained mass distributions of $B^+ \to D^+_{(s)}\overline{D}^0$ candidates; left are from simulation and right are from data. Top and third rows are with $\overline{D}^0 \to K^+\pi^-$ final states, second and bottom rows with $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ final states. Overlaid are fits to a double Crystal Ball function plus background components. For the simulation, the shape parameters are floating, while for the data all shape parameters have been fixed to the values obtained from the fit to the simulated events.

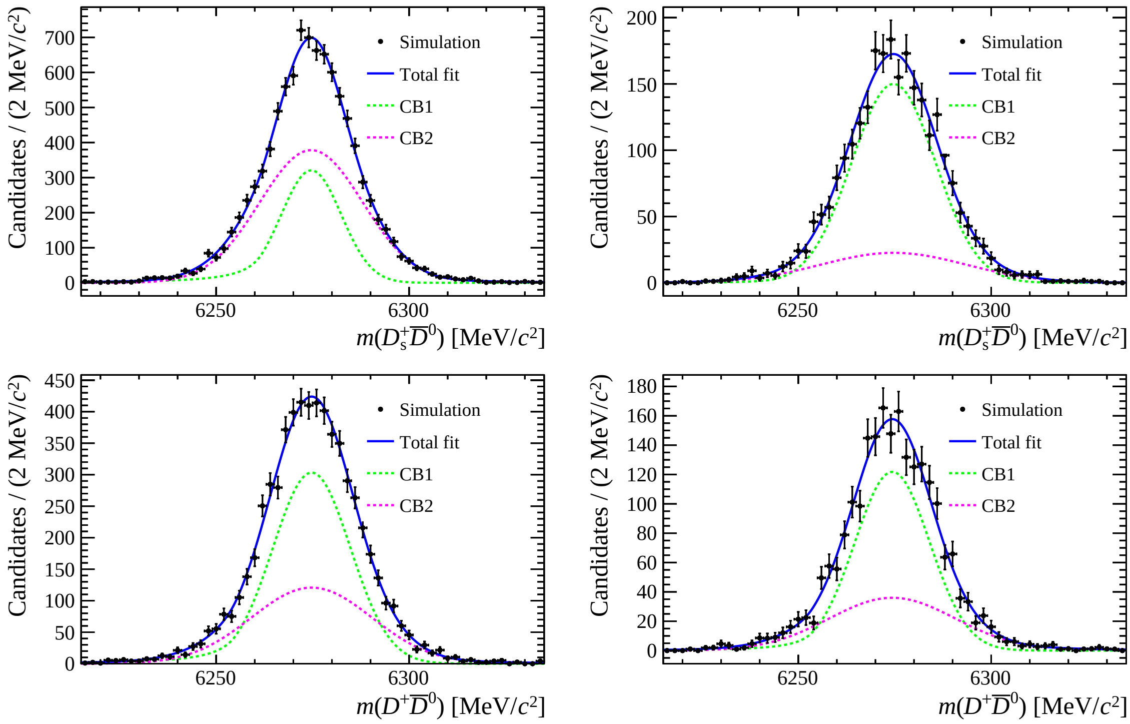


Figure 5.2: Constrained mass distributions of $B_c^+ \to D_{(s)}^+ \overline{D}^0$ candidates from simulation; left are with $\overline{D}^0 \to K^+\pi^-$ final states and right are with $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ final states. Overlaid are fits to a double Crystal Ball function with all shape parameters floating.

Gaussian functions. The normalisation of each Gaussian function is given by a_i, however, this is fixed to 0 if there are less than 10 entries within 2σ of the mean.

Figure 5.3 shows the fits to the constrained mass distribution of $B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$ simulation. The angular distribution of $D^* \to D\pi^0$ decays results in a twin peak structure, whereas decays of $D^* \to D\gamma$ results in a single wide peak. Decays of D_s^{*+} are dominated by the γ mode, hence the twin peak structure is less prominent in the $B_c^+ \to D_s^{*+} D$ model than in the $B_c^+ \to D^{*+} D$ or $B_c^+ \to D_{(s)}^+ D^*$ models.

5.2.3 $B_c^+ \to D_{(s)}^{*+} D^*$ model

The polarisation of B_c^+ decays to final states containing two excited charm mesons is unknown. A 2 : 1 ratio of the transverse to longitudinal polarisations is assumed. As before, the parametrisation is given by equation 5.11. In this case, since the constrained mass distributions of $B_c^+ \to D_{(s)}^{*+} D^*$ decays are wider than $B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$ decays,

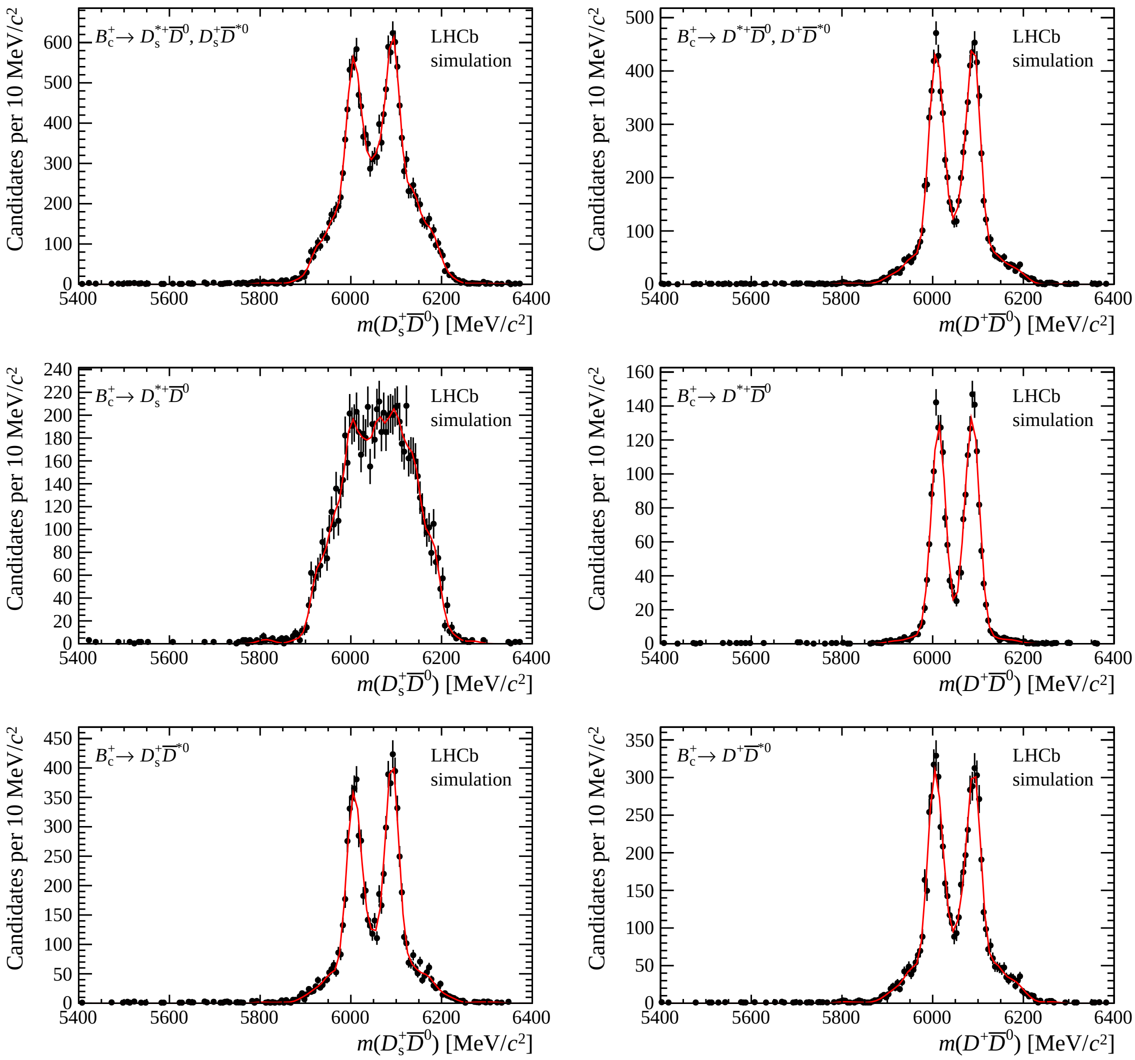


Figure 5.3: Constrained mass distributions of $B_c^+ \to D_{(s)}^{*+}\overline{D}^0$, $D_{(s)}^+\overline{D}^{*0}$, $\overline{D}^0 \to K^+\pi^-$ candidates from simulation. The top row figures show the weighted sum of the two components (shown in the middle and bottom row). The middle and bottom row figures show the individual components, used for the systematic studies. Fits to a sum of Gaussian functions are overlaid.

a width and spacing of $20\,\mathrm{MeV}/c^2$ is used. Figure 5.4 shows the fits to the constrained mass distributions of $B_c^+ \to D_{(s)}^{*+} D^*$ simulation.

5.2.4 $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ model

Potential sources of peaking background arise from B decays to the same final state as the signal that occur without one or both of the intermediate charm mesons. These are known as "single charm", *e.g.* $B^+ \to K^+ K^- \pi^+ \overline{D}^0$, and "charmless", *e.g.* $B^+ \to K^+ K^- \pi^+ K^+ \pi^-$, backgrounds. As these backgrounds don't peak in the charm masses, they are expected to be heavily suppressed by the charm mass windows and the requirement on the BDT response. However, there may be a small component remaining in the final selection.

The presence of single charm or charmless backgrounds can be determined by checking for B^+ peaks in data from the charm mass sidebands ($25 < |m(D) - m_D| < 75\,\mathrm{MeV}/c^2$). A small B^+ peak was observed using data from the D_s^+ sidebands, indicating the presence of $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ decays. This background could be removed by tightening the decay time significance requirement as was done to remove $B^+ \to \pi^+ \pi^+ \pi^- \overline{D}^0$ decays; however, the D_s^+ lifetime is short (half that of D^+), so this would result in a relatively large loss of signal. Instead, $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ decays are modelled in the fit.

Although the width of $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ and $B^+ \to D_s^+ \overline{D}^0$ decays is expected to be the same in the invariant B mass, it differs between the two in the constrained B mass. This is due to the charm mass constraint, which improves the mass resolution of signal decays, but degrades it for single charm and charmless backgrounds. The width is determined by fitting a Gaussian function to the constrained mass of $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ simulation, as shown in figure 5.5.

The constrained B mass fixes the charm meson masses to their known values, which results in events in the upper mass sideband ($25 < m(D) - m_D < 75\,\mathrm{MeV}/c^2$) being shifted $50\,\mathrm{MeV}/c^2$ down, and events in the lower mass sideband ($25 < m_D - m(D) < 75\,\mathrm{MeV}/c^2$) being shifted $50\,\mathrm{MeV}/c^2$ up. Instead of fitting these two displaced peaks, the $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ yield is determined from the invariant B mass. The invariant B mass is fitted with the sum of a Gaussian function and an exponential function, as shown in figure 5.6. The $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ yield in the signal window is taken as the yield from the fit scaled down by a factor of two to correct for the larger size of the sidebands compared to the signal window. The yield of $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ is $\mathcal{O}(1\%)$ the size of the $B^+ \to D_s^+ \overline{D}^0$ yield. The model for $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ decays in the final fit is a Gaussian

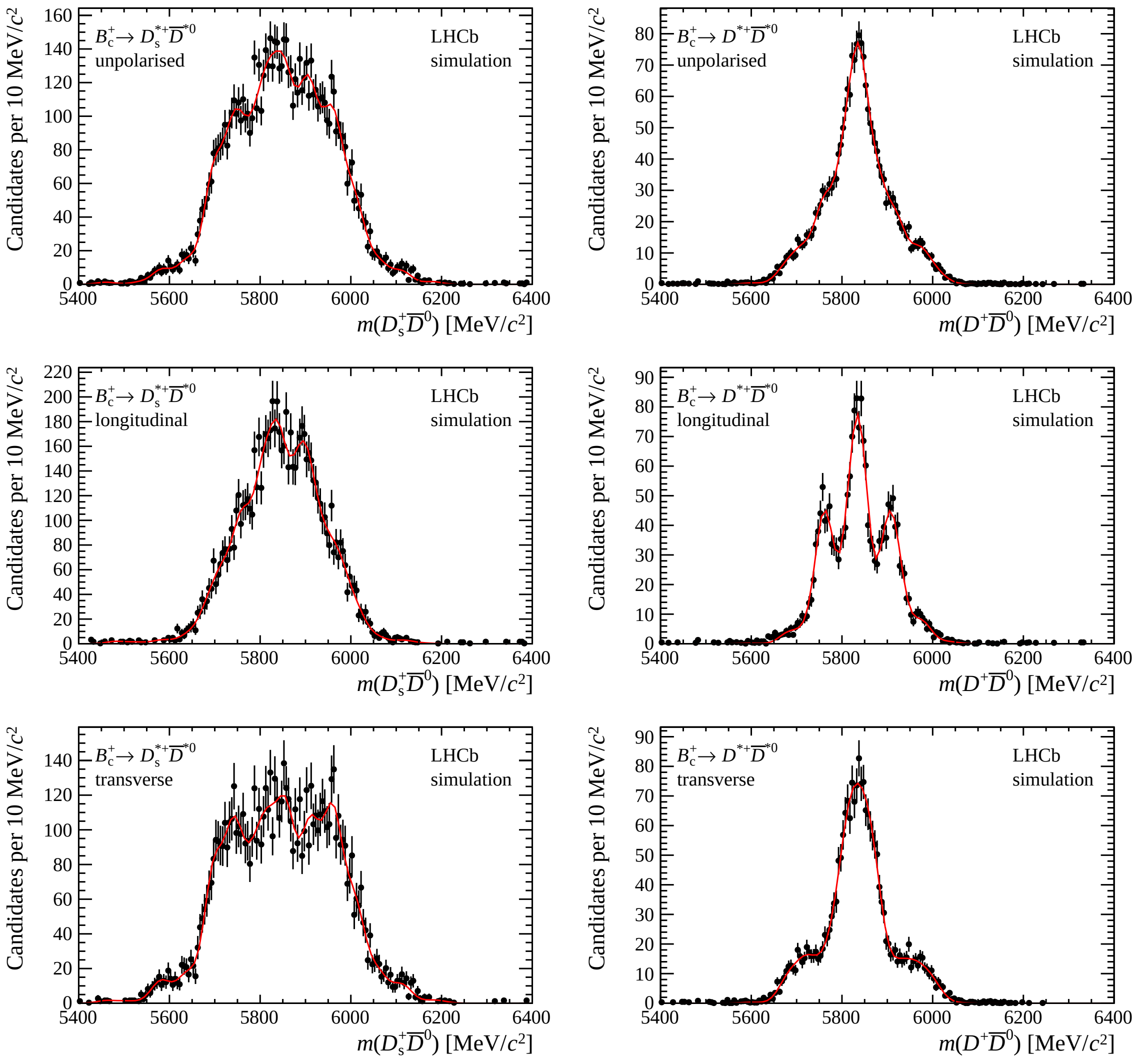


Figure 5.4: Constrained mass distributions of $B_c^+ \to D_{(s)}^{*+}\overline{D}^{*0}$, $\overline{D}^0 \to K^+\pi^-$ candidates from simulation. The top row figures show the weighted sum of the two components (shown in the middle and bottom row). The middle and bottom row figures show the individual components, used for the systematic studies. Fits to a sum of Gaussian functions are overlaid.

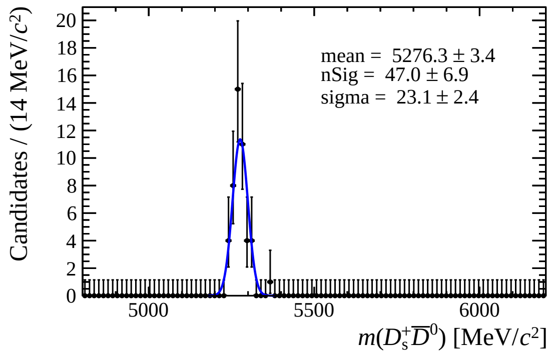


Figure 5.5: Distribution of the constrained mass of $B^+ \to K^+K^-\pi^+\overline{D}^0$ simulation, reconstructed as $B^+ \to D_s^+\overline{D}^0$. A fit to the data with a Gaussian function in the range $5200 - 5360\,\mathrm{MeV}/c^2$ is overlaid.

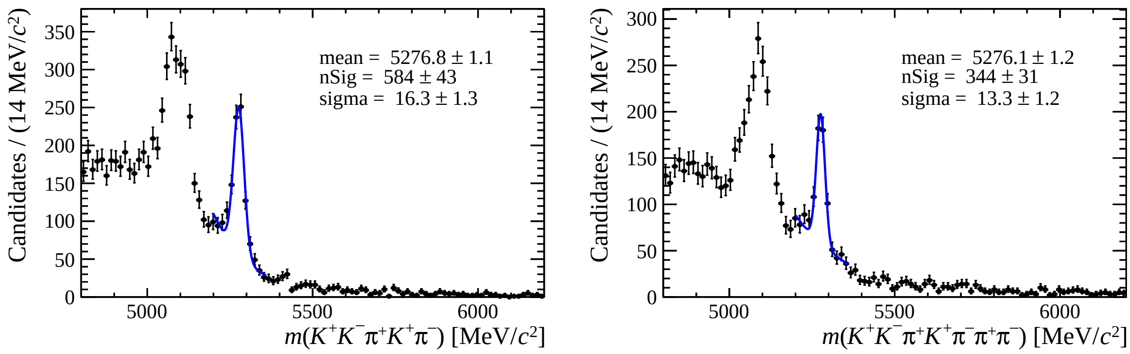


Figure 5.6: Distributions of the invariant mass of $B^+ \to D_s^+\overline{D}^0$ candidates using data from the D_s^+ sidebands ($25 < |m(D_s^+) - m_{D_s^+}| < 75\,\mathrm{MeV}/c^2$). A fit to the data with the sum of a Gaussian function and an exponential function in the range $5200 - 5360\,\mathrm{MeV}/c^2$ is overlaid.

function,

$$\mathrm{Gaus}(m|\sigma,\mu) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{5.12}$$

with its normalisation and width σ fixed, and its mean μ shared with $B^+ \to D_{(s)}^+\overline{D}^0$ decays.

As a cross check to make sure the B^+ peak in the D_s^+ sidebands is truly from $B^+ \to K^+K^-\pi^+\overline{D}^0$ decays and not due to the tails of the $B^+ \to D_s^+\overline{D}^0$ signal, the sidebands are divided evenly into near ($25 < |m(D) - m_D| < 50\,\mathrm{MeV}/c^2$) and far ($50 < |m(D) - m_D| < 75\,\mathrm{MeV}/c^2$) sub-samples. If the peak were due to the tails of the signal then the sub-samples nearer the peak would contain significantly more events than those further away. The $B^+ \to K^+K^-\pi^+\overline{D}^0$ yields are consistent in both, confirming the nature of the background as single charm.

Analogously, the the single charm $B_c^+ \to K^+K^-\pi^+D^0$ decay may affect the $B_c^+ \to D_s^+D^0$ channel. However, as this background is expected to be $\mathcal{O}(1\%)$ the size of the signal, it is considered negligible due to the small signal yields expected.

5.3 Efficiencies

To determine the branching fraction of the B_c^+ decays relative to the B^+ decays, knowledge of the relative efficiency $\varepsilon_{B_c^+}/\varepsilon_{B^+}$ is required (see equations 5.1, 5.2 and 5.3). Although considering the same final state results in the cancellation of many effects, differences between B^+ and B_c^+ can still have a large effect on the relative efficiency. First, the B_c^+ lifetime is approximately a third that of B^+, resulting in a lower reconstruction efficiency. Second, the B_c^+ mass is $\sim 1\,\text{GeV}/c^2$ greater than the B^+, resulting in daughters with harder momentum spectra. Finally, differences in the B production spectrum (p_{T}, y) affect the geometric acceptance.

The relative efficiency is determined from simulation. Only events with all $B_{(c)}^+$ daughters within the LHCb acceptance were simulated. Correcting the simulation for the efficiency of this acceptance requirement ε_{dau}, gives the number of events generated in 4π. However, the f_c/f_u measurements at 7 and 8 TeV correspond to the kinematic region $2.0 < y < 4.5$ and $p_{\text{T}} < 20\,\text{GeV}/c$. A second correction is therefore applied to impose a kinematic region similar to that of the f_c/f_u measurements, $2.0 < y < 4.5$ and $p_{\text{T}} > 4\,\text{GeV}/c$, with an efficiency ε_{kin}. This is the same kinematic region as is used for the data. The reconstruction efficiency is therefore given by

$$\varepsilon = \frac{N_{\text{rec}}}{N_{\text{gen,dau}}}\frac{\varepsilon_{\text{dau}}}{\varepsilon_{\text{kin}}}, \tag{5.13}$$

where N_{rec} is the number of reconstructed events in the final selection and $N_{\text{gen,dau}}$ is the number of events generated in the LHCb acceptance.

The values of ε_{dau} and ε_{kin} are calculated from "generator level" MC, *i.e.* without simulation of the detector response. The efficiency of the acceptance requirement is lower for B_c^+ decays than for B^+ decays ($\sim 13\%$ and $\sim 16\%$, respectively) due to the larger opening angle between the two charm mesons. The kinematic efficiency of B_c^+ decays is also lower than for B^+ decays ($\sim 19.4\%$ and $\sim 20.4\%$, respectively) since B_c^+ mesons are produced more centrally. This results in a B_c^+ to B^+ ratio of $\varepsilon_{\text{kin}}/\varepsilon_{\text{dau}} \sim 1.1$.

In order to reduce the sensitivity to any remaining mismodelling in the simulation, the relative efficiencies are calculated using events generated with the same simulation version. The `Sim08` simulation is used to determine the efficiency of fully reconstructed $B_c^+ \to D_{(s)}^+ \overline{D}^0$ decays relative to $B^+ \to D_{(s)}^+ \overline{D}^0$ decays and the `Sim09` simulation is used to determine the efficiency of partially reconstructed $B_c^+ \to D_{(s)}^{(*)+} \overline{D}^{(*)0}$ decays relative to fully reconstructed $B_c^+ \to D_{(s)}^+ \overline{D}^0$ decays. The efficiencies of the B_c^+ decays relative to B^+ decays are shown in table 5.1. The uncertainties are due to the finite size of the simulation samples. Although the selection was optimised using fully reconstructed B_c^+ decays, the partially reconstructed B_c^+ decays have large efficiencies $75 - 90\%$ the size of those of the fully reconstructed B_c^+ decays.

Table 5.1: Ratio of total efficiencies of B_c^+ decays relative to the corresponding fully reconstructed B^+ decays, $\varepsilon(B_c^+)/\varepsilon(B^+)$. The quoted uncertainties are statistical only.

	Reconstructed state			
	$D_s^+ \overline{D}^0$ with $\overline{D}^0 \to$		$D^+ \overline{D}^0$ with $\overline{D}^0 \to$	
Decay channel	$K^+\pi^-$	$K^+\pi^-\pi^+\pi^-$	$K^+\pi^-$	$K^+\pi^-\pi^+\pi^-$
$B_c^+ \to D_{(s)}^+ \overline{D}^0$	0.420 ± 0.005	0.373 ± 0.009	0.441 ± 0.007	0.398 ± 0.010
$B_c^+ \to D_{(s)}^{*+} \overline{D}^0, D_{(s)}^+ \overline{D}^{*0}$	0.372 ± 0.006	0.317 ± 0.010	0.381 ± 0.008	0.337 ± 0.011
$B_c^+ \to D_{(s)}^{*+} \overline{D}^{*0}$	0.339 ± 0.006	0.278 ± 0.009	0.342 ± 0.007	0.297 ± 0.010

5.4 Mass fits

For a general fit PDF composed of a signal component S of fraction f and a background component B with fraction $1 - f$,

$$\int_{m_{\min}}^{m_{\max}} fS(m|p) + (1-f)B(m|p)dm = 1, \tag{5.14}$$

where m is the observable fitted in range $m_{\min} - m_{\max}$ and p are the floating parameters. For an unbinned fit, the likelihood function is given by,

$$\mathcal{L}(p) = \prod_{i}^{N_{\rm data}} fS(m_i|p) + (1-f)B(m_i|p), \tag{5.15}$$

where i is the entry in the data set. The values of the floating parameters that maximise the likelihood function are those that are the most likely to have produced the data. The errors returned by the fit are computed from the second derivatives of the likelihood function.

It is computationally more efficient to sum than to multiply and to minimise than to maximise, so in reality, the negative logarithm of the likelihood is minimised,

$$-\log(\mathcal{L}(p)) = -\sum_{i}^{N_{\rm data}} \log(fS(m_i|p) + (1-f)B(m_i|p)). \tag{5.16}$$

As yields and branching fractions are the parameters of interest, it is more convenient to work in the "extended" likelihood formalism. Instead of being normalised to unity, the fit PDF is normalised to the number of signal and background events,

$$\int_{m_{\min}}^{m_{\max}} N_S S(m|p) + N_B B(m|p) dm = N_S + N_B, \tag{5.17}$$

where $N_{\rm exp} = N_S + N_B$ is the expected number of events. The negative log likelihood includes an additional Poisson term that describes the probability of observing the number of events in the data set,

$$-\log(\mathcal{L}(p)) = -\sum_{i}^{N_{\rm data}} \log(N_S S(m_i|p) + N_B B(m_i|p)) - \log(\mathrm{Poisson}(N_{\rm obs}|N_{\rm exp})), \tag{5.18}$$

where $N_{\rm obs}$ is the number of observed events. The Poisson term has the usual definition,

$$\mathrm{Poisson}(N_{\rm obs}|N_{\rm exp}) = \frac{N_{\rm obs}^{N_{\rm exp}}}{N_{\rm obs}!} \exp(-N_{\rm exp}). \tag{5.19}$$

Unbinned, extended, maximum likelihood fits are performed to the constrained mass distribution of $B_{(c)}^+ \to D_{(s)}^{(*)+} D^{(*)}$ candidates using the RooFit package [172] within Root.

The minimisation algorithm used in RooFit is Minuit [173]. The fit PDF is

$$\begin{aligned}\mathrm{PDF_{fit}}(m|p) = N(B^+\!\to D^+_{(s)}D)\mathrm{DCB}(m|\sigma,\mu) + N(B_c^+\!\to D^+_{(s)}D)\mathrm{DCB}(m)\\ + N(B_c^+ \to D^{*+}_{(s)}D, D^+_{(s)}D^*)f(m) + N(B_c^+ \to D^{*+}_{(s)}D^*)f(m)\\ + N(B^+ \to K^+K^-\pi^+\overline{D}{}^0)\mathrm{Gaus}(m|\mu) + N_{\mathrm{exp}}\exp(mx) + N_{\mathrm{const}},\end{aligned} \quad (5.20)$$

where the floating parameters are: the yields N; the width σ of the $B^+\!\to D^+_{(s)}D$ model; the mean μ of the B^+ models; and the exponent x. The $B^+\!\to K^+K^-\pi^+\overline{D}{}^0$ component is only included in the fit to $B^+_{(c)}\!\to D_s^{(*)+}\overline{D}{}^{(*)0}$ candidates. As the $B^+\!\to D^+_{(s)}D^0$ yield is due to cross-feed from $B^+\!\to D^+_{(s)}\overline{D}{}^0$ decays (see section 4.6), the $B^+\!\to D^+_{(s)}D^0$ shape parameters are fixed from the fit to $B^+\!\to D^+_{(s)}\overline{D}{}^0$.

The two $\overline{D}{}^0$ channels are fitted simultaneously, resulting in four independent fits to the eight constrained mass distributions. In these fits, the background parameters and B^+ yields are free to vary independently in each $\overline{D}{}^0$ channel; however, only the sum of the B_c^+ yields, $N^{\mathrm{tot}}_{B_c^+} = N^{K\pi}_{B_c^+} + N^{K\pi\pi\pi}_{B_c^+}$, is allowed to float. The yield in each channel is determined by taking into account the relative efficiencies and B^+ yields,

$$N^{K\pi}_{B_c^+} = \frac{N^{K\pi}_{B^+}\varepsilon^{K\pi}_{B_c^+}/\varepsilon^{K\pi}_{B^+}}{N^{K\pi}_{B^+}\varepsilon^{K\pi}_{B_c^+}/\varepsilon^{K\pi}_{B^+} + N^{K\pi\pi\pi}_{B^+}\varepsilon^{K\pi\pi\pi}_{B_c^+}/\varepsilon^{K\pi\pi\pi}_{B^+}} N^{\mathrm{tot}}_{B_c^+}, \quad (5.21)$$

$$N^{K\pi\pi\pi}_{B_c^+} = \frac{N^{K\pi\pi\pi}_{B^+}\varepsilon^{K\pi\pi\pi}_{B_c^+}/\varepsilon^{K\pi\pi\pi}_{B^+}}{N^{K\pi}_{B^+}\varepsilon^{K\pi}_{B_c^+}/\varepsilon^{K\pi}_{B^+} + N^{K\pi\pi\pi}_{B^+}\varepsilon^{K\pi\pi\pi}_{B_c^+}/\varepsilon^{K\pi\pi\pi}_{B^+}} N^{\mathrm{tot}}_{B_c^+}. \quad (5.22)$$

For a simultaneous fit, the individual likelihood functions are combined by taking the product,

$$\mathcal{L}(p^{K\pi}, p^{K\pi\pi\pi}) = \mathcal{L}^{K\pi}(p^{K\pi})\mathcal{L}^{K\pi\pi\pi}(p^{K\pi\pi\pi}). \quad (5.23)$$

The negative log likelihood of the simultaneous fit is therefore given by

$$-\log(\mathcal{L}(p^{K\pi}, p^{K\pi\pi\pi})) = -\log(\mathcal{L}^{K\pi}(p^{K\pi})) - \log(\mathcal{L}^{K\pi\pi\pi}(p^{K\pi\pi\pi})). \quad (5.24)$$

The fit is performed in the range $5230 < m(D^+_{(s)}D) < 6700\,\mathrm{MeV}/c^2$. The lower limit is chosen to be above any contributions from partially reconstructed B^+ or B^0 decays. The upper limit is determined by the stripping requirement.

Before the fits were performed on the unblinded data, the analysis procedure was finalised and internally reviewed using fake data (see section 5.5). The results of the

Table 5.2: Signal yields from the fits of $B \to D_{(s)}^+ D$ decays. Samples with $\overline{D}^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ are fitted simultaneously. The uncertainties are statistical only.

	Reconstructed state			
Decay channel	$D_s^+\overline{D}^0$	$D_s^+ D^0$	$D^+\overline{D}^0$	$D^+ D^0$
$B^+ \to D_{(s)}^+ D$	$33\,734 \pm 187$	476 ± 27	1866 ± 46	37 ± 11
$B_c^+ \to D_{(s)}^+ D$	5 ± 5	-4 ± 3	6 ± 6	2 ± 4
$B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$	-1 ± 14	-4 ± 10	1 ± 13	-10 ± 9
$B_c^+ \to D_{(s)}^{*+} D^*$	34 ± 28	73 ± 19	68 ± 23	-8 ± 14

unblinded fits are shown in figure 5.7 and the corresponding signal yields are listed in table 5.2. The significance of the B_c^+ signals is calculated using Wilks' theorem [174],

$$\text{significance} = \sqrt{2 \ln \left(\frac{\mathcal{L}_{\text{alt}}}{\mathcal{L}_{\text{null}}} \right)} \tag{5.25}$$

where $\mathcal{L}_{\text{alt}}$ is the maximum likelihood value from the fit with the signal yield floating and $\mathcal{L}_{\text{null}}$ is the value from the fit with the signal yield fixed to zero. After the inclusion of systematic uncertainties (see section 5.6), none of the B_c^+ signals are above 2σ in significance. For some of the B_c^+ yields, a small negative value is fitted, which appears as a slight dip in the fit PDF. This is expected when searching for very small signals and is due to a downward fluctuation in the background. As the y axis is logarithmic, negative components are not displayed.

The total B^+ yield is the sum of the yields fitted in each $\overline{D}^0$ channel with their statistical uncertainties added in quadrature. As expected, large yields are observed in the normalisation channel, $B^+ \to D_{(s)}^+\overline{D}^0$. The yield of $B^+ \to D_s^+\overline{D}^0$ is greater than that of $B^+ \to D^+\overline{D}^0$ due to the factor ~ 25 larger branching fraction [3].

5.5 Fit validation

It is important to verify that the B_c^+ yields returned in the fit correspond to the true yields in data. However, as the true yields are unknown, the fit is validated using "toys" (fake data) of known distribution and yield. As small B_c^+ signals are expected, the background-only hypothesis is chosen. If the distribution of the yields returned by the

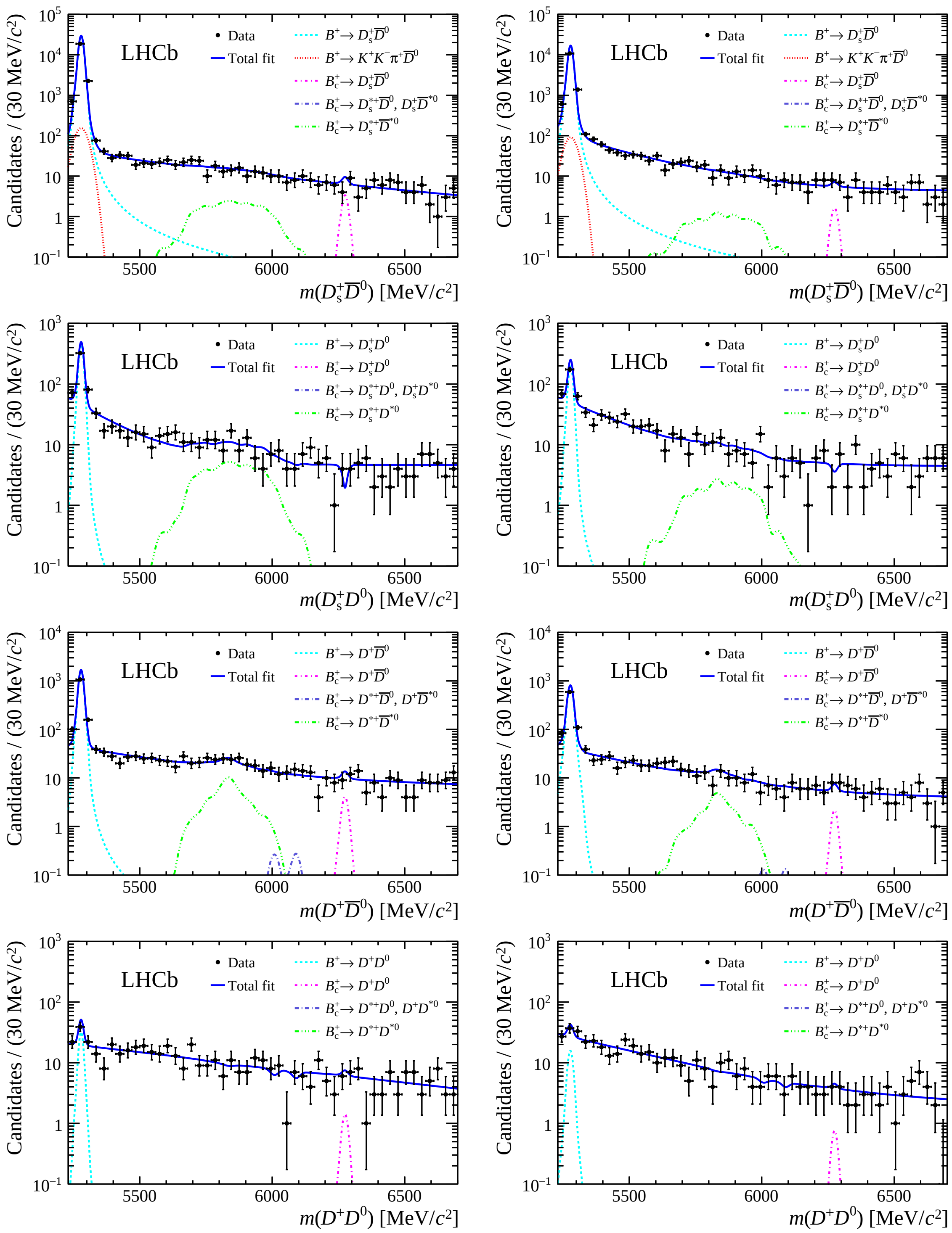


Figure 5.7: Fits to the constrained mass distribution of $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ candidates. Left plots are for $\overline{D}^0 \to K^+\pi^-$ decays, right for $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$.

fits to the toys are not consistent with zero, this indicates a bias in the fit. The procedure to generate the toys is:

1. The fit is performed on data in the region $5230 < m(D_{(s)}^+ D) < 6235\,\mathrm{MeV}/c^2$, where the upper limit is below any expected contributions from fully reconstructed $B_c^+ \to D_{(s)}^+ D$ decays. In this fit, the B_c^+ signal yields are fixed to zero.

2. The number of background events expected in the signal window, $6235 < m(D_{(s)}^+ D) < 6315\,\mathrm{MeV}/c^2$, is calculated by extrapolation of the combinatorial background model (the sum of an exponential function and a constant).

3. Background-only toys are generated according to the combinatorial background model. The number of generated events is drawn from a Poisson distribution with mean equal to the number of expected background events. Events in the signal window are then replaced by these background-only toys.

4. The fit is performed in the full fit region $5230 < m(D_{(s)}^+ D) < 6700\,\mathrm{MeV}/c^2$ with the B_c^+ yields fixed to zero. As the fit is performed with the additional knowledge of the data above the signal window, the description of the combinatorial background model in the signal window is expected to improve. Steps 2 and 3 are then repeated.

5. The fit is performed in the full fit region with the B_c^+ yields floating.

The generation of the toys and fit over the full region are repeated 1,000 times for each channel. Toys are also separately generated for the partially reconstructed B_c^+ decays. The signal window is defined as $5900 < m(D_{(s)}^+ D) < 6235\,\mathrm{MeV}/c^2$ for B_c^+ decays with one excited charm meson in the final state, $5550 < m(D_{(s)}^+ D) < 6150\,\mathrm{MeV}/c^2$ for $B_c^+ \to D_s^{*+} D^*$ decays and $5650 < m(D_{(s)}^+ D) < 6050\,\mathrm{MeV}/c^2$ for $B_c^+ \to D^{*+} D^*$ decays. The wider window for $B_c^+ \to D_s^{*+} D^*$ decays is due to the wider shape of the constrained mass distribution, as shown in figure 5.4.

Figures 5.8 and 5.9 show the results of the toy studies for fully reconstructed $B_c^+ \to D_{(s)}^+ D$ decays. The "pulls" are defined as the B_c^+ yields normalised by their uncertainties. If the mean and width of the pulls distribution are zero and one, respectively, then the fit is unbiased and has the correct error coverage. As the means have small non-zero values, a slight bias is observed. This has been subtracted from the yields reported in table 5.2. In general, the widths of the pulls distributions are compatible with unity within 2σ. As the fit biases are smaller than the uncertainties on the yields, no significant signals are observed.

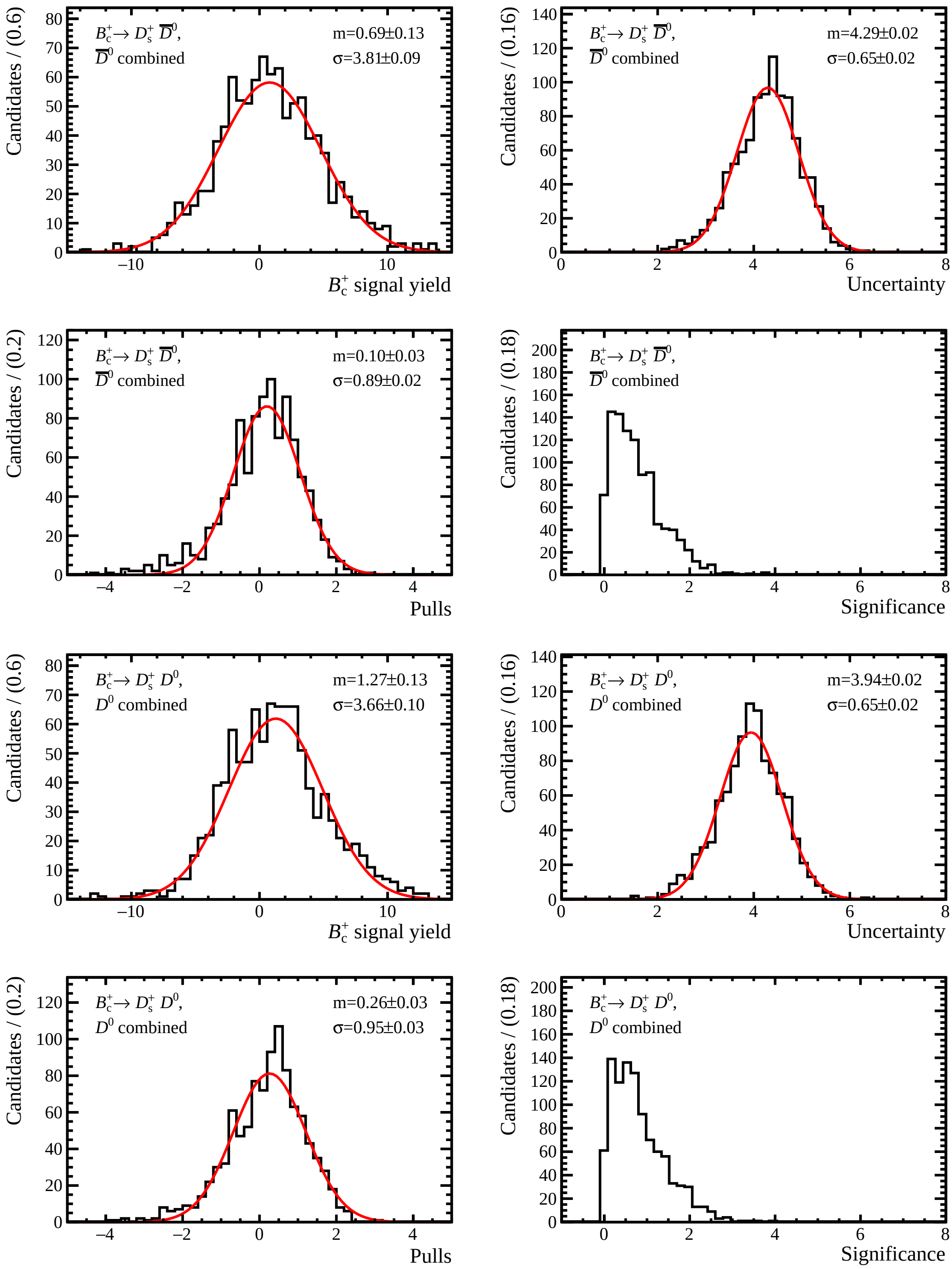


Figure 5.8: Distribution of the $B_c^+ \to D_s^+ D$ signal yields, uncertainties, pulls and significance from fitting to toys generated under the background-only hypothesis. Fits with a Gaussian function are overlaid.

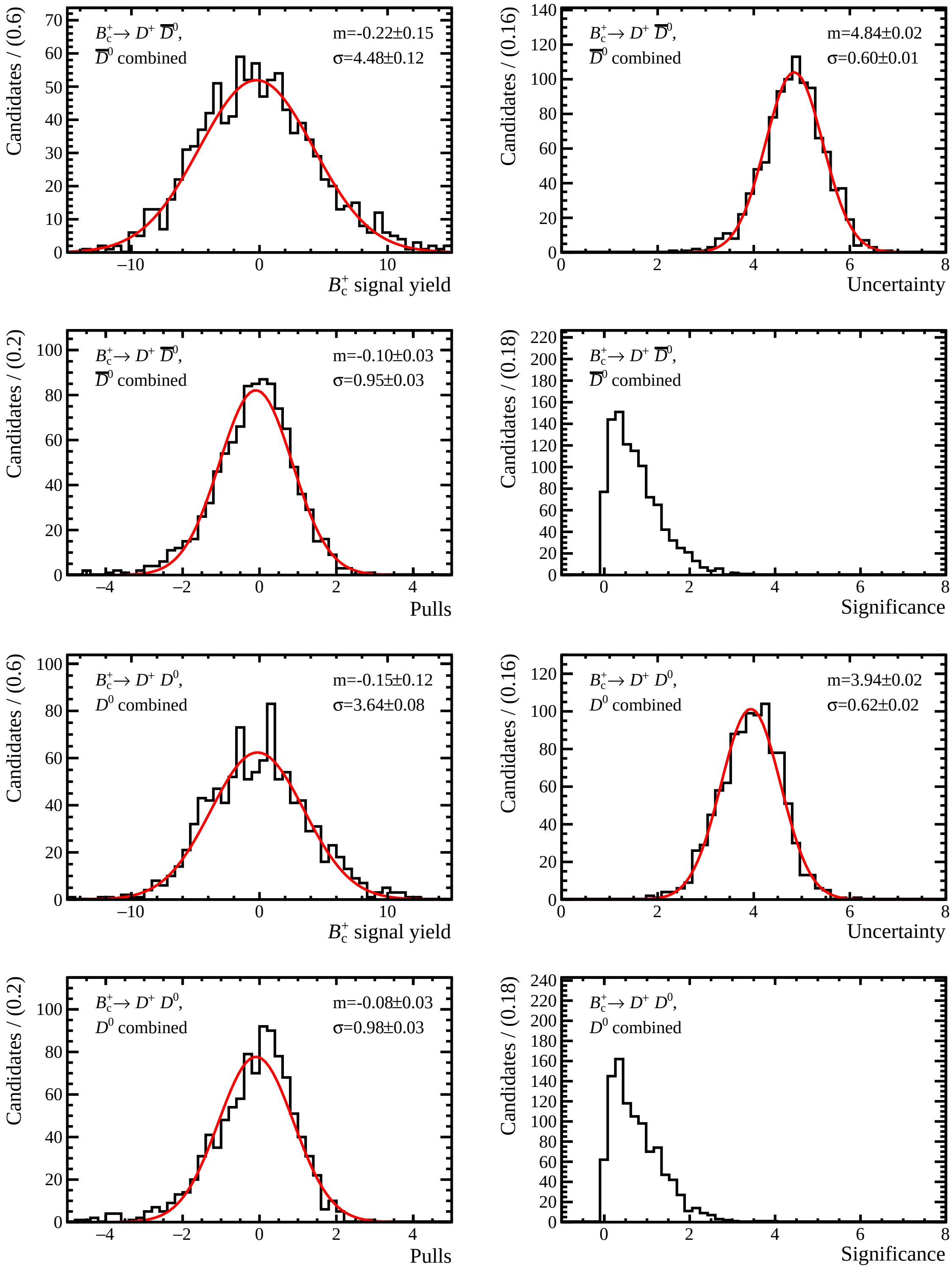


Figure 5.9: Distribution of the $B_c^+ \to D^+ D$ signal yields, uncertainties, pulls and significance from fitting to toys generated under the background-only hypothesis. Fits with a Gaussian function are overlaid.

5.6 Systematic uncertainties

The systematic uncertainties can be broadly split into two categories: those on the B_c^+ yields and those on the normalisation (the B^+ yields and relative efficiencies). The systematic uncertainties on the yields are expressed in absolute units of events. Since the normalisation is well defined, *i.e.* large and positive, the systematic uncertainties are expressed as percentages.

Many of the systematic uncertainties are determined by varying the fit model. The corresponding systematic uncertainty is given by

$$\sigma_{\text{sys}} = \sqrt{(N_{\text{def}} - N_{\text{var}})^2 + |\sigma^2_{\text{def}} - \sigma^2_{\text{var}}|}, \tag{5.26}$$

which takes into account both the change in yield N and uncertainty σ.

5.6.1 Systematic uncertainties on the B_c^+ yields

The systematic uncertainties on the B_c^+ yields are summarised in table 5.3. This section describes each contribution in turn.

Signal shape

The shapes of the fully reconstructed $B_c^+ \to D_{(s)}^+ D$ signals are fixed from a fit to simulation. However, this neglects the statistical uncertainty on the shape parameters due to the finite size of the simulation samples. The fit is repeated 1,000 times varying the shape parameters according to Gaussian distributions that take into account the correlations between them.

Signal model

The constrained mass distribution of $B_c^+ \to D_{(s)}^+ D$ candidates is described by the sum of two Crystal Ball functions. As this is an empirical choice, there is an associated systematic uncertainty to account for the limited knowledge of the true model. The fit is repeated using the sum of two Gaussian functions as the B_c^+ signal model. This model is chosen since it is a simplified version of the DCB, but is still able to provide a fairly good description of the data.

Table 5.3: Systematic uncertainties on the B_c^+ yields, for the combined fit to both the $\overline{D}{}^0 \to K^+\pi^-$ and the $\overline{D}{}^0 \to K^+\pi^-\pi^+\pi^-$ decay channels. The total systematic uncertainty is calculated as the quadratic sum of the individual components.

	Reconstructed state			
Source	$D_s^+\overline{D}{}^0$	$D_s^+D^0$	$D^+\overline{D}{}^0$	D^+D^0
$B_c^+ \to D_{(s)}^+ D$				
Signal shape	0.25	0.28	0.31	0.13
Signal model	0.40	0.34	0.61	0.44
B_c^+ mass	0.64	0.62	0.79	0.51
Background model	1.12	1.75	1.88	0.56
Fit bias	0.70	1.28	0.27	0.19
Total	1.54	2.30	2.17	0.91
$B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$				
Signal composition	7.6	5.5	7.1	5.7
Background model	11.9	17.5	16.4	4.5
Fit bias	5.5	9.4	3.9	1.3
Total	15.2	20.6	18.3	7.4
$B_c^+ \to D_{(s)}^{*+} D^*$				
Polarisation	23	14	9	5
Background model	43	98	37	9
Fit bias	10	7	8	1
Total	49	99	39	10

B_c^+ mass

The mean of the $B_c^+ \to D_{(s)}^+ D$ model is fixed to the world-average value $m_{B_c^+} = 6274.9 \pm 0.8\,\text{MeV}/c^2$ [3]. However, this neglects the uncertainty. The position of the mean in the data also has an uncertainty due to the momentum scale calibration (see section 4.5.1) of 0.03% [175]. Applying this to the Q-value of the decay gives $0.03\% \times (m_{B_c^+} - m_{D_{(s)}^+} - m_{\overline{D}{}^0}) \sim 0.8\,\text{MeV}/c^2$. Adding the two uncertainties in quadrature results in a total uncertainty of $1.1\,\text{MeV}/c^2$. The fit is repeated twice, each time shifting the position of the B_c^+ mean up and down within its uncertainty. The larger variation of the two is taken as the systematic uncertainty.

Background model

The model chosen to describe the combinatorial background is the sum of an exponential with a constant. To evaluate the uncertainty associated with this choice, the fit is repeated using an exponential function as the background model. This is the dominant systematic uncertainty for both the fully and partially reconstructed B_c^+ modes. The partially reconstructed B_c^+ modes are particularly sensitive to the choice of background model due to their wide constrained mass distributions.

Fit bias

The fits to the toys show a small but non-zero value for the mean of the B_c^+ yields, indicating a fit bias (see figures 5.8 and 5.9). The bias is subtracted from the yields in the final fit. The associated systematic uncertainty is taken as the quadratic sum of the value of the fit bias and its uncertainty.

Signal composition

For the constrained mass model of B_c^+ decays with one excited charm meson in the final state, it is assumed that $\mathcal{B}(B_c^+ \to D_{(s)}^{*+} D) = \mathcal{B}(B_c^+ \to D_{(s)}^+ D^*)$. However, since these decays have not been observed, their relative branching fractions are unknown. The model is varied using the shape of each individual decay in turn, the distributions of which can be seen in figure 5.3. The larger of the two variations is taken as the systematic uncertainty.

Polarisation

For the constrained mass model of the decay of B_c^+ with two excited charm mesons in the final state, it is assumed that there is a $2:1$ ratio of the transverse : longitundinal polarisations. However, this is also unmeasured. The model is varied with each individual polarisation in turn, the distributions of which are shown in figure 5.4. The larger of the two variations is taken as the systematic uncertainty.

Table 5.4: Systematic uncertainties, in %, on the normalisation of the B_c^+ branching fraction determination. The total systematic uncertainty is calculated as the quadratic sum of the common and channel-specific components.

		Reconstructed state			
		$D_s^+ D$, with $D^0 \to$		$D^+ D$, with $D^0 \to$	
Channel	Source	$K^-\pi^+$	$K^-\pi^+\pi^-\pi^+$	$K^-\pi^+$	$K^-\pi^+\pi^-\pi^+$
Common	B^+ stat.	0.7	0.9	3.1	4.3
	B^+ signal shape	0.0	0.0	0.0	0.3
	B^+ signal model	0.1	0.2	0.1	0.3
	Background model	0.0	0.6	1.6	1.3
	$B^+ \to K^+ K^- \pi^+ \overline{D}^0$	1.4	1.4	—	—
	B_c^+ lifetime	1.5	1.5	1.5	1.5
	PID	2.4	0.9	1.2	3.2
	D^0 model	—	1.1	—	0.7
$B_c^+ \to D_{(s)}^+ D$	Simulation stat.	1.2	2.4	1.6	2.5
	Total	3.5	3.6	4.3	6.3
$B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*$	Simulation stat.	1.7	3.3	2.0	3.3
	Signal composition	1.0	0.8	0.7	2.6
	Total	3.8	4.3	4.5	7.1
$B_c^+ \to D_{(s)}^{*+} D^*$	Simulation stat.	1.7	3.4	2.0	3.3
	Polarisation	1.5	0.4	1.4	1.3
	$\mathcal{B}\ (D^{*+} \to D^+ \pi^0/\gamma)$	—	—	1.5	1.5
	Total	3.9	4.4	4.9	6.9

5.6.2 Systematic uncertainties on the normalisation

The systematic uncertainties on the normalisation are summarised in table 5.4. This section describes each contribution in turn.

B^+ statistics

The B^+ yields have associated statistical uncertainties (see table 5.2). Since the B^+ yields are used in the normalisation, these are treated as systematic uncertainties. As large yields are observed for $B^+ \to D_s^+ \overline{D}^0$ decays, this is a small effect. However, for Cabibbo suppressed $B^+ \to D^+ \overline{D}^0$ decays, this is the dominant systematic uncertainty.

B^+ signal shape

The tail parameters and ratio of widths of the DCB model used for $B^+ \to D_{(s)}^+ \overline{D}^0$ decays are fixed from simulation, neglecting their statistical uncertainties. To evaluate the systematic uncertainty, the same procedure as for the $B_c^+ \to D_{(s)}^+ D$ models is used. The fit is repeated 1,000 times, varying the fixed B^+ shape parameters according to their covariance matrix. The resulting systematic uncertainty is smaller than in the B_c^+ case for two reasons. First, fewer parameters are fixed from the fit to simulation and second, the B^+ simulation samples are larger due to the larger efficiencies, resulting in smaller variations.

B^+ signal model

The choice of the DCB function as the $B^+ \to D_{(s)}^+ \overline{D}^0$ constrained mass model has an associated systematic uncertainty. To evaluate this, the model is varied with the sum of two Gaussian functions, where the ratio of widths is fixed from a fit to simulation. This leaves the same number of parameters floating as for the nominal fit.

Background model

To evaluate the uncertainty in the choice of background model, the fit is repeated using an exponential function. The $B^+ \to D^+ \overline{D}^0$ decays are more sensitive than the $B^+ \to D_s^+ \overline{D}^0$ decays to the choice of background model due to the smaller yields.

$B^+ \to K^+ K^- \pi^+ \overline{D}^0$

The fit to the constrained mass distributions of $B^+ \to D_s^+ \overline{D}^0$ candidates includes a component for the single charm background $B^+ \to K^+ K^- \pi^+ \overline{D}^0$. The $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ yield is determined using data from the D_s^+ sidebands and its width from simulation. However, there may be non-linearities in the charm mass distribution of $B^+ \to K^+ K^- \pi^+ \overline{D}^0$ decays or discrepancies between data and simulation. A systematic uncertainty 100% the size of the yield is therefore assigned.

B_c^+ lifetime

The lifetime of the B_c^+ in the `Sim08` simulation is corrected to the world-average value $\tau_{\mathrm{WA}}(B_c^+) = (507 \pm 9)\,\mathrm{fs}$ (see section 4.5.2). The uncertainty is taken into account by reweighting the simulation such that the lifetime is shifted up and down by one standard deviation. The corresponding change in the relative efficiency is between $1.3 - 1.6\%$, depending on the channel under consideration. As the lifetime is a property of the B_c^+ and not expected to be dependent on the final state, an uncertainty of 1.5% is applied uniformly.

Particle identification

As the final state is the same between the signal and normalisation channel, inaccuracies in the modelling of the $\mathrm{DLL}_{K\pi}$ distributions are expected to cancel to first order. However, differences in the momentum spectra of particles produced in B^+ and B_c^+ decays may lead to small differences in the PID efficiencies. A conservative estimate of this uncertainty is taken by evaluating the relative efficiency with and without any corrections to the $\mathrm{DLL}_{K\pi}$ distributions (see section 4.5.2).

D^0 model

The decay of $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ is generated according to the kinematics of the decay, however, there is resonant structure expected from $K^*(892)^0 \to K^+\pi^-$, $\rho(770)^0 \to \pi^+\pi^-$, $a_1(1260)^+ \to \pi^+\pi^-\pi^+$ and $K_1(1270)^+ \to K^+\pi^-\pi^+$. As the same charm decay modes are used for the signal and normalisation, this effect is expected to cancel to first order. To evaluate the systematic uncertainty, each resonance is replicated in turn by weighting the simulation with the data/simulation ratio of the invariant masses of the final states, *i.e.* $m(K^+\pi^-)$, $m(\pi^+\pi^-)$, $m(\pi^+\pi^-\pi^+)$ and $m(K^+\pi^-\pi^+)$, as can be seen in figure A.6. The uncertainty is taken as the quadratic sum of the differences in relative efficiency. There is no resonant structure in two-body decays, so this systematic uncertainty is only relevant for channels with $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ decays.

Simulation statistics

The relative efficiencies are determined from simulation samples of a finite size. Therefore, they have an associated statistical uncertainty, as shown in table 5.1. As the relative efficiency forms part of the normalisation, the statistical uncertainties are treated as systematic uncertainties. This is one of the dominant systematic uncertainties in the partially reconstructed B_c^+ modes due to the lower reconstruction efficiencies.

Signal composition

The relative efficiency of the decays of B_c^+ with one excited charm meson in the final state is the average efficiency of $B_c^+ \to D_{(s)}^{*+} D$ and $B_c^+ \to D_{(s)}^{+} D^*$ decays. There may be differences between the two, arising from the different branching fractions of the $D^* \to D\pi^0/\gamma$ decays. The relative efficiency is re-evaluated using $B_c^+ \to D_{(s)}^{*+} D$ and $B_c^+ \to D_{(s)}^{+} D^*$ decays individually, with the largest difference taken as the systematic uncertainty.

Polarisation

The calculation of the relative efficiency of B_c^+ decays to two excited charm mesons assumes a 2 : 1 ratio of the transverse : longitundinal polarisations. The relative efficiency is re-evaluated with each of the individual polarisations in turn, with the largest difference taken as the systematic uncertainty.

$\mathcal{B}(D^{*+} \to D^+\pi^0/\gamma)$

As shown in equation 5.3, the branching fraction of decays of B_c^+ to two excited charm mesons are corrected for the D^* branching fractions. For modes with D_s^{*+} or $\overline{D}^{*0}$, the branching fractions to final states containing D_s^+ and $\overline{D}^0$, respectively, and a neutral particle are 100%. However, for D^{*+}, there are additional decays to $D^0\pi^+$ with a branching fraction of $(67.7 \pm 0.5)\%$ [3]. The uncertainty on this branching fraction is taken into account as a systematic uncertainty.

5.7 Branching fractions

The relationship between the B_c^+ branching fractions and yields are described by equations 5.1, 5.2 and 5.3. The fit is repeated with the left-hand side of these equations as the floating parameters by multiplying the B_c^+ yields with the B^+ yields and relative efficiencies.

The systematic uncertainties are included in the fit as Gaussian constraints. For the systematic uncertainties on the B_c^+ yields, a Gaussian constrained factor with mean zero and width equal to the total systematic uncertainty is added to each of the yields. For the systematic uncertainties on the normalisation, a Gaussian constraint of width equal to the total systematic uncertainty is applied directly on the normalisation factors (the product of the relative efficiencies and B^+ yields) entering equations 5.1, 5.2 and 5.3. When the constraints are included in the fit, the likelihood function becomes

$$\mathcal{L}_c(p) = \mathcal{L}(p) \prod_i^{N_{\text{constraints}}} \text{Gaus}(p_i), \tag{5.27}$$

where $\text{Gaus}(p_i)$, is the Gaussian constraint for parameter p_i. The negative log likelihood is therefore given by

$$-\log(\mathcal{L}_c(p)) = -\log(\mathcal{L}(p)) - \sum_i^{N_{\text{constraints}}} \log(\text{Gaus}(p_i)). \tag{5.28}$$

The inclusion of the systematic uncertainties in the fit increases the uncertainty on the B_c^+ branching fractions.

The results for the fully reconstructed B_c^+ branching fractions are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0)}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (3.0 \pm 3.7) \times 10^{-4}, \tag{5.29}$$
$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^+ D^0)}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (-3.8 \pm 2.6) \times 10^{-4}, \tag{5.30}$$
$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^+ \overline{D}^0)}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (8.0 \pm 7.5) \times 10^{-3}, \tag{5.31}$$
$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^+ D^0)}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (2.9 \pm 5.3) \times 10^{-3}. \tag{5.32}$$

The results for the branching fractions of B_c^+ decays with one excited charm meson in the final state are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+}\overline{D}^0)+\mathcal{B}(B_c^+\to D_s^+\overline{D}^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} = (-0.1\pm 1.5)\times 10^{-3}, \tag{5.33}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+}D^0)+\mathcal{B}(B_c^+\to D_s^+D^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} = (-0.3\pm 1.9)\times 10^{-3}, \tag{5.34}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to (D^{*+}\to D^+\pi^0/\gamma)\overline{D}^0)+\mathcal{B}(B_c^+\to D^+\overline{D}^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} = (0.2\pm 3.2)\times 10^{-2}, \tag{5.35}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to (D^{*+}\to D^+\pi^0/\gamma)D^0)+\mathcal{B}(B_c^+\to D^+D^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} = (-1.5\pm 1.7)\times 10^{-2}. \tag{5.36}$$

The results for the branching fractions of B_c^+ decays with two excited charm mesons in the final state are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+}\overline{D}^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} = (3.2\pm 4.3)\times 10^{-3}, \tag{5.37}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+}D^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} = (7.0\pm 9.2)\times 10^{-3}, \tag{5.38}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D^{*+}\overline{D}^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} = (3.4\pm 2.3)\times 10^{-1}, \tag{5.39}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D^{*+}D^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} = (-4.1\pm 9.1)\times 10^{-2}. \tag{5.40}$$

The fitted values of all parameters allowed to float can be seen in appendix C.1. As there is no evidence for signal above 2σ in significance, upper limits are placed on the branching fractions.

5.7.1 Limits

The upper limit is the value of the B_c^+ branching fraction such that there is a given small probability α to observe as many, or fewer, events. The values of α chosen are 0.10 and 0.05, corresponding to upper limits at 90 and 95% confidence level, respectively. The upper limits are calculated using the asymptotic CL_s method [176], implemented within the RooStats package [177] of Root.

Asymptotic CL$_s$ method

The CL$_s$ method is a *frequentist* method for setting upper limits, commonly used in experimental particle physics. A frequentist confidence interval contains the true value of the parameter of interest with some probability $1-\alpha$, *i.e.* if the experiment were repeated an infinite number of times, $1-\alpha$ of the intervals would contain the true value.[1] An upper limit is the upper bound of a confidence interval with a lower bound of zero.

To test the compatibility of a given value of the parameter of interest μ with the data, the profile likelihood ratio is measured,

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \tag{5.41}$$

where θ are the "nuisance parameters" (parameters floating in the fit that aren't the parameters of interest, *e.g.* the width of the B^+ model). $\hat{\mu}$ and $\hat{\theta}$ indicate the values for which the likelihood function is maximised and $\hat{\hat{\theta}}$ are the values of the nuisance parameters that maximise the likelihood for a given value of μ. If $\lambda(\mu)$ is near 1, then the data and hypothesised value of μ are in good agreement. A more convenient measure of compatibility is the "test statistic", which for an upper limit is given by

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \text{for } \hat{\mu} \le \mu \\ 0 & \text{for } \hat{\mu} > \mu, \end{cases} \tag{5.42}$$

where large values of q_μ indicate poor agreement between the data and μ. Values of μ less than the one obtained in the fit are not considered less compatible with the data, so the test statistic is set to zero.

When setting limits, the "null" hypothesis is the model with signal and background. This is tested against the alternative hypothesis, which is the background-only model. The p-value of the signal plus background hypothesis is

$$p_{s+b} = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu|s+b)dq_\mu, \tag{5.43}$$

[1] This differs from a Bayesian credibility interval, which is defined as the interval in which the parameter of interest lies with some probability $1-\alpha$. In general, the frequentist approach can be thought of as making statements of likelihood on the data given the parameters, whereas the Bayesian approach considers the likelihood of the parameters given the data.

where $q_{\mu,\mathrm{obs}}$ is the observed test statistic and $f(q_\mu|s+b)$ is the PDF of the test statistic assuming the signal plus background hypothesis. p_{s+b} is the probability, under the signal plus background hypothesis, of measuring a value of the test statistic greater than or equal to the observed value. Similarly, the p-value of the background-only hypothesis is

$$p_b = \int_{-\infty}^{q_{\mu,\mathrm{obs}}} f(q_\mu|b)dq_\mu, \tag{5.44}$$

where $f(q_\mu|b)$ is the PDF of the test statistic assuming the background-only hypothesis. p_b is the probability, under the background-only hypothesis, of measuring a value of the test statistic less than or equal to the observed value. A hypothesis can be excluded if the observed p-value is below a specified threshold, α.

The "CL_{s+b}" method excludes the signal plus background model at $1-\alpha$ confidence level for the value of μ at which $\mathrm{CL}_{s+b} \equiv p_{s+b} = \alpha$. However, this method can be used to exclude hypotheses to which there is little to no sensitivity. This occurs when the expected number of signal events is much less than the expected number of background events and the distributions of $f(q_\mu|s+b)$ and $f(q_\mu|b)$ are almost overlapping.

To prevent this, the more stringent "CL_s" method is used. The CL_s method excludes the signal plus background hypothesis at $1-\alpha$ confidence level for the value of μ at which $\mathrm{CL}_s \equiv \mathrm{CL}_{s+b}/\mathrm{CL}_b = p_{s+b}/(1-p_b) = \alpha$. The p-value of the signal plus background hypothesis is effectively penalised by a factor that increases with decreasing sensitivity, resulting in a more conservative value for the upper limit. The CL_s method "overcovers" (has a coverage greater than the stated $1-\alpha$) by an amount that depends on the value of μ.

There are two methods to obtain the distributions for the test statistic under the signal plus background and background-only hypotheses. The first is using toy MC, which can be time consuming and computationally expensive. The second is using analytical expressions that are based on the assumption that $\mathcal{O}(1/\sqrt{N})$ terms, where N is the size of the data sample, can be neglected. With these analytical expressions, the results for the CL_s method can be obtained with the use of a single representative data set (the "Asimov" data set). The asymptotic CL_s method is based on the Asimov data set.

Results

The distribution of the CL_s p-values are shown in figure 5.10 for each B_c^+ branching fraction. The 90 (95)% upper limit is the value of the branching fraction at which the

p-value is equal to 0.10 (0.05). If there is an upward (downward) fluctuation of the background, the value of the limit is greater (less) than the expected limit. The upper limits on the fully reconstructed B_c^+ branching fractions at 90 (95%) confidence level are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^+\overline{D}^0)}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} < 0.9\,(1.1)\times 10^{-3}, \tag{5.45}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^+ D^0)}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} < 3.7\,(4.7)\times 10^{-4}, \tag{5.46}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D^+\overline{D}^0)}{\mathcal{B}(B^+\to D^+\overline{D}^0)} < 1.9\,(2.2)\times 10^{-2}, \tag{5.47}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D^+ D^0)}{\mathcal{B}(B^+\to D^+\overline{D}^0)} < 1.2\,(1.4)\times 10^{-2}. \tag{5.48}$$

The upper limits on the branching fractions of B_c^+ decays with one excited charm meson in the final state at 90 (95%) confidence level are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+}\overline{D}^0)+\mathcal{B}(B_c^+\to D_s^+\overline{D}^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} < 2.8\,(3.4)\times 10^{-3}, \tag{5.49}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+} D^0)+\mathcal{B}(B_c^+\to D_s^+ D^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} < 3.0\,(3.6)\times 10^{-3}, \tag{5.50}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to (D^{*+}\to D^+\pi^0/\gamma)\overline{D}^0)+\mathcal{B}(B_c^+\to D^+\overline{D}^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} < 5.5\,(6.6)\times 10^{-2}, \tag{5.51}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to (D^{*+}\to D^+\pi^0/\gamma)D^0)+\mathcal{B}(B_c^+\to D^+ D^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} < 2.2\,(2.8)\times 10^{-2}. \tag{5.52}$$

The upper limits on the branching fractions of B_c^+ decays with two excited charm mesons in the final state at 90 (95%) confidence level are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+}\overline{D}^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} < 1.1\,(1.3)\times 10^{-2}, \tag{5.53}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D_s^{*+} D^{*0})}{\mathcal{B}(B^+\to D_s^+\overline{D}^0)} < 2.0\,(2.4)\times 10^{-2}, \tag{5.54}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D^{*+}\overline{D}^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} < 6.5\,(7.3)\times 10^{-1}, \tag{5.55}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+\to D^{*+} D^{*0})}{\mathcal{B}(B^+\to D^+\overline{D}^0)} < 1.3\,(1.6)\times 10^{-1}. \tag{5.56}$$

Although the limits are approximately an order of magnitude larger in the $B_c^+\to D^{(*)+}D^{(*)}$ modes than the $B_c^+\to D_s^{(*)+}D^{(*)}$ modes, this is largely due to the difference in B^+ branching fractions [3]. When the limits are multiplied through by $\mathcal{B}(B^+)$, they are approximately equal, as expected from the similarities in the background levels and size of the systematic uncertainties. As the B_c^+ decays to two excited charm mesons are corrected for the D^* branching fractions, the limits for $B_c^+ \to D^{*+} D^*$ decays are a factor of three less sensitive than $B_c^+ \to D_s^{*+} D^*$ decays due to $\mathcal{B}(D^{*+} \to D^0 \pi^+) \sim 68\%$ [3].

5.7.2 Interpretation

Some theoretical predictions for the B_c^+ branching fractions are listed in table 2.3. The largest prediction for the fully reconstructed modes is [20]

$$\mathcal{B}(B_c^+ \to D^+ \overline{D}{}^0) = (3.2 \pm 0.7) \times 10^{-5}. \tag{5.57}$$

In order to compare this with the upper limit, input is required for f_c/f_u and $\mathcal{B}(B^+ \to D^+ \overline{D}{}^0)$. The branching fraction has been measured to be $\mathcal{B}(B^+ \to D^+ \overline{D}{}^0) = (3.8 \pm 0.4) \times 10^{-4}$ [3] and f_c/f_u is estimated to be between $0.2 - 1.1\%$ (see section 2.3.2.2). Assuming an optimistic value of $f_c/f_u = 1.1\%$, the limit becomes

$$\mathcal{B}(B_c^+ \to D^+ \overline{D}{}^0) < 6.6\,(7.6) \times 10^{-4}, \tag{5.58}$$

which is well above the theoretical expectation. With the addition of data collected in Run II and beyond, the expected B_c^+ yields will increase (see table 2.3). By the end of Run III in 2023, it is likely that $B_c^+ \to D^+ \overline{D}{}^0$ will be observed.

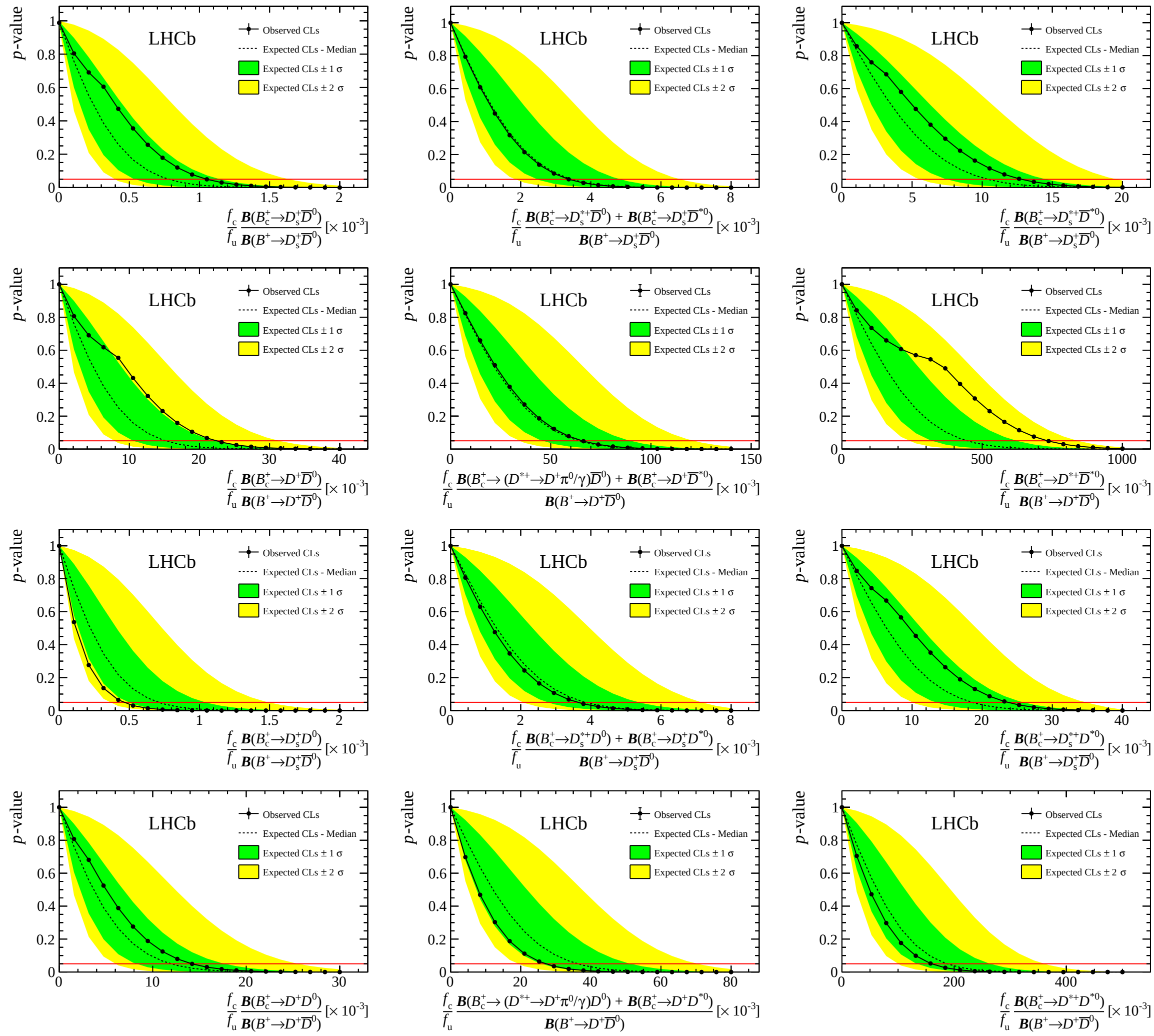


Figure 5.10: CL$_s$ p-value distributions for the measured values of $\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_{(s)}^{(*)+} D^{(*)})}{\mathcal{B}(B^+ \to D_{(s)}^+ \overline{D}^0)}$. The intersections of the horizontal red lines with the black solid curves correspond to the limits at 95% confidence level.

Chapter 6

Measurement of the *CP* asymmetry in $B^+ \to D^+_{(s)} \overline{D}^0$ decays

"Symmetry is overrated. Overrated is symmetry."
— Larry Wall

This chapter describes the measurement of the *CP* asymmetry in $B^+ \to D^+_{(s)} \overline{D}^0$ decays using Run I data collected by the LHCb experiment and has been published in reference [2]. Some predictions for $\mathcal{A}^{CP}(B^+ \to D^+_{(s)} \overline{D}^0)$ in the SM are listed in table 2.4. Large values of $\mathcal{O}(10\%)$ are predicted by some new physics models [18, 19]. First, an overview of the analysis strategy is given (section 6.1). This is followed by a description of the raw asymmetry (section 6.2), detection asymmetry (section 6.3), and production asymmetry (section 6.4) and lastly, the results for the *CP* asymmetry (section 6.5) are presented.

6.1 Analysis strategy

The *CP* asymmetry in $B^+ \to D^+_{(s)} \overline{D}^0$ decays is defined as

$$\mathcal{A}^{CP}(B^+ \to D^+_{(s)} \overline{D}^0) \equiv \frac{\Gamma(B^- \to D^-_{(s)} D^0) - \Gamma(B^+ \to D^+_{(s)} \overline{D}^0)}{\Gamma(B^- \to D^-_{(s)} D^0) + \Gamma(B^+ \to D^+_{(s)} \overline{D}^0)}. \tag{6.1}$$

where Γ is the decay rate. The determination of $\mathcal{A}^{CP}(B^+ \to D^+_{(s)} \overline{D}^0)$ is based on the raw asymmetry,

$$A_{\text{raw}} \equiv \frac{N(B^- \to D^-_{(s)} D^0) - N(B^+ \to D^+_{(s)} \overline{D}^0)}{N(B^- \to D^-_{(s)} D^0) + N(B^+ \to D^+_{(s)} \overline{D}^0)}, \tag{6.2}$$

where N stand for the signal yields. The raw asymmetry contains contributions from CP violation, as well as experimental effects such as the collision environment and the interaction of the final state particles with the detector. If the asymmetries are small such that terms containing products of asymmetries can be neglected, the CP asymmetry can be expressed as

$$\mathcal{A}^{CP} = A_{\text{raw}} - A_D - A_P, \tag{6.3}$$

where the detection asymmetry, A_D, is the asymmetry in the efficiencies, ε,

$$A_D \equiv \frac{\varepsilon(B^- \to D^-_{(s)} D^0) - \varepsilon(B^+ \to D^+_{(s)} \overline{D}^0)}{\varepsilon(B^- \to D^-_{(s)} D^0) + \varepsilon(B^+ \to D^+_{(s)} \overline{D}^0)}, \tag{6.4}$$

and the production asymmetry, A_P, is the asymmetry of the B^+ and B^- production cross-sections, σ,

$$A_P \equiv \frac{\sigma(B^-) - \sigma(B^+)}{\sigma(B^-) + \sigma(B^+)}. \tag{6.5}$$

6.2 Raw asymmetry

To determine the yields, fits are performed to the constrained mass distributions of $B^+ \to D^+_{(s)} \overline{D}^0$ candidates, separated by charge, in the range $5230 < m(D^+_{(s)} \overline{D}^0) < 5330\,\text{MeV}/c^2$. The fit model contains three components: the $B^+ \to D^+_{(s)} \overline{D}^0$ signal; the single charm $B^+ \to K^- \pi^+ \pi^+ \overline{D}^0$ background; and the combinatorial background. The parametrisation of the $B^+ \to D^+_{(s)} \overline{D}^0$ and $B^+ \to K^- \pi^+ \pi^+ \overline{D}^0$ models is the same as in the B_c^+ search (see sections 5.2.1 and 5.2.4), however, small differences are expected since the shape parameters are derived from samples with differing MVA requirements. For the $B^+ \to K^- \pi^+ \pi^+ \overline{D}^0$ model, the yield is determined separately for each charge, whereas the width is assumed to be charge independent. For the combinatorial background, an exponential function is used. This is a simpler model than the one used in the B_c^+ search (the sum of an exponential function and a constant) due to the reduced fit range.

Before unblinding, the $B^+ \to D^+ \overline{D}^0$ mode was separated by charge randomly based on the run number. Only after the analysis procedure was finalised and internally reviewed was the fit performed on the unblinded data. The results of the unblinded fits are shown in figure 6.1 with the corresponding yields and raw asymmetries listed in table 6.1. The fitted values of all parameters allowed to float can be seen in appendix C.2. Separate fits are performed for each of the $\overline{D}^0$ channels and the results are combined by summing the yields and adding the uncertainties in quadrature. The yields are larger for the $B^+ \to D_s^+ \overline{D}^0$ mode than for the $B^+ \to D^+ \overline{D}^0$ mode due to the factor ~ 25 larger branching fraction [3], resulting in smaller statistical uncertainties. The raw asymmetry in $B^+ \to D_s^+ \overline{D}^0$ decays is significantly non-zero, whereas the raw asymmetry in $B^+ \to D^+ \overline{D}^0$ is compatible with zero.

Although the choice of fit model may result in small biases in the yields, this is expected to affect both charges in the same way and hence cancel to first order in the measurement of the raw asymmetry. To check for any residual bias, the fit is repeated with several variations: the sum of two Gaussian functions as the B^+ signal model; a single Gaussian function as the B^+ signal model; a linear function as the combinatorial background model; and a uniform function as the combinatorial background model. For the high-statistics $B^+ \to D_s^+ \overline{D}^0$ mode, the effect on the raw asymmetry is below 0.1%. Slightly larger variations of $0.1 - 0.2\%$ are observed in the $B^+ \to D^+ \overline{D}^0$ mode, however, this is expected due to the larger statistical uncertainties. Overall, no bias due to the fit model is observed.

Table 6.1: Yields and raw asymmetries for $B^+ \to D_s^+ \overline{D}^0$ decays.

Channel	$N(B^-)$	$N(B^+)$	$A_{\text{raw}}(\%)$
$B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	13659 ± 129	14209 ± 132	-2.0 ± 0.7
$B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	7717 ± 103	7945 ± 104	-1.5 ± 0.9
$B^+ \to D_s^+ \overline{D}^0$, combined	21375 ± 165	22153 ± 168	-1.8 ± 0.5
$B^+ \to D^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$	678 ± 32	660 ± 31	1.3 ± 3.3
$B^+ \to D^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	369 ± 24	345 ± 24	3.4 ± 4.7
$B^+ \to D^+ \overline{D}^0$, combined	1047 ± 40	1005 ± 39	2.0 ± 2.7

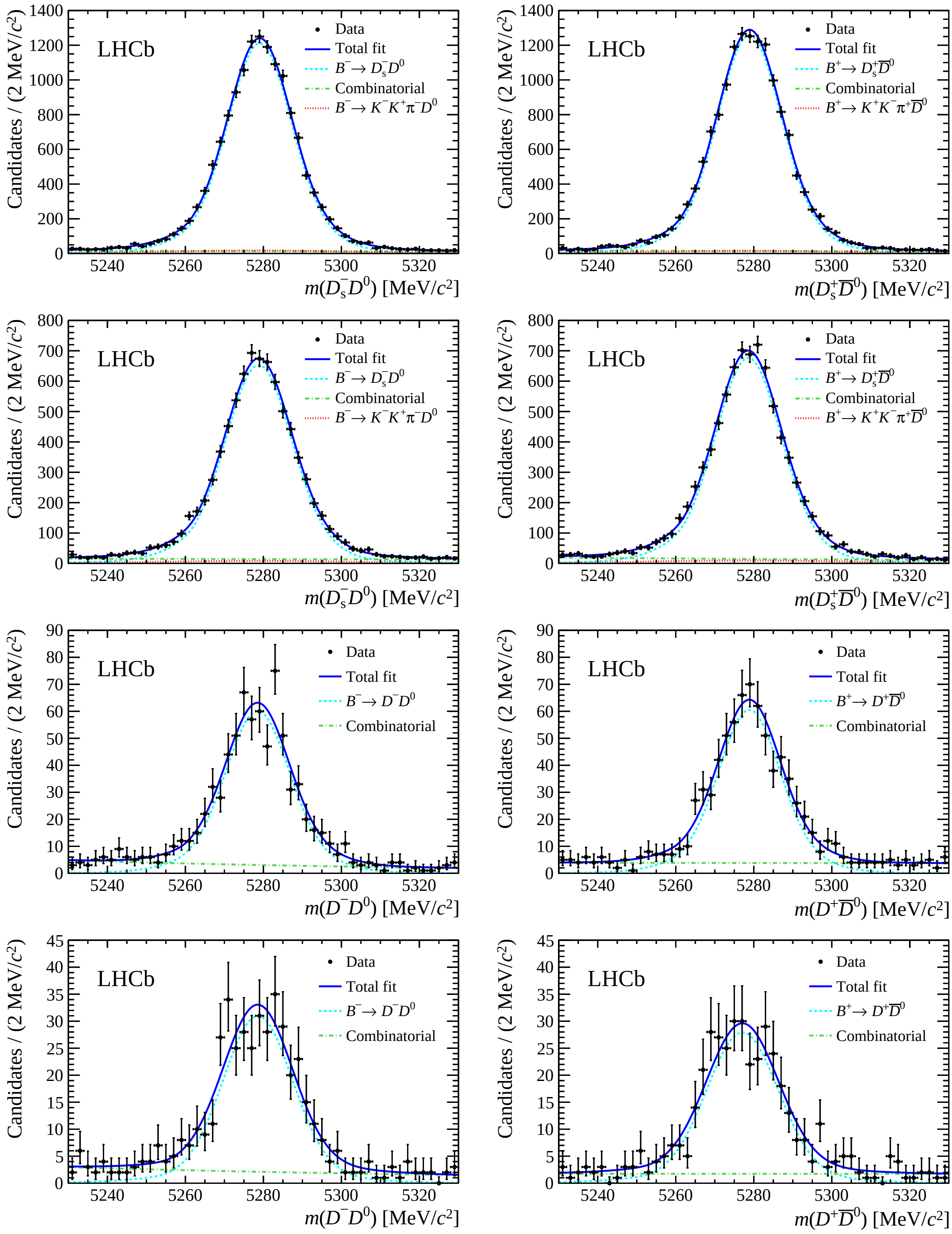


Figure 6.1: Constrained mass distributions of $B^+ \to D^+_{(s)} \overline{D}^0$ candidates, separated by charge. The top and third row plots are with $\overline{D}^0 \to K^+\pi^-$ decays and the second and last row are with $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ decays.

6.3 Detection asymmetry

The presence of a left-right asymmetry in the detection efficiencies can be determined by measuring the raw asymmetry separately for each magnet polarity. If one side of the detector is more efficient than the other, this favours particles of a particular charge that are bent into that side by the magnetic field, creating an asymmetry. This effect is cancelled when combining magnet polarities. Time-dependent detection asymmetries can be identified by determining the raw asymmetry for different data-taking periods. The raw asymmetry for the high-statistics $B^+\to D^+_s\overline{D}^0$ mode, separated by magnet polarity and data-taking year, is shown in table 6.2. None of the asymmetries show a significant difference from the combination, indicating that the left-right and time-dependent asymmetries are small.

Four sources of detection asymmetry are considered: the trigger asymmetry (section 6.3.1); the tracking asymmetry (section 6.3.2); the PID asymmetry (section 6.3.3); and the kaon detection asymmetry (section 6.3.4). Since the asymmetries are small, the total detection asymmetry can be expressed as

$$A_D = A_{\text{trigger}} + A_{\text{tracking}} + A_{\text{PID}} + A_D(K^-\pi^+). \tag{6.6}$$

The contribution of each component to the detection asymmetry is summarised in table 6.3.

6.3.1 Trigger asymmetry

The trigger is made up of two stages: the L0 trigger, implemented in hardware; and the high level trigger, implemented in software (see section 3.3.3). From other studies of CP violation [178], the hardware trigger is known to result in $\mathcal{O}(0.1\%)$ asymmetries, while the software trigger results in $\mathcal{O}(0.01\%)$ asymmetries, which are considered negligible compared to the statistical uncertainties. At L0, the B candidates are required to be either TIS on the global trigger or TOS on the hadron trigger. The trigger selection is described in full in section 4.3.

Table 6.2: Yields and raw asymmetries for $B^+ \to D^+_s \overline{D}^0$ decays, split by $\overline{D}^0$ decay mode, data-taking year and magnet polarity.

Year	Magnet polarity	$\overline{D}^0$	$N(B^-)$	$N(B^+)$	$A_{\text{raw}}(\%)$
2011	*MagUp*	$\overline{D}^0 \to K^+\pi^-$	1649 ± 45	1704 ± 46	-1.6 ± 1.9
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	957 ± 35	978 ± 37	-1.1 ± 2.6
		combined	2606 ± 57	2682 ± 59	-1.4 ± 1.6
2011	*MagDown*	$\overline{D}^0 \to K^+\pi^-$	2376 ± 54	2523 ± 55	-3.0 ± 1.6
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	1321 ± 42	1304 ± 44	0.6 ± 2.3
		combined	3696 ± 68	3828 ± 70	-1.7 ± 1.3
2011	combined	$\overline{D}^0 \to K^+\pi^-$	4024 ± 71	4231 ± 72	-2.5 ± 1.2
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	2286 ± 55	2283 ± 57	0.1 ± 1.7
		combined	6310 ± 90	6513 ± 92	-1.6 ± 1.0
2012	*MagUp*	$\overline{D}^0 \to K^+\pi^-$	4781 ± 77	5073 ± 79	-3.0 ± 1.1
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	2738 ± 62	2958 ± 64	-3.9 ± 1.6
		combined	7519 ± 99	8031 ± 101	-3.3 ± 0.9
2012	*MagDown*	$\overline{D}^0 \to K^+\pi^-$	4860 ± 77	4870 ± 78	-0.1 ± 1.1
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	2696 ± 61	2711 ± 59	-0.3 ± 1.6
		combined	7556 ± 98	7581 ± 98	-0.2 ± 0.9
2012	combined	$\overline{D}^0 \to K^+\pi^-$	9641 ± 108	9979 ± 111	-1.7 ± 0.8
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	5430 ± 87	5660 ± 87	-2.1 ± 1.1
		combined	15071 ± 139	15639 ± 141	-1.8 ± 0.6
2011/12	*MagUp*	$\overline{D}^0 \to K^+\pi^-$	6430 ± 89	6777 ± 91	-2.6 ± 1.0
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	3702 ± 71	3943 ± 74	-3.1 ± 1.3
		combined	10132 ± 114	10720 ± 117	-2.8 ± 0.8
2011/12	*MagDown*	$\overline{D}^0 \to K^+\pi^-$	7227 ± 94	7432 ± 102	-1.4 ± 0.9
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	4019 ± 74	4012 ± 73	0.1 ± 1.3
		combined	11246 ± 119	11444 ± 125	-0.9 ± 0.8
2011/12	combined	$\overline{D}^0 \to K^+\pi^-$	13659 ± 129	14209 ± 132	-2.0 ± 0.7
		$\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	7717 ± 103	7945 ± 104	-1.5 ± 0.9
		combined	21375 ± 165	22153 ± 168	-1.8 ± 0.5

Table 6.3: Detection asymmetries for $B^+ \to D^+_{(s)} \overline{D}^0$ decays. The total is the sum of the individual components, with their uncertainties added in quadrature.

Channel	$A_{\text{trigger}}(\%)$	$A_{\text{tracking}}(\%)$	$A_{\text{PID}}(\%)$	$A_D(K^-\pi^+)(\%)$	total (%)
$B^+ \to D_s^+ \overline{D}^0$	0.0 ± 0.0	0.2 ± 0.1	0.0 ± 0.1	-1.0 ± 0.2	-0.8 ± 0.2
$B^+ \to D^+ \overline{D}^0$	0.0 ± 0.0	0.2 ± 0.1	0.0 ± 0.1	0.0 ± 0.0	0.2 ± 0.1

L0 global TIS

The L0 global TIS asymmetry has been measured with $B \to \overline{D}^0 \mu^+ \nu_\mu X$ decays using Run I data collected by the LHCb experiment [179]. In a $B \to \overline{D}^0 \mu^+ \nu_\mu X$ event, one of the b quarks from the $b\bar{b}$ pair will hadronise into a B meson and decay via $B \to \overline{D}^0 \mu^+ \nu_\mu X$. The rest of the event, including the presence of the b quark that doesn't decay via the signal channel, is identical between $B^+ \to D^+_{(s)} \overline{D}^0$ and $B \to \overline{D}^0 \mu^+ \nu_\mu X$. Hence, the L0 global TIS asymmetry is also valid for $B^+ \to D^+_{(s)} \overline{D}^0$. The L0 global TIS asymmetry multiplied by the TIS fraction of the signal ($\sim 60\%$) results in an asymmetry of 0.04% and is considered negligible.

L0 hadron TOS

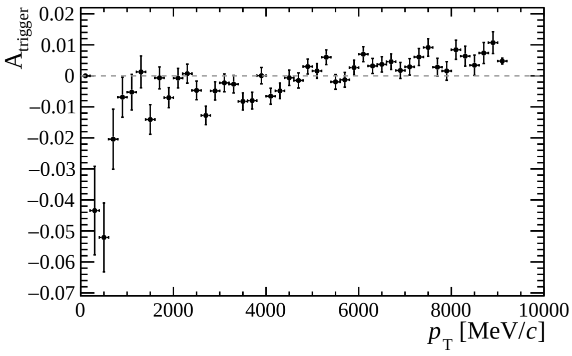
If the HCAL has a non-uniform response, the TOS requirement on the L0 hadron trigger can result in an asymmetry. The performance of the hadron trigger is not well reproduced in simulation, so the efficiency of the TOS requirement is determined from data using a $\overline{D}^0 \to K^+ \pi^-$ calibration sample [180]. The efficiency is given by

$$\varepsilon_{\text{trigger}} = \frac{N(\text{TIS and TOS})}{N(\text{TIS})}, \tag{6.7}$$

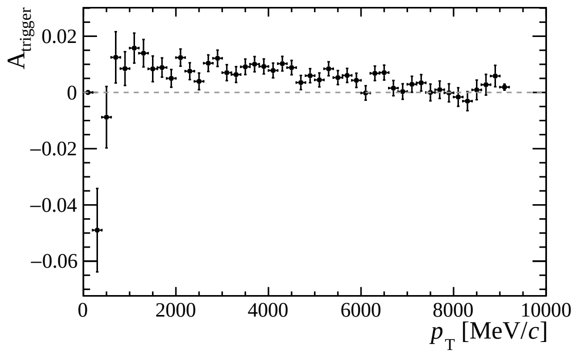
where $N(\text{TIS and TOS})$ is the number of candidates passing both the L0 global TIS and L0 hadron TOS requirements and $N(\text{TIS})$ is the number passing only L0 global TIS. The asymmetry is calculated separately for $h = K, \pi$ with

$$A_{\text{trigger}}(h) = \frac{\varepsilon(h^-) - \varepsilon(h^+)}{\varepsilon(h^-) + \varepsilon(h^+)}, \tag{6.8}$$

and is shown in figure 6.2 as a function of p_{T}. At low p_{T}, there are large negative asymmetries of up to -5% and at high p_{T}, there are small ($< 1\%$) asymmetries.



(a) Kaon



(b) Pion

Figure 6.2: L0 hadron TOS trigger asymmetry as a function of p_{T} in data collected in 2012 with both magnet polarities combined [180].

The asymmetries for the individual hadrons need to be folded with the p_{T} distributions of the final state particles from $B^+ \to D^+_{(s)} \overline{D}^0$ decays and combined into an overall asymmetry for the B^+ meson. In data, only events that pass the L0 trigger requirements are saved, so generator level (without simulation of the detector response) MC was used to model the p_{T} spectra of the tracks without any trigger requirements. A total of $70,000$ events are generated and momentum requirements from the stripping and preselection are applied. The efficiencies of the final state particles are combined using the logical OR,

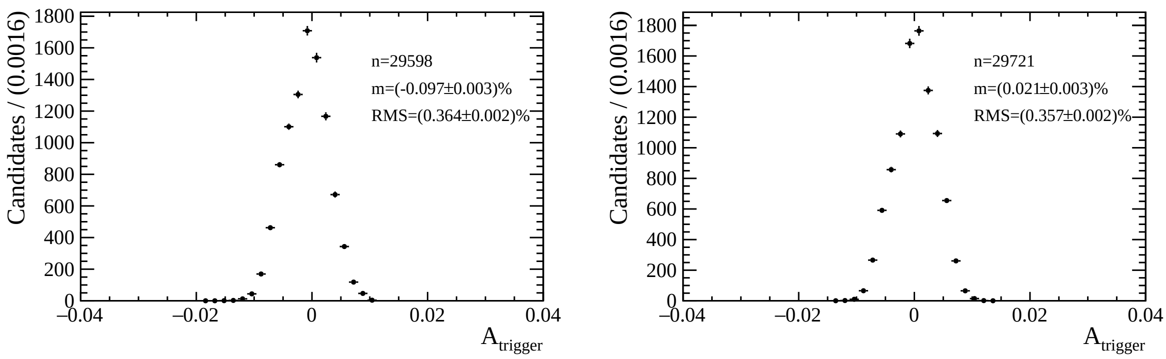
$$\varepsilon(B) = 1 - \prod_{i=1}^{N_{\mathrm{tracks}}} (1 - \varepsilon(h_i)), \tag{6.9}$$

since the event is saved if any of the tracks fire the trigger. It is assumed that the $\varepsilon(h_i)$ are independent. For each event, the efficiency is calculated assuming the B meson was both positively and negatively charged to give an event-by-event asymmetry,

$$A_{\mathrm{trigger}}(B) = \frac{\varepsilon(B^-) - \varepsilon(B^+)}{\varepsilon(B^-) + \varepsilon(B^+)}. \tag{6.10}$$

This is weighted with the average efficiency, $(\varepsilon(B^-) + \varepsilon(B^+))/2$, since events with a higher efficiency are more likely to pass the trigger selection and contribute to the final asymmetry.

Figure 6.3 shows an example of the weighted event-by-event asymmetry, where the mean gives the size of the L0 hadron TOS asymmetry. Averaging by data-taking year, magnet polarity, $\overline{D}^0$ channel and multiplying by the TOS fraction of the signal (64%



(a) $B^+ \to D_s^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ (b) $B^+ \to D^+\overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$

Figure 6.3: Event-by-event L0 hadron TOS asymmetry, weighted by the average efficiency, in data collected in 2012 with both magnet polarities combined.

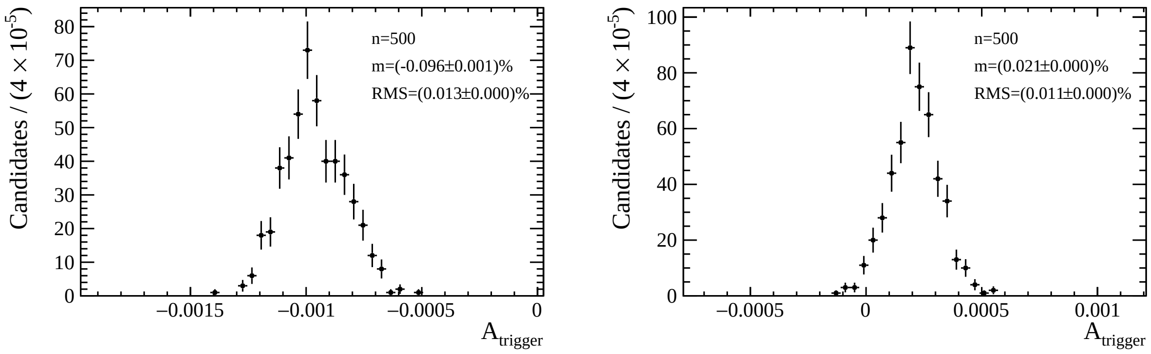
for $\overline{D}^0 \to K^+\pi^-$ and 61% for $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$) results in an asymmetry of -0.045% for $B^+ \to D_s^+\overline{D}^0$ and 0.030% for $B^+ \to D^+\overline{D}^0$. The uncertainty on the mean is $\mathcal{O}(0.001\%)$ and is considered negligible.

Only momentum cuts from the preselection and stripping are applied to the generator level MC. The track p_{T} distributions are used as input to the MVA, as well as other slightly correlated variables; hence there may be differences between the p_{T} distributions in MC and data. To estimate the size of this effect, the asymmetry is re-evaluated without applying any momentum cuts. The resulting change is less than 0.03%.

The efficiencies calculated using the $\overline{D}^0 \to K^+\pi^-$ calibration sample have an associated uncertainty. To take this into account, the calculation is repeated varying the efficiencies according to a Gaussian distribution of width the value of the uncertainty. Figure 6.4 shows an example of the distribution of the asymmetry recalculated in this way 500 times. The root mean square (RMS) of the distributions in all channels and data-taking years is 0.03% or less. As the size of the L0 hadron TOS asymmetry and its uncertainties are below 0.05%, no correction is applied to the *CP* asymmetry measurement.

6.3.2 Tracking asymmetry

The tracking efficiency has been studied in Run I data using pions from $D^{*+} \to \pi^+(D^0 \to K^-\pi^+\pi^-\pi^+)$ decays [181]. The efficiency was determined as a function of momentum by comparing the partially reconstructed D^{*+} yields with the fully reconstructed yields when a pion is



(a) $B^+\to D_s^+\overline{D}^0$, $\overline{D}^0\to K^+\pi^-$ **(b)** $B^+\to D^+\overline{D}^0$, $\overline{D}^0\to K^+\pi^-$

Figure 6.4: L0 hadron TOS asymmetry, calculated by varying the efficiencies according to Gaussian distributions, using data collected in 2012 with both magnet polarities combined.

added,

$$\varepsilon_{\text{tracking}}(\pi) = \frac{N_{\text{full}}(\pi)}{N_{\text{partial}}(\pi)}. \tag{6.11}$$

Figure 6.5 shows the asymmetry derived from these efficiencies. Although each individual magnet polarity shows a strong dependence on momentum, a large cancellation is expected from their combination.

A total of 70,000 generator level MC events are used to model the momentum distributions of the $B^+\to D^+_{(s)}\overline{D}^0$ tracks. Momentum requirements from the stripping and preselection are applied. As the tracking efficiency applies to all final state tracks,

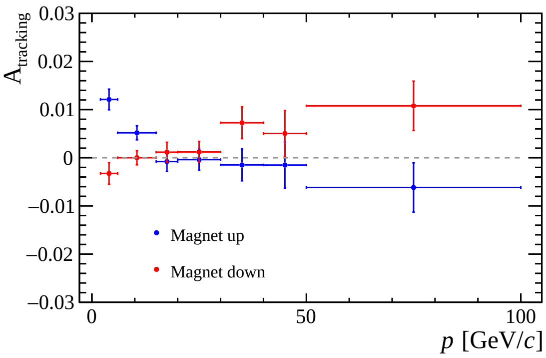
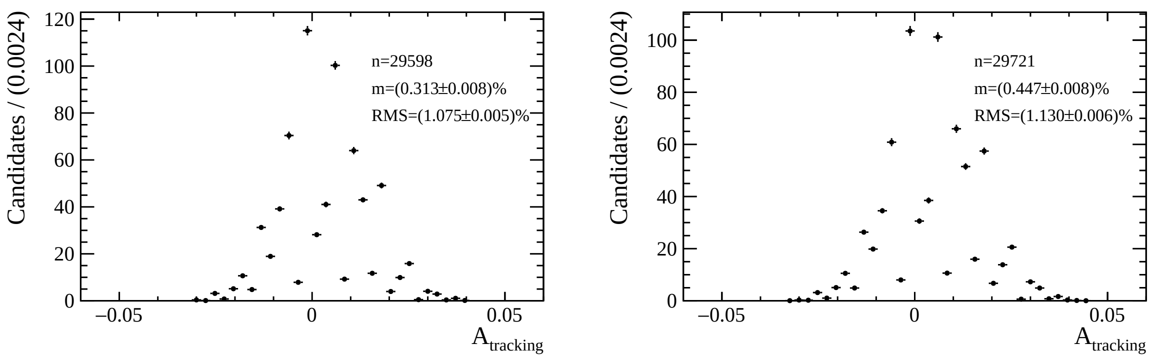


Figure 6.5: Pion tracking asymmetry in data collected in 2012 [181].



(a) $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ **(b)** $B^+ \to D^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$

Figure 6.6: Event-by-event tracking asymmetry, weighted by the average efficiency, in data collected in 2012 with the *MagUp* polarity.

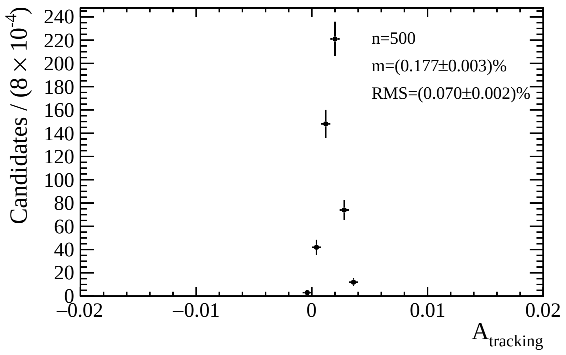
the efficiencies are combined using the logical AND,

$$\varepsilon(B) = \prod_{i=1}^{N_{\text{tracks}}} \varepsilon(h_i), \tag{6.12}$$

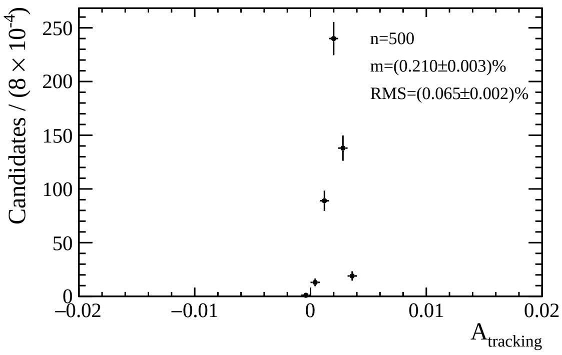
where the kaons in the final state were treated as pions. As for the L0 hadron TOS asymmetry (see section 6.3.1), an event-by-event asymmetry is calculated assuming the B meson is both positively and negatively charged, then weighted with the average efficiency.

Figure 6.6 shows an example of the distribution of the weighted event-by-event asymmetry, where the mean gives the size of the tracking asymmetry. As there are fewer bins in the measurement of the tracking asymmetry than in the L0 hadron TOS asymmetry, the distribution of the event-by-event asymmetry is not smoothly-varying. Averaging by data-taking year, magnet polarity and $\overline{D}^0$ channel results in an asymmetry of +0.18% for $B^+ \to D_s^+ \overline{D}^0$ and +0.21% for $B^+ \to D^+ \overline{D}^0$. The uncertainty on the mean is $\mathcal{O}(0.01\%)$ and is considered negligible.

The same sources of systematic uncertainty as for the L0 hadron TOS asymmetry are considered. To evaluate the effect of only applying momentum requirements from the preselection and stripping, the calculation is repeated without any requirements. The resulting change in the asymmetry is less than 0.03%. To take into account the uncertainties on the efficiencies, the calculation is repeated varying the efficiencies according to Gaussian distributions 500 times. Figure 6.7 shows the distribution of



(a) $B^+ \to D_s^+ \overline{D}^0$



(b) $B^+ \to D^+ \overline{D}^0$

Figure 6.7: Tracking asymmetry, calculated by varying the efficiencies according to Gaussian distributions, with data-taking years, magnet polarities and $\overline{D}^0$ channels combined.

the asymmetry. This results in an uncertainty of 0.07% for both $B^+ \to D_s^+ \overline{D}^0$ and $B^+ \to D^+ \overline{D}^0$ decays and is the dominant uncertainty on the tracking asymmetry.

6.3.3 PID asymmetry

The PID information, which is based on the response of the RICH detectors, is dependent on magnet polarity and track charge. The efficiencies of a number of PID requirements were studied using samples of $\overline{D}^0 \to K^+\pi^-$ decays, selected using only kinematical and topological information, as a function of p, η and the number of tracks (nTracks) in the event [182]. By folding these efficiencies with the distributions of p, η and nTracks from $B^+ \to D^+_{(s)} \overline{D}^0$ simulation without any PID requirements applied, the efficiency of any given $\text{DLL}_{K\pi}$ requirement can be determined. The efficiency of the requirements in the preselection ($\text{DLL}_{K\pi}(K) > -5$ and $\text{DLL}_{K\pi}(\pi) < 10$) is $\sim 95\%$. The $\text{DLL}_{K\pi}$ distributions are used as input to the MVA, so the precise requirements placed on them are unknown. However, as the efficiency of the MVA requirement is $\sim 90\%$, the minimum PID efficiency is the product of the two, $95\% \times 90\% = \sim 85\%$.

Table 6.4 shows the efficiencies of $\text{DLL}_{K\pi}(K) > x$ and $\text{DLL}_{K\pi}(\pi) < -x$ requirements for $x = 0, 10, 20$, on all final state particles from $B^+ \to D^+_{(s)} \overline{D}^0$ decays. After combining magnet polarities, the asymmetry induced by the $x = 0$ requirement is 0.11% or less for both channels and data-taking years. Since the efficiency of this requirement is lower than the minimum PID efficiency, the asymmetry is expected to be smaller than 0.11%. Therefore, no correction is applied to the CP asymmetry measurement and a conservative systematic uncertainty of 0.1% is assigned.

It could be argued that the loss of signal efficiency from applying the PID requirements is due to a small number of events experiencing a requirement tighter than $x = 0$, *i.e.* $x = 20$. Table 6.4 shows that, after combining magnet polarities, the PID asymmetry for the $x = 20$ requirement, is 0.35% or less for both channels and data-taking years. As the minimum PID efficiency is $\sim 85\%$ and the efficiency of the $x = 20$ requirement is extremely low, at most 15% of events can experience this requirement. The asymmetry is therefore diluted to $15\% \times 35\% \sim 0.05\%$, which is well covered by the assigned 0.1% systematic uncertainty.

6.3.4 Kaon detection asymmetry

Positively charged kaons are composed of u and $\bar{s}$ quarks, whereas negatively charged kaons are composed of s and $\bar{u}$ quarks. As the LHCb detector largely consists of u and d quarks in the form of protons and neutrons, K^- mesons interact more strongly with the detector material than K^+ mesons due to $u\bar{u}$ annihilation. This effect is momentum dependent since, as the momentum of the kaon increases, the sea quarks become more likely to interact with the detector material than the valence quarks, and the detection asymmetry decreases.

The final states of $B^+\to D^+\overline{D}^0$ and $B^+\to D_s^+\overline{D}^0$ decays are $K^-\pi^+\pi^+K^+\pi^-(\pi^+\pi^-)$ and $K^+K^-\pi^+K^+\pi^-(\pi^+\pi^-)$, respectively. As there are two oppositely charged kaons in the final state of the $B^+\to D^+\overline{D}^0$ decay, the kaon detection asymmetry is expected to largely cancel. However, as there are an odd number of kaons in the final state of the $B^+\to D_s^+\overline{D}^0$ decay, there is no such cancellation and an asymmetry is expected.

The difference in detection asymmetry between kaons and pions, $A_D(K^-\pi^+)$, has been measured as a function of kaon momentum by comparing the yield of $D^+\to K^-\pi^+\pi^-$ decays with the yield of $D^+\to \overline{K}^0\pi^+$ decays [183]. The raw asymmetry in $D^+\to K^-\pi^+\pi^+$ decays is,

$$A_{\text{raw}}(K^-\pi^+\pi^+) = A_P(D^+) + A_D(K^-\pi^+) + A_D(\pi^+), \tag{6.13}$$

where $A_P(D^+)$ is the D^+ production asymmetry and $A_D(\pi^+)$ is the pion detection asymmetry. The raw asymmetry in $D^+\to \overline{K}^0\pi^+$ decays is,

$$A_{\text{raw}}(\overline{K}^0\pi^+) = A_P(D^+) + A_D(\pi^+) - A_D(K^0), \tag{6.14}$$

Table 6.4: Efficiency of $\mathrm{DLL}_{K\pi}\ (K) > x$ and $\mathrm{DLL}_{K\pi}\ (\pi) < -x$ cuts for the $B^+ \to D^+_{(s)}\overline{D}^0$ final states with $\overline{D}^0 \to K^+\pi^-$ decays, split by charge and magnet polarity. The combination of magnet up and magnet down takes into account that approximately 43% of events in 2011 were recorded with the magnet up polarity, whereas in 2012 the split is approximately even.

$\varepsilon_{\mathrm{PID}}$ (%)		$x=0$			$x=10$			$x=20$		
		MagUp	*MagDown*	comb.	*MagUp*	*MagDown*	comb.	*MagUp*	*MagDown*	comb.
2011	$B^+ \to D^+\overline{D}^0$	62.09	62.23	62.17	14.89	14.93	14.91	2.199	2.225	2.214
	$B^- \to D^-D^0$	62.38	62.16	62.25	14.92	14.86	14.89	2.210	2.225	2.219
	A_{PID} (%)	0.23	-0.06	0.06	0.10	-0.23	-0.07	0.249	0.000	0.113
	$B^+ \to D_s^+\overline{D}^0$	67.26	67.34	67.31	17.39	17.79	17.62	2.341	2.440	2.397
	$B^- \to D_s^-D^0$	67.42	67.49	67.46	17.50	17.66	17.59	2.366	2.414	2.393
	A_{PID} (%)	0.12	0.11	0.11	0.32	-0.37	-0.09	0.531	-0.536	-0.084
2012	$B^+ \to D^+\overline{D}^0$	65.13	66.59	65.86	15.87	17.11	16.49	2.382	2.700	2.541
	$B^- \to D^-D^0$	65.34	66.52	65.93	16.04	17.08	16.56	2.420	2.698	2.559
	A_{PID} (%)	0.16	-0.05	0.05	0.53	-0.09	0.21	0.791	-0.037	0.353
	$B^+ \to D_s^+\overline{D}^0$	69.50	70.79	70.15	18.24	19.64	18.94	2.497	2.832	2.665
	$B^- \to D_s^-D^0$	69.65	70.42	70.04	18.38	19.48	18.93	2.525	2.777	2.651
	A_{PID} (%)	0.11	-0.26	-0.08	0.38	-0.41	-0.03	0.558	-0.981	-0.263

where $A_D(K^0)$ is the detection asymmetry of $K^0 \to \pi^+\pi^-$ decays. The measurement of $A_D(K^0)$ takes into account CP violation, $K^0 - \overline{K}^0$ mixing and the difference in the interaction cross-section of K^0 and $\overline{K}^0$ with the detector material. As $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to \overline{K}^0\pi^+$ decays are Cabibbo favoured, it is assumed there is negligible CP violation. Hence, $A_D(K^-\pi^+)$ is obtained with,

$$A_D(K^-\pi^+) = A_{\text{raw}}(K^-\pi^+\pi^+) - A_{\text{raw}}(\overline{K}^0\pi^+) - A_D(K^0). \tag{6.15}$$

Table 6.5 shows the asymmetry calculated by folding the measurement of $A_D(K^-\pi^+)$ with the momentum spectra of kaons from simulated $B^+ \to D^+_{(s)}\overline{D}^0$ decays. The uncertainties on the measurement of $A_D(K^-\pi^+)$ are taken into account by repeating the calculation with the values of $A_D(K^-\pi^+)$ shifted up or down by 1σ. As expected, $A_D(K^-\pi^+)$ in $B^+ \to D^+\overline{D}^0$ decays is consistent with zero, whereas a significantly non-zero value of $(-1.0 \pm 0.2)\%$ is measured for $B^+ \to D^+_s\overline{D}^0$ decays. As the kaons are treated as pions in the evaluation of the tracking asymmetry, the sum of $A_D(K^-\pi^+)$ and A_{tracking} isolates the kaon detection asymmetry.

6.4 Production asymmetry

The content of B^+ mesons is u and $\bar{b}$ quarks, whereas B^- mesons consist of b and $\bar{u}$ quarks. As the LHC is a pp, or $uuduud$, collider, the production of B^+ mesons is favoured over B^- mesons. Hence, from equation 6.5, a negative production asymmetry is expected. The raw asymmetry needs to be corrected for this effect in order to determine the CP asymmetry.

The B^+ production asymmetry has been measured in Run I data using $B^+ \to \overline{D}^0\pi^+$ decays [179]. The raw asymmetry is related to the production asymmetry as follows,

$$A_{\text{raw}}(B^+ \to \overline{D}^0\pi^+) = A_P(B^+) + A_D(\overline{D}^0\pi^+) + \mathcal{A}^{CP}(B^+ \to \overline{D}^0\pi^+), \tag{6.16}$$

where $A_D(\overline{D}^0\pi^+)$ is the detection asymmetry of the $\overline{D}^0\pi^+$ final state and $\mathcal{A}^{CP}(B^+ \to \overline{D}^0\pi^+)$ is the CP asymmetry. The CP asymmetry was estimated from measurements of the CKM angle γ and the hadronic parameters of $B^+ \to \overline{D}^0\pi^+$ decays. Averaging the production asymmetries measured at 7 and 8 TeV results in a value of $A_P(B^+) = (-0.5 \pm 0.4)\%$. In the measurement of $A_P(B^+)$, no significant dependence on the production spectrum (p_{T}, y) was observed.

Table 6.5: Difference in kaon and pion detection asymmetry for $B^+ \to D^+_{(s)}\overline{D}^0$ decays. The values quoted are for the underlined particle. The asymmetry of the B meson is the sum of the individual kaon asymmetries, multiplied by their charge.

Particle	$\langle p \rangle$ (GeV/c)	$A_D(K^-\pi^+)$ (%)
$B^+ \to D^+_s(\underline{K^+}K^-\pi^+)\overline{D}^0(K^+\pi^-)$	25.6	-1.10 ± 0.14
$B^+ \to D^+_s(K^+\underline{K^-}\pi^+)\overline{D}^0(K^+\pi^-)$	25.8	-1.10 ± 0.14
$B^+ \to D^+_s(K^+K^-\pi^+)\overline{D}^0(\underline{K^+}\pi^-)$	32.6	-1.03 ± 0.16
$\underline{B^+} \to D^+_s(K^+K^-\pi^+)\overline{D}^0(K^+\pi^-)$		-1.03 ± 0.16
$B^+ \to D^+_s(\underline{K^+}K^-\pi^+)\overline{D}^0(K^+\pi^-\pi^+\pi^-)$	27.5	-1.07 ± 0.15
$B^+ \to D^+_s(K^+\underline{K^-}\pi^+)\overline{D}^0(K^+\pi^-\pi^+\pi^-)$	28.0	-1.07 ± 0.15
$B^+ \to D^+_s(K^+K^-\pi^+)\overline{D}^0(\underline{K^+}\pi^-\pi^+\pi^-)$	30.2	-1.05 ± 0.15
$\underline{B^+} \to D^+_s(K^+K^-\pi^+)\overline{D}^0(K^+\pi^-\pi^+\pi^-)$		-1.05 ± 0.15
$B^+ \to D^+_s\overline{D}^0$, $\overline{D}^0$ combined		-1.04 ± 0.16
$B^+ \to D^+(\underline{K^-}\pi^+\pi^+)\overline{D}^0(K^+\pi^-)$	28.2	-1.06 ± 0.15
$B^+ \to D^+(K^-\pi^+\pi^+)\overline{D}^0(\underline{K^+}\pi^-)$	32.1	-1.02 ± 0.16
$\underline{B^+} \to D^+(K^-\pi^+\pi^+)\overline{D}^0(K^+\pi^-)$		0.04 ± 0.01
$B^+ \to D^+(\underline{K^-}\pi^+\pi^+)\overline{D}^0(K^+\pi^-\pi^+\pi^-)$	30.4	-1.03 ± 0.16
$B^+ \to D^+(K^-\pi^+\pi^+)\overline{D}^0(\underline{K^+}\pi^-\pi^+\pi^-)$	29.9	-1.03 ± 0.15
$\underline{B^+} \to D^+(K^-\pi^+\pi^+)\overline{D}^0(K^+\pi^-\pi^+\pi^-)$		0.00 ± 0.00
$B^+ \to D^+\overline{D}^0$, $\overline{D}^0$ combined		0.02 ± 0.01

6.5 Results

The CP asymmetry is determined by subtracting the production (see section 6.4) and detection asymmetries (see table 6.3) from the raw asymmetry, according to equation 6.3. The results are

$$\mathcal{A}^{CP}(B^+ \to D^+_s\overline{D}^0) = (-0.4 \pm 0.5 \pm 0.5)\%, \tag{6.17}$$

$$\mathcal{A}^{CP}(B^+ \to D^+\overline{D}^0) = (+2.3 \pm 2.7 \pm 0.4)\%, \tag{6.18}$$

where the first uncertainty is statistical and the second systematic. The dominant source of systematic uncertainty originates from the production asymmetry. As the production asymmetry is only dependent on the flavour of the B meson and not the decay mode, it cancels in the difference of the two measurements,

$$\mathcal{A}^{CP}(B^+ \to D_s^+ \overline{D}^0) - \mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0) = (-2.7 \pm 2.8 \pm 0.2)\%. \tag{6.19}$$

This is the first ever measurement of $\mathcal{A}^{CP}(B^+ \to D_s^+ \overline{D}^0)$ and the measurement of $\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0)$ is the most precise to date, improving upon the precision of the previous world-average value, $(-3 \pm 7)\%$ [3], by a factor of 2.6. The results are consistent with zero and hence there is no evidence of *CP* violation. The measurements of $\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0)$ and $\mathcal{A}^{CP}(B^+ \to D_s^+ \overline{D}^0)$ place constraints on the range of allowed values predicted by R-parity violating supersymmetry [18] and the fourth generation model [19], respectively.

Chapter 7

Summary and outlook

"Endings to be useful must be inconclusive."
— Samuel R. Delany

The SM provides an incredibly successful description of nature, however, it is known to be incomplete. One of its more glaring flaws is its prediction that we should not exist, or rather, that almost all matter and antimatter should have annihilated in the early universe. This is because the amount of CP violation included in the SM is too small to generate a large enough baryon asymmetry to explain astronomical observations [11]. Therefore, there must be additional, new physics sources of CP violation and physicists at the LHCb experiment are studying the decays of particles containing heavy quarks to try and find them. In this thesis, two such analyses involving doubly charmed B decays have been described. In this chapter, the results are summarised (section 7.1), followed by a discussion of the future prospects (section 7.2).

7.1 Summary

CP violation is included in the SM as a complex phase in the 3×3 unitary CKM matrix. The unitarity of the CKM matrix provides several constraints that can be represented geometrically as triangles. The least well-known angle of the unitarity triangle with sides of similar sizes is γ. It is also the only angle that can be measured with tree-level decays, making it an important SM benchmark since it is unlikely to be affected by new physics. Measurements of γ have been made with B^+, B^0 and B_s^0 decays, but not yet with B_c^+

decays [57]. Sensitivity to γ arises in the interference of $b \to c$ and $b \to u$ decays, which correspond to decays of $B_c^+ \to D_{(s)}^{(*)+} \overline{D}^{(*)0}$ and $B_c^+ \to D_{(s)}^{(*)+} D^{(*)0}$, respectively [12–17].

A search for $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ decays has been performed using Run I data collected by the LHCb experiment [1]. The B_c^+ decays were measured relative to the high-yield $B^+ \to D_{(s)}^+ \overline{D}^0$ normalisation channel. No evidence for signal was found. The measured branching fractions and upper limits set at 90 (95%) for the fully reconstructed B_c^+ decays are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^+ \overline{D}^0)}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (3.0 \pm 3.7) \times 10^{-4}\ [< 0.9\,(1.1) \times 10^{-3}], \tag{7.1}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^+ D^0)}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (-3.8 \pm 2.6) \times 10^{-4}\ [< 3.7\,(4.7) \times 10^{-4}], \tag{7.2}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^+ \overline{D}^0)}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (8.0 \pm 7.5) \times 10^{-3}\ [< 1.9\,(2.2) \times 10^{-2}], \tag{7.3}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^+ D^0)}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (2.9 \pm 5.3) \times 10^{-3}\ [< 1.2\,(1.4) \times 10^{-2}]. \tag{7.4}$$

The results for the B_c^+ decays with one excited charm meson in the final state are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^{*+} \overline{D}^0) + \mathcal{B}(B_c^+ \to D_s^+ \overline{D}^{*0})}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (-0.1 \pm 1.5) \times 10^{-3}\ [< 2.8\,(3.4) \times 10^{-3}], \tag{7.5}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^{*+} D^0) + \mathcal{B}(B_c^+ \to D_s^+ D^{*0})}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (-0.3 \pm 1.9) \times 10^{-3}\ [< 3.0\,(3.6) \times 10^{-3}], \tag{7.6}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to (D^{*+} \to D^+\pi^0/\gamma) \overline{D}^0) + \mathcal{B}(B_c^+ \to D^+ \overline{D}^{*0})}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (0.2 \pm 3.2) \times 10^{-2}\ [< 5.5\,(6.6) \times 10^{-2}], \tag{7.7}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to (D^{*+} \to D^+\pi^0/\gamma) D^0) + \mathcal{B}(B_c^+ \to D^+ D^{*0})}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (-1.5 \pm 1.7) \times 10^{-2}\ [< 2.2\,(2.8) \times 10^{-2}]. \tag{7.8}$$

The results for the B_c^+ decays with two excited charm mesons in the final state are

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^{*+} \overline{D}^{*0})}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (3.2 \pm 4.3) \times 10^{-3}\ [< 1.1\,(1.3) \times 10^{-2}], \tag{7.9}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D_s^{*+} D^{*0})}{\mathcal{B}(B^+ \to D_s^+ \overline{D}^0)} = (7.0 \pm 9.2) \times 10^{-3}\ [< 2.0\,(2.4) \times 10^{-2}], \tag{7.10}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^{*+} \overline{D}^{*0})}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (3.4 \pm 2.3) \times 10^{-1}\ [< 6.5\,(7.3) \times 10^{-1}], \tag{7.11}$$

$$\frac{f_c}{f_u}\frac{\mathcal{B}(B_c^+ \to D^{*+} D^{*0})}{\mathcal{B}(B^+ \to D^+ \overline{D}^0)} = (-4.1 \pm 9.1) \times 10^{-2}\ [< 1.3\,(1.6) \times 10^{-1}]. \tag{7.12}$$

The limits are an order of magnitude greater than the largest theoretical prediction. Hence, the experimental measurements and theoretical predictions are consistent.

A measurement of the *CP* asymmetry in $B^+ \to D^+_{(s)} \overline{D}^0$ decays has been performed using Run I data collected by the LHCb experiment [2]. A non-zero *CP* asymmetry requires two interfering contributions with different weak and strong phases. Decays of $B^+ \to D^+_{(s)} \overline{D}^0$ receive contributions from both tree-level and loop-level diagrams and are predicted to have a *CP* asymmetry of $\mathcal{O}(1\%)$ in the SM [18, 19, 78–80]. New physics can enhance the *CP* asymmetry in these decays to $\mathcal{O}(10\%)$ [18, 19]. The *CP* asymmetry was measured by determining the asymmetry in the signal yields, then correcting this to account for the production and detection asymmetries. The results are

$$\mathcal{A}^{CP}(B^+ \to D^+_s \overline{D}^0) = (-0.4 \pm 0.5 \pm 0.5)\%, \tag{7.13}$$

$$\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0) = (+2.3 \pm 2.7 \pm 0.4)\%, \tag{7.14}$$

where the first uncertainty is statistical and the second systematic. This is the first measurement of $\mathcal{A}^{CP}(B^+ \to D^+_s \overline{D}^0)$ and the measurement of $\mathcal{A}^{CP}(B^+ \to D^+ \overline{D}^0)$ improves upon the precision of the world average value by a factor of 2.6 [3]. The results are consistent with *CP* symmetry and can be used to constrain new physics contributions [18, 19]. All in all, the SM wins again.

7.2 Outlook

To date, the LHCb detector has only collected $\sim 1\%$ of its total expected yield in the channels of interest (see table 2.2). The immediate next step is to repeat the analyses with the addition of Run II data, which is expected to have yields 3.5 times the size of those in the Run I data. With the combined Run I and II data set, a yield of 6.5 events is predicted in the channel with the largest theory prediction, $B^+_c \to D^+ \overline{D}^0$, using optimistic assumptions (see table 2.3). With an upward fluctuation, evidence of this decay could be seen. With the addition of Run III data in 2023, this estimation increases to 47 events, at which point observation is likely. Once the branching fraction of $B^+_c \to D^+ \overline{D}^0$ decays has been measured, more accurate predictions can be made for $B^+_c \to D^+_s D$ decays. In 2033, when the full data set up to Run V has been collected, it may be possible to make a measurement of γ with an uncertainty of around 20°.

Currently, some sensitivity is lost in the $B_c^+ \to D^{*+} D^{(*)}$ channels since the decay of $D^{*+} \to D^0 \pi^+$ with a branching fraction of $\sim 68\%$ is not being reconstructed. However, this mode could be added in a future update. The $B_c^+ \to D^{*+} D$ decay would therefore be fully reconstructed and have a sensitivity similar to $B_c^+ \to D_{(s)}^+ D$ decays. The analysis of this mode would also provide information to help distinguish $B_c^+ \to D^{*+} D$ and $B_c^+ \to D^+ D^*$ in the limits set with partially reconstructed $D^{*+} \to D^+ \pi^0/\gamma$ decays (see equations 7.7 and 7.8). In addition, the CP asymmetry in $B^+ \to D^{*+} \overline{D}^0$ decays could be measured.

The dominant systematic uncertainty in the partially reconstructed B_c^+ modes is the choice of background model due to their wide B mass distributions (see table 5.3). Instead of placing a requirement on the MVA response based on a figure of merit, the fit could be performed simultaneously over bins of MVA response. The addition of low signal purity bins to the fit would help to constrain the background model, reducing the systematic uncertainty.

It is an exciting time to be a particle physicist. The field has made enormous leaps forward, both theoretically and experimentally, since Ernest Rutherford discovered the structure of the atom just over one hundred years ago. The LHC, the most powerful particle accelerator ever built, is subjecting the SM to its most rigorous testing yet. The LHCb experiment now faces some healthy competition from the Belle II experiment [184], a "B factory" that started collecting data earlier this year. Although Belle II doesn't operate at a high enough centre-of-mass energy to produce B_c^+ mesons, it will be able to make measurements of many key flavour physics observables using B^+, B^0 and B_s^0 decays [185, 186]. If there is new physics at the electroweak scale, it hopefully won't be able to hide much longer.

Appendix A

Data/simulation comparison

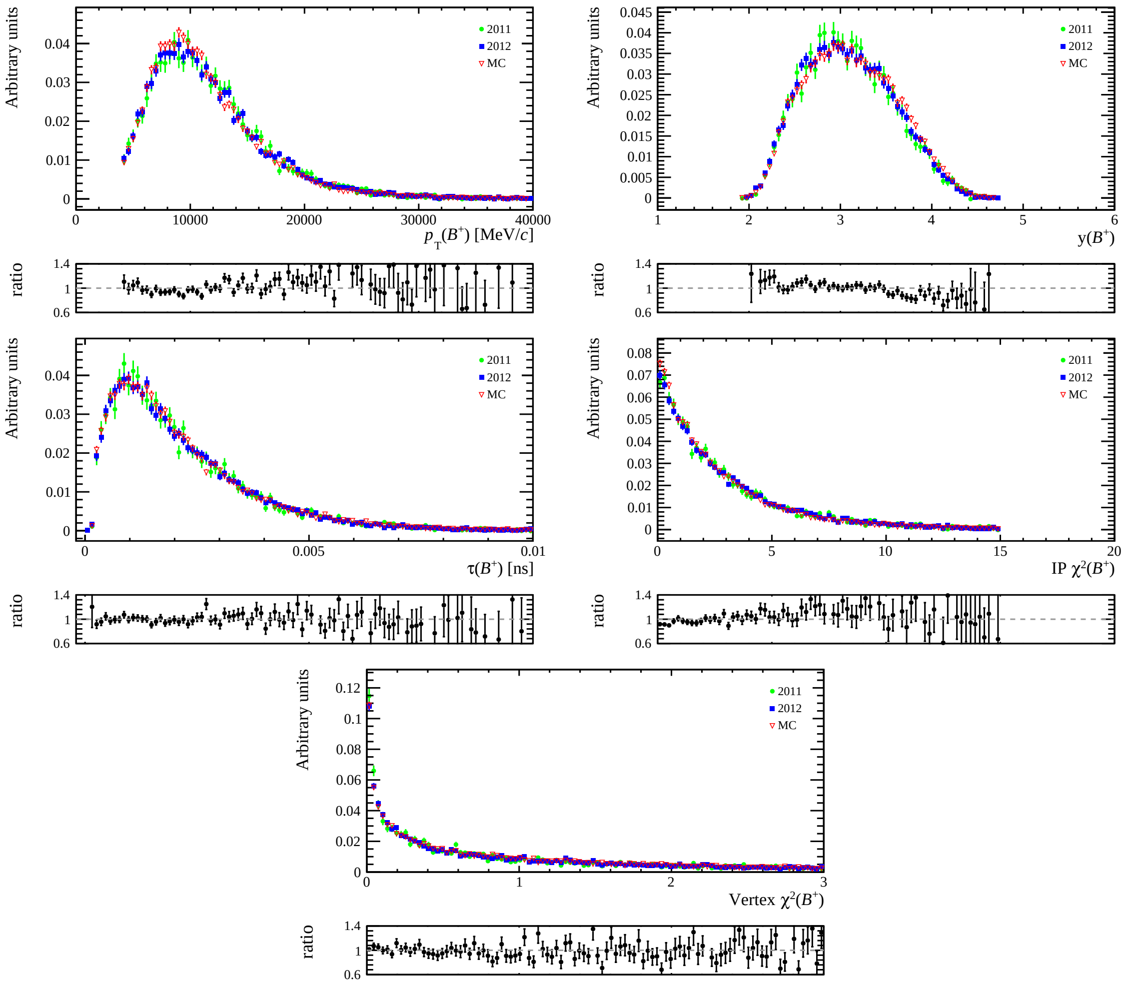


Figure A.1: Comparison of variables related to the reconstructed B^+ between data and simulation distributions with $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ decays.

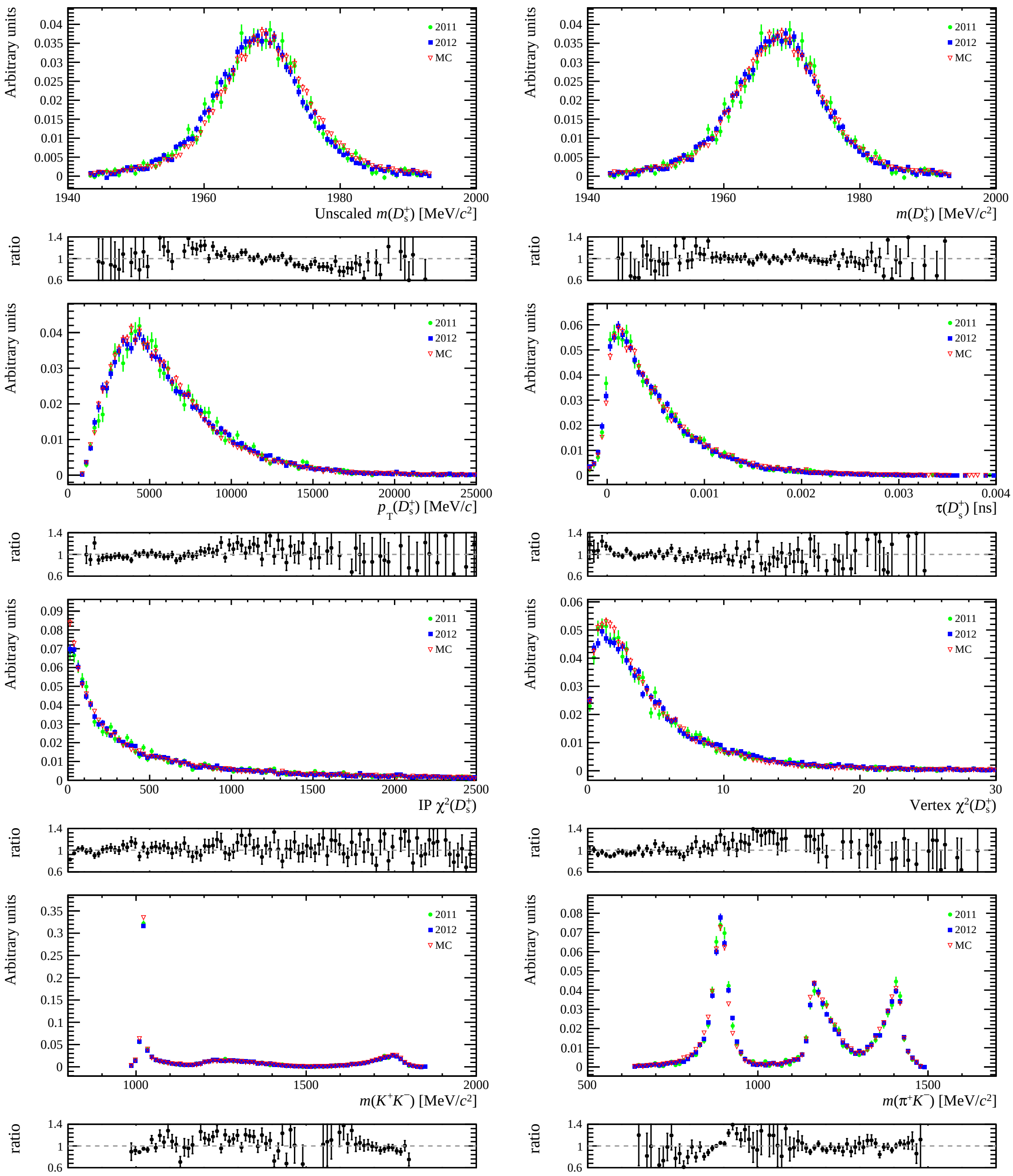


Figure A.2: Comparison of variables related to the reconstructed D_s^+ between data and simulation with $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ decays.

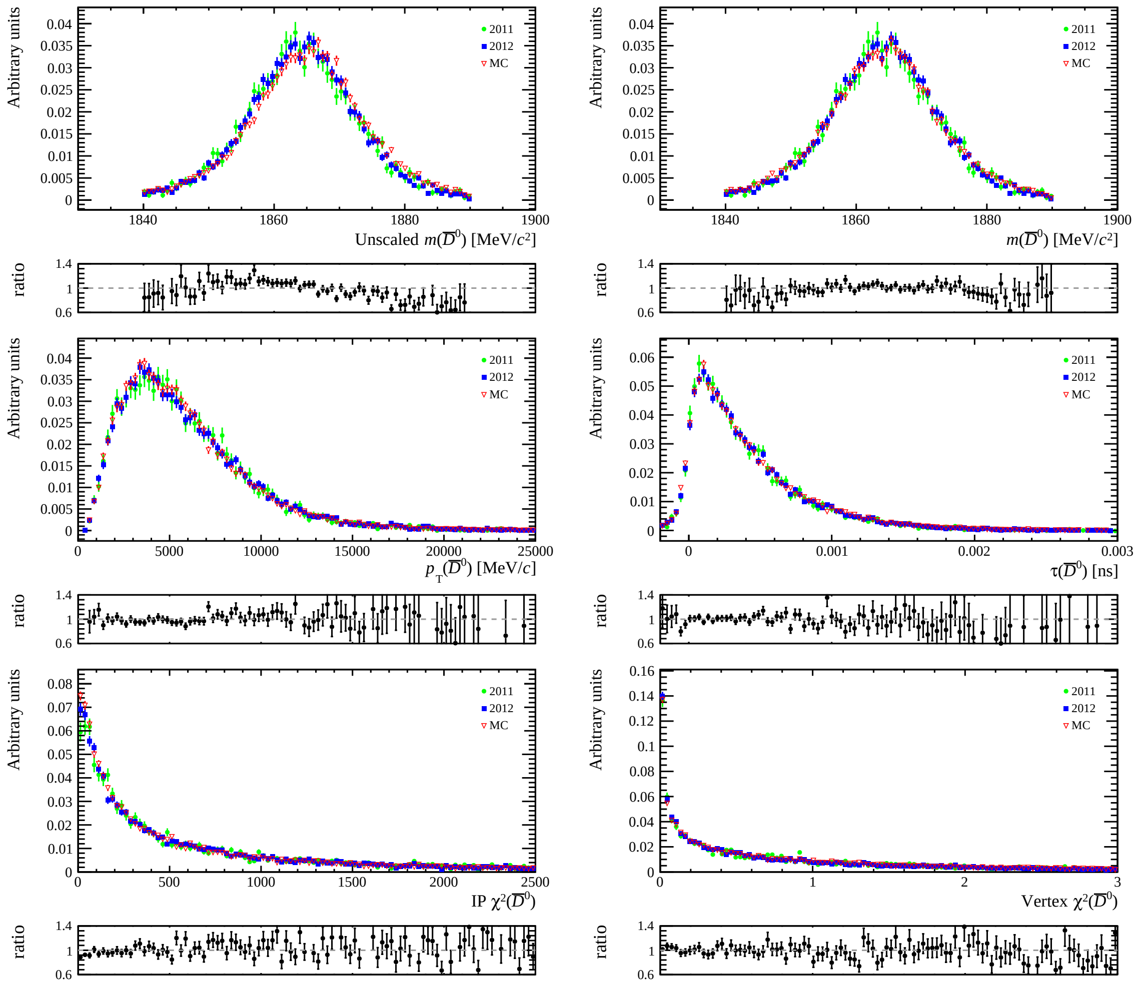


Figure A.3: Comparison of variables related to the reconstructed $\overline{D}^0$ between data and simulation with $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ decays.

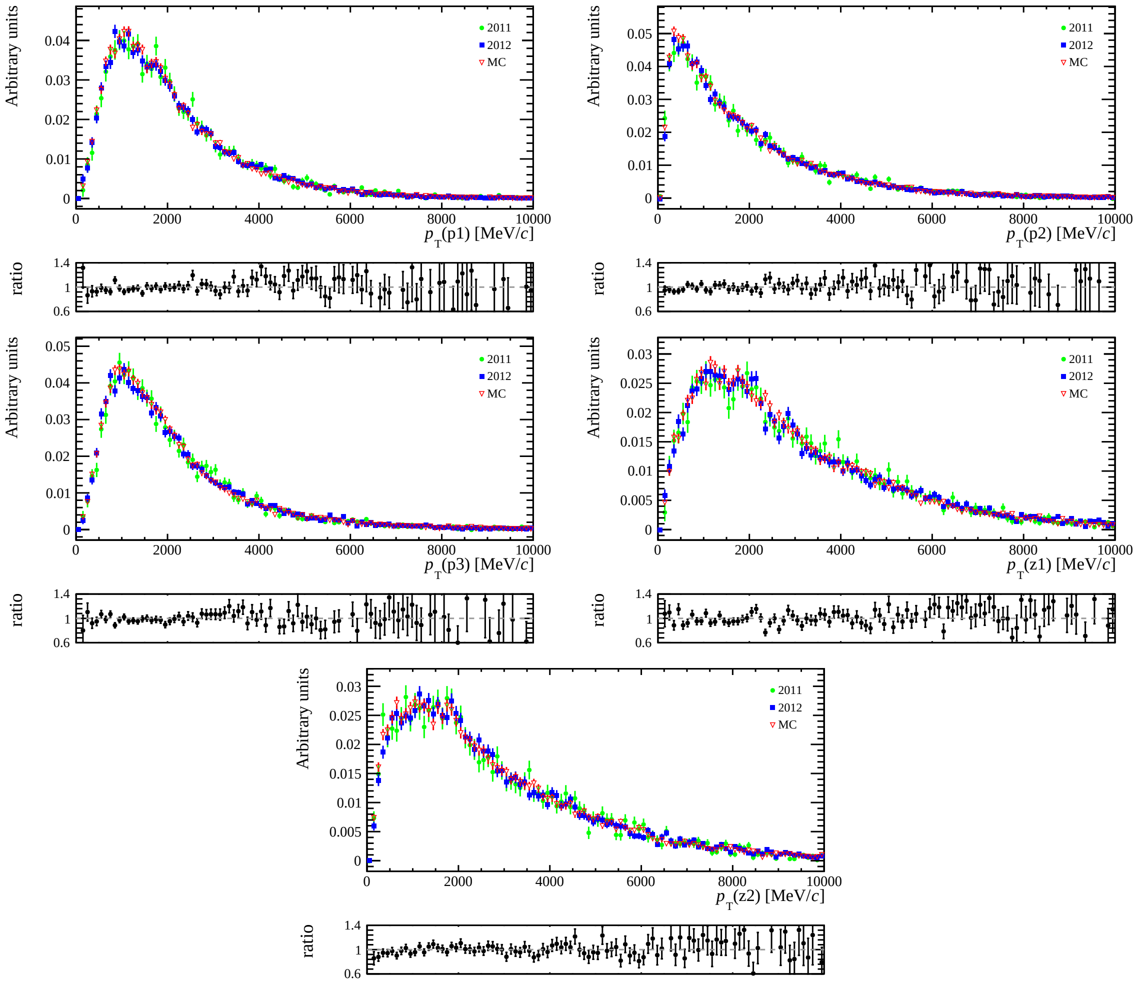


Figure A.4: Comparison of p_{T} distributions of the tracks between data and simulation with $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ decays. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D_s^+ respectively and similarly, the labels z1 and z2 stand for the K^+ and π^- daughters of $\overline{D}^0$, respectively.

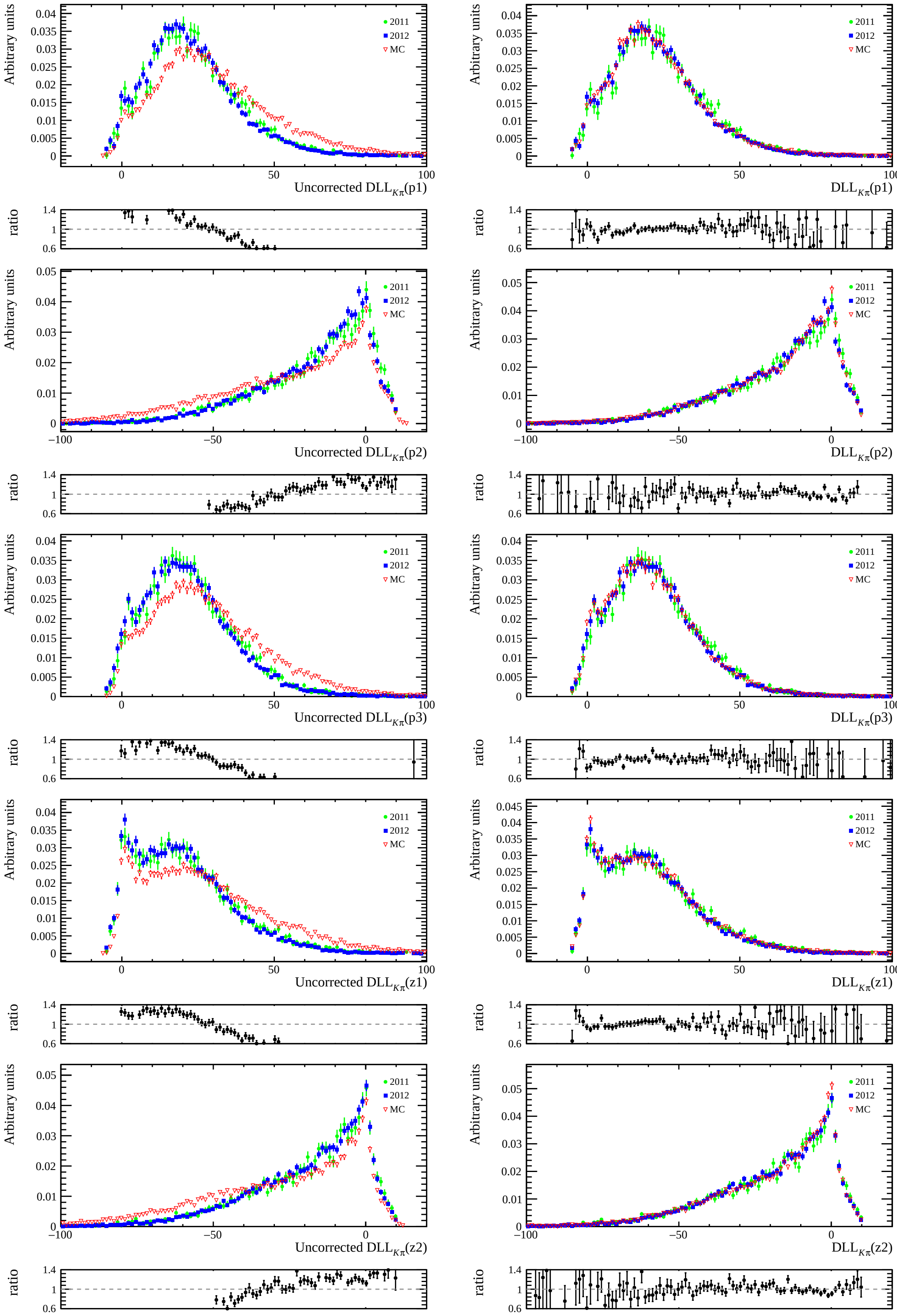


Figure A.5: Comparison of variables related to the particle identification between data and simulation with $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-$ decays before (left) and after (right) correction. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D_s^+ respectively and similarly, the labels z1 and z2 stand for the K^+ and π^- daughters of $\overline{D}^0$, respectively.

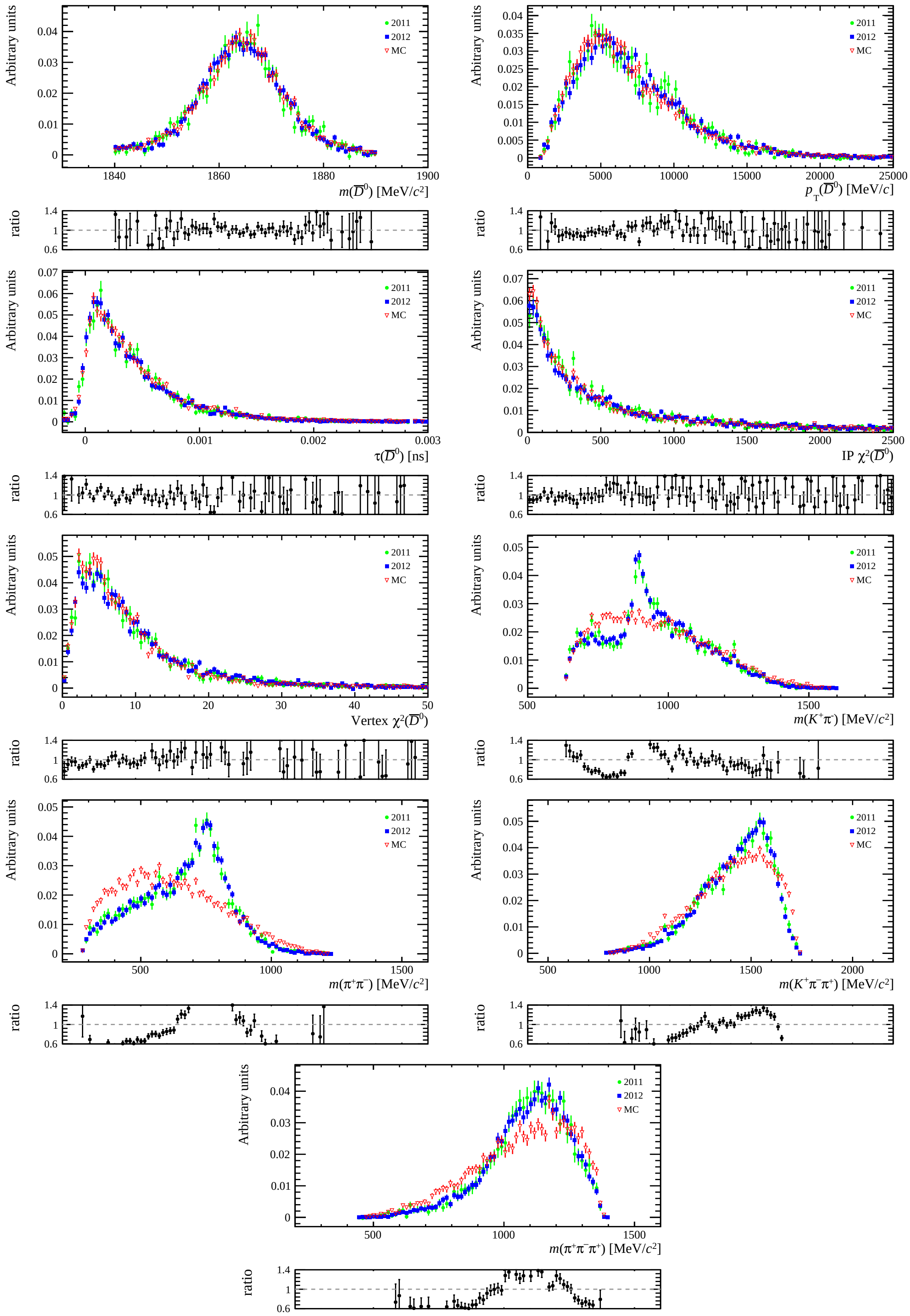


Figure A.6: Comparison of variables related to the reconstructed $\overline{D}^0$ between data and simulation with $B^+ \to D_s^+ \overline{D}^0$, $\overline{D}^0 \to K^+\pi^-\pi^+\pi^-$ decays.

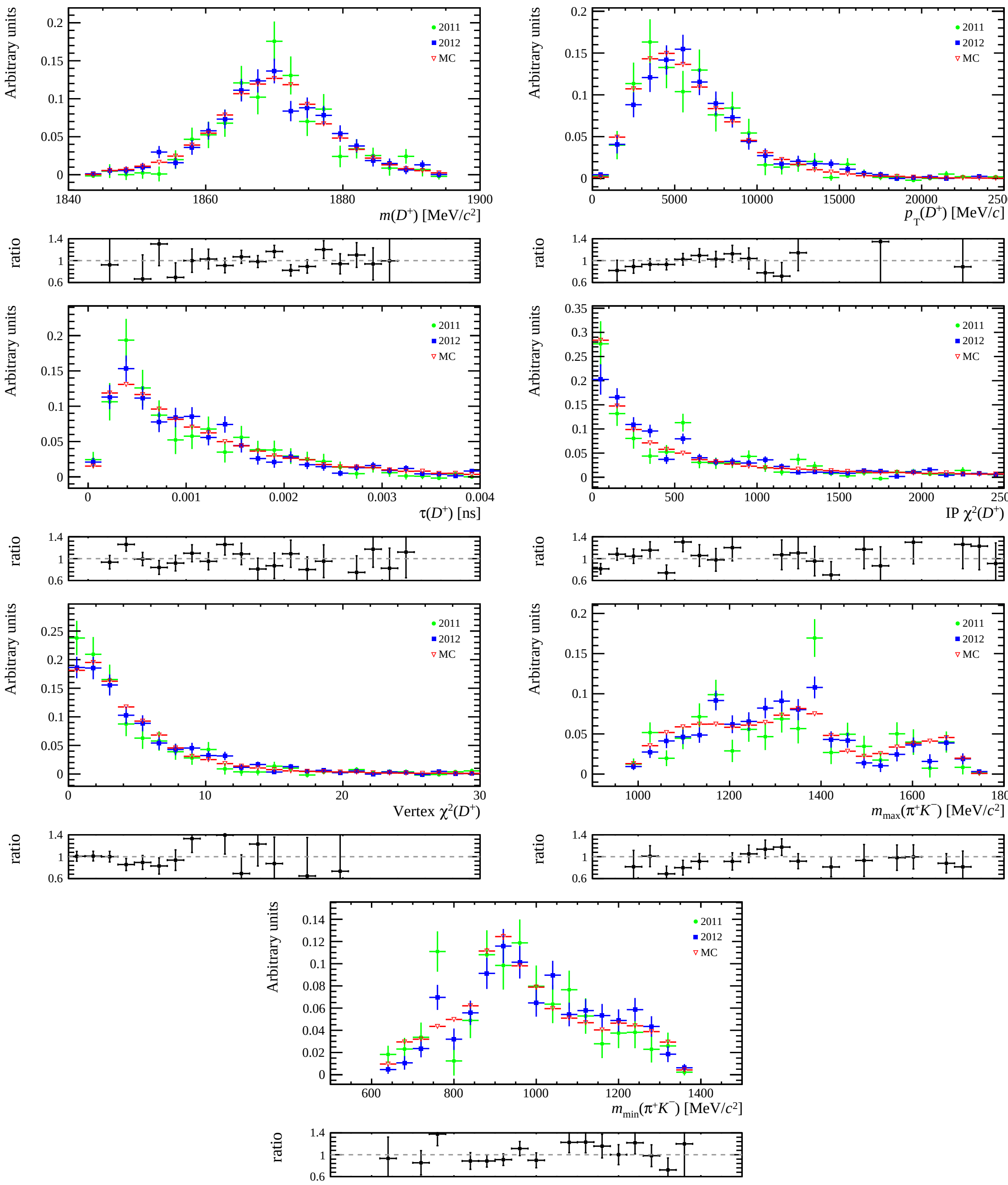


Figure A.7: Comparison of variables related to the reconstructed D^+ between data and simulation with $B^+\to D^+\overline{D}^0$, $\overline{D}^0\to K^+\pi^-$ decays.

Appendix B

MVA input variables

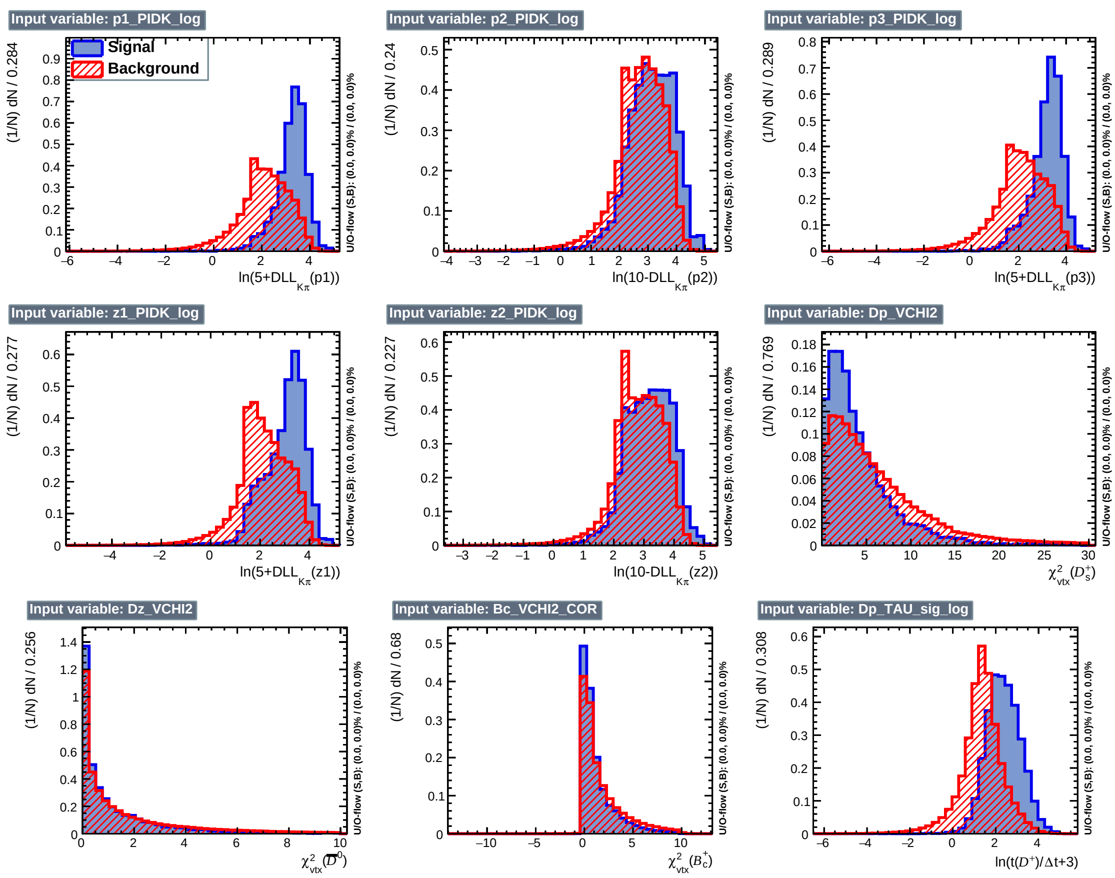


Figure B.1: Distributions of the input variables to the MVA for $B_c^+ \to D_s^+ \overline{D}^0$ candidates with $\overline{D}^0 \to K^+\pi^-$ decays. Blue is signal and red is background. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D_s^+ respectively and similarly, the labels z1 and z2 stand for the K^+ and π^- daughter of $\overline{D}^0$, respectively.

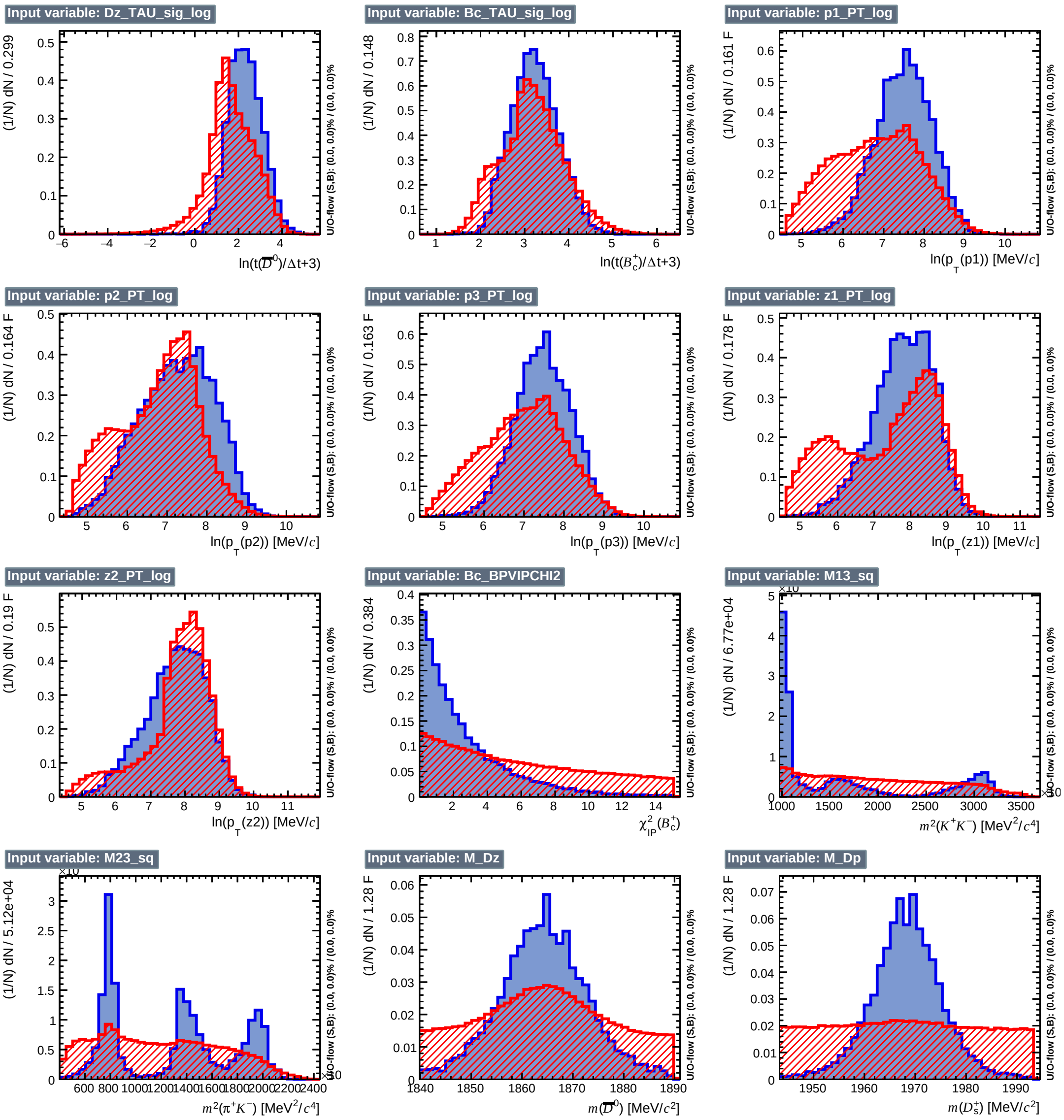


Figure B.2: Distributions of the input variables to the MVA for $B_c^+ \to D_s^+ \overline{D}^0$ candidates with $\overline{D}^0 \to K^+\pi^-$ decays. Blue is signal and red is background. The labels p1, p2 and p3 stand for the K^+, π^+ and K^- daughters of D_s^+ respectively and similarly, the labels z1 and z2 stand for the K^+ and π^- daughter of $\overline{D}^0$, respectively.

Appendix C

Fit results

C.1 B_c^+ search

Table C.1: Results of the fits to the constrained mass distributions of $B_c^+ \to D_{(s)}^{(*)+} D^{(*)}$ candidates.

Parameter	Reconstructed state			
	$D_s^+ \overline{D}^0$	$D_s^+ D^0$	$D^+ \overline{D}^0$	$D^+ D^0$
$\frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D_{(s)}^+ D)}{\mathcal{B}(B^+ \to D_{(s)}^+ \overline{D}^0)} (\times 10^{-3})$	0.30 ± 0.37	-0.38 ± 0.26	8.00 ± 7.54	2.86 ± 5.27
$\frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to (D_{(s)}^{*+} \to D_{(s)}^+ \pi^0/\gamma) D) + \mathcal{B}(B_c^+ \to D_{(s)}^+ D^*)}{\mathcal{B}(B^+ \to D_{(s)}^+ \overline{D}^0)} (\times 10^{-3})$	0.11 ± 1.55	-0.34 ± 1.84	2 ± 32	-15 ± 17
$\frac{f_c}{f_u} \frac{\mathcal{B}(B_c^+ \to D_{(s)}^{*+} D^*)}{\mathcal{B}(B^+ \to D_{(s)}^+ \overline{D}^0)} (\times 10^{-3})$	3.20 ± 4.30	7.05 ± 9.18	340 ± 230	-41 ± 91
$\sigma_{\text{yield}}^{\text{syst}}(B_c^+ \to D_{(s)}^+ D)$	0.00 ± 1.50	0.01 ± 2.19	0.00 ± 2.12	0.00 ± 0.87
$\sigma_{\text{yield}}^{\text{syst}}(B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*)$	-1.3 ± 13.7	0.0 ± 19.2	-0.1 ± 17.9	-0.02 ± 7.02
$\sigma_{\text{yield}}^{\text{syst}}(B_c^+ \to D_{(s)}^{*+} D^*)$	-4.3 ± 40.4	-2.6 ± 81.7	1.2 ± 38.1	0.00 ± 9.86
$N^{K\pi}(B^+)$	$21,340 \pm 150$	326 ± 21	$1,197 \pm 37$	23.7 ± 7.2
$N^{K\pi\pi\pi}(B^+)$	$12,390 \pm 110$	150 ± 16	669 ± 28	13.8 ± 7.9
$\mu^{K\pi}(B^+)$	5278.8 ± 0.1	—	5278.4 ± 0.3	—
$\mu^{K\pi\pi\pi}(B^+)$	5278.5 ± 0.1	—	5277.6 ± 0.5	—
$\sigma^{K\pi}(B^+)$	7.61 ± 0.05	—	7.31 ± 0.23	—
$\sigma^{K\pi\pi\pi}(B^+)$	7.43 ± 0.07	—	9.44 ± 0.4	—
$\sigma_{\text{norm}}^{\text{syst}}(B_c^+ \to D_{(s)}^+ D)$	$13,580 \pm 320$	$13,590 \pm 320$	795 ± 26	795 ± 26
$\sigma_{\text{norm}}^{\text{syst}}(B_c^+ \to D_{(s)}^{*+} D, D_{(s)}^+ D^*)$	$11,850 \pm 320$	$11,850 \pm 320$	683 ± 25	683 ± 24
$\sigma_{\text{norm}}^{\text{syst}}(B_c^+ \to D_{(s)}^{*+} D^*)$	$10,670 \pm 300$	$10,670 \pm 300$	197 ± 7	197 ± 7
$N_{\text{const}}^{K\pi}$	55 ± 68	226 ± 27	303 ± 85	5 ± 148
$N_{\text{const}}^{K\pi\pi\pi}$	206 ± 33	215 ± 26	182 ± 42	91 ± 40
$N_{\text{exp}}^{K\pi}$	621 ± 77	289 ± 26	541 ± 93	474 ± 15
$N_{\text{exp}}^{K\pi\pi\pi}$	793 ± 43	410 ± 31	431 ± 45	354 ± 44
$x^{K\pi} (\times 10^{-3})$	-1.92 ± 0.33	-6.14 ± 0.90	-2.34 ± 0.56	-1.17 ± 0.43
$x^{K\pi\pi\pi} (\times 10^{-3})$	-0.42 ± 0.40	-4.48 ± 0.48	-3.05 ± 0.49	-2.58 ± 0.45

C.2 *CP* asymmetry measurement

Table C.2: Results of the fits to the constrained mass distributions of $B^+ \to D^+_{(s)}\overline{D}^0$ candidates.

Channel	Parameter $N(B)$	$\mu(B)$	$\sigma(B)$	N_{exp}	x $(\times 10^{-3})$
$B^- \to D_s^- D^0,\ D^0 \to K^-\pi^+$	$13,659 \pm 129$	5279.1 ± 0.1	7.79 ± 0.07	$10,516 \pm 4,352$	-3.65 ± 2.01
$B^- \to D_s^- D^0,\ D^0 \to K^-\pi^+\pi^-\pi^+$	$7,717 \pm 103$	5278.8 ± 0.1	8.05 ± 0.11	$9,059 \pm 978$	-2.16 ± 1.69
$B^+ \to D_s^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-$	$14,209 \pm 132$	5278.9 ± 0.1	7.79 ± 0.07	$11,379 \pm 4,607$	-3.70 ± 1.92
$B^+ \to D_s^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	$7,945 \pm 104$	5278.6 ± 0.1	8.00 ± 0.11	$18,417 \pm 9,560$	-5.47 ± 1.78
$B^- \to D^- D^0,\ D^0 \to K^-\pi^+$	678 ± 32	5278.5 ± 0.4	7.77 ± 0.39	$13,322 \pm 16,192$	-9.13 ± 3.40
$B^- \to D^- D^0,\ D^0 \to K^-\pi^+\pi^-\pi^+$	369 ± 24	5278.6 ± 0.6	8.78 ± 0.57	$4,055 \pm 5,356$	-6.81 ± 4.04
$B^+ \to D^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-$	660 ± 31	5278.9 ± 0.4	7.51 ± 0.36	$3,035 \pm 2,910$	-0.53 ± 2.75
$B^+ \to D^+ \overline{D}^0,\ \overline{D}^0 \to K^+\pi^-\pi^+\pi^-$	345 ± 24	5277.1 ± 0.7	9.11 ± 0.65	$1,622 \pm 3,518$	-0.13 ± 4.57

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Search for B_c^+ decays to two charm mesons*, Nucl. Phys. **B930** (2018) 563, `arXiv:1712.04702`.

[2] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in $B^+ \to D_s^+ \overline{D}^0$ and $B^+ \to D^+ \overline{D}^0$ decays*, JHEP **05** (2018) 060, `arXiv:1803.10990`.

[3] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[4] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, `arXiv:1207.7214`.

[5] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, `arXiv:1207.7235`.

[6] D. Hanneke, S. F. Hoogerheide, and G. Gabrielse, *Cavity control of a single-electron quantum cyclotron: measuring the electron magnetic moment*, Phys. Rev. **A83** (2011) 052122, `arXiv:1009.4831`.

[7] Planck collaboration, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, `arXiv:1807.06209`.

[8] P. J. E. Peebles and B. Ratra, *The cosmological constant and dark energy*, Rev. Mod. Phys. **75** (2003) 559, `arXiv:astro-ph/0207347`.

[9] M. P. Hobson, G. P. Efstathiou, and A. N. Lasenby, *General Relativity: An introduction for physicists*, Cambridge University Press, 2006.

[10] 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. Nauk 161, no.5, 61 (1991)].

[11] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pène, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795, `arXiv:hep-ph/9312215`.

[12] M. Masetti, *CP violation in B_c decays*, Phys. Lett. **B286** (1992) 160.

[13] R. Fleischer and D. Wyler, *Exploring CP violation with B_c decays*, Phys. Rev. D **62** (2000) 057503.

[14] A. K. Giri, R. Mohanta, and M. P. Khanna, *Determination of the angle γ from nonleptonic $B_c \to D_s D^0$ decays*, Phys. Rev. **D65** (2002) 034016, `arXiv:hep-ph/0104009`.

[15] A. K. Giri, B. Mawlong, and R. Mohanta, *Determining the CKM angle gamma with B_c decays*, Phys. Rev. **D75** (2007) 097304, `arXiv:hep-ph/0611212`, [Erratum: Phys. Rev. D76, 099902 (2007)].

[16] M. A. Ivanov, J. G. Korner, and O. N. Pakhomova, *The nonleptonic decays $B_c^+ \to D_s^+ \bar{D}^0$ and $B_c^+ \to D_s^+ D^0$ in a relativistic quark model*, Phys. Lett. **B555** (2003) 189, `arXiv:hep-ph/0212291`.

[17] V. V. Kiselev, *Gold plated mode of CP violation in decays of B_c meson from QCD sum rules*, J. Phys. **G30** (2004) 1445, `arXiv:hep-ph/0302241`.

[18] C. S. Kim, R.-M. Wang, and Y.-D. Yang, *Studying double charm decays of $B_{u,d}$ and B_s mesons in the MSSM with R-parity violation*, Phys. Rev. **D79** (2009) 055004, `arXiv:0812.4136`.

[19] Y.-G. Xu and R.-M. Wang, *Studying the fourth generation quark contributions to the double charm decays $B_{(s)} \to D_{(s)}^{(*)} D_s^{(*)}$*, Int. J. Theor. Phys. **55** (2016) 5290.

[20] R. Zhou, Z. Zou, and C.-D. Lü, *Double charm decays of B_c meson in the perturbative QCD approach*, Phys. Rev. D **86** (2012) 074019.

[21] M. A. Ivanov, J. G. Körner, and P. Santorelli, *Exclusive semileptonic and nonleptonic decays of the B_c meson*, Phys. Rev. D **73** (2006) 054024.

[22] P. A. M. Dirac and R. H. Fowler, *The quantum theory of the electron*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **117** (1928) 610.

[23] P. A. M. Dirac and R. H. Fowler, *Quantum mechanics of many-electron systems*, Proceedings of the Royal Society of London. Series A, Containing Papers of a

Mathematical and Physical Character **123** (1929) 714.

[24] P. A. M. Dirac, *Quantised singularities in the electromagnetic field*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **133** (1931) 60.

[25] C. D. Anderson, *The positive electron*, Phys. Rev. **43** (1933) 491.

[26] R. P. Feynman, *The theory of positrons*, Phys. Rev. **76** (1949) 749.

[27] ALPHA collaboration, M. Ahmadi *et al.*, *Observation of the 1S-2S transition in trapped antihydrogen*, Nature **541** (2016) 506.

[28] A. G. Cohen, A. De Rujula, and S. L. Glashow, *A matter - antimatter universe?*, Astrophys. J. **495** (1998) 539, `arXiv:astro-ph/9707087`.

[29] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and antimatter in the universe*, New J. Phys. **14** (2012) 095012, `arXiv:1204.4186`.

[30] G. Krnjaic, *Can the baryon asymmetry arise from initial conditions?*, Phys. Rev. **D96** (2017) 035041, `arXiv:1606.05344`.

[31] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, *On the anomalous electroweak baryon number nonconservation in the early universe*, Phys. Lett. **155B** (1985) 36.

[32] G. 't Hooft, *Symmetry breaking through bell-jackiw anomalies*, Phys. Rev. Lett. **37** (1976) 8.

[33] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *decay of the* K_2^0 *meson*, Phys. Rev. Lett. **13** (1964) 138.

[34] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in charm decays*, `arXiv:1903.08726`, submitted to Phys. Rev. Lett.

[35] LHCb collaboration, R. Aaij *et al.*, *Measurement of matter-antimatter differences in beauty baryon decays*, Nature Physics **13** (2017) 391, `arXiv:1609.05216`.

[36] M. Dine, *TASI lectures on the strong CP problem*, in *Flavor physics for the millennium. Proceedings, Theoretical Advanced Study Institute in elementary particle physics, TASI 2000, Boulder, USA, June 4-30, 2000*, pp. 349–369, 2000, `arXiv:hep-ph/0011376`.

[37] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[38] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[39] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[40] A. Banerjee, *Electroweak phase transition in the early universe and baryogenesis*, 2011.

[41] P. B. Arnold and O. Espinosa, *The effective potential and first order phase transitions: beyond leading-order*, Phys. Rev. **D47** (1993) 3546, `arXiv:hep-ph/9212235`, [Erratum: Phys. Rev. D 50, 6662 (1994)].

[42] R. Jora, *Baryon asymmetry in the standard model revisited*, `arXiv:1806.00597`.

[43] Super-Kamiokande collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, `arXiv:hep-ex/9807003`.

[44] B. Pontecorvo, *Neutrino experiments and the problem of conservation of leptonic charge*, Sov. Phys. JETP **26** (1968) 984, [Zh. Eksp. Teor. Fiz. 53, 1717 (1967)].

[45] B. Pontecorvo, *Mesonium and anti-mesonium*, Sov. Phys. JETP **6** (1957) 429, [Zh. Eksp. Teor. Fiz. 33, 549 (1957)].

[46] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Progress of Theoretical Physics **28** (1962) 870.

[47] DUNE collaboration, B. Abi *et al.*, *The DUNE far detector interim design report volume 1: physics, technology and strategies*, `arXiv:1807.10334`.

[48] Hyper-Kamiokande collaboration, K. Abe *et al.*, *Hyper-Kamiokande Design Report*, `arXiv:1805.04163`.

[49] P. Di Bari, *An introduction to leptogenesis and neutrino properties*, Contemp. Phys. **53** (2012) 315, `arXiv:1206.3168`.

[50] J. M. Cline, *Baryogenesis*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*, 2006, `arXiv:hep-ph/0609145`.

[51] J. M. Cline, M. Joyce, and K. Kainulainen, *Supersymmetric electroweak baryogenesis*, JHEP **07** (2000) 018, `arXiv:hep-ph/0006119`.

[52] L.-L. Chau and W.-Y. Keung, *Comments on the parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[53] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[54] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`.

[55] J. Brod, A. Lenz, G. Tetlalmatzi-Xolocotzi, and M. Wiebusch, *New physics effects in tree-level decays and the precision in the determination of the quark mixing angle* γ, Phys. Rev. **D92** (2015) 033002, `arXiv:1412.1446`.

[56] J. Brod and J. Zupan, *The ultimate theoretical error on* γ *from* $B \to DK$ *decays*, JHEP **01** (2014) 051, `arXiv:1308.5663`.

[57] LHCb collaboration, *Update of the LHCb combination of the CKM angle* γ *using* $B \to DK$ *decays*, LHCb-CONF-2018-002, 2018.

[58] M. Rama, *Effect of* $D-\bar{D}$ *mixing in the extraction of* γ *with* $B^- \to D^0K^-$ *and* $B^- \to D^0\pi^-$ *decays*, Phys. Rev. **D89** (2014) 014021, `arXiv:1307.4384`.

[59] M. Gronau and D. Wyler, *On determining a weak phase from CP asymmetries in charged B decays*, Phys. Lett. **B265** (1991) 172.

[60] M. Gronau and D. London, *How to determine all the angles of the unitarity triangle from* $B^0_d \to DK_S$ *and* $B^0_s \to D\phi$, Phys. Lett. **B253** (1991) 483.

[61] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing* CP *violation in* $B^\pm \to \mathrm{KD}$ *and measuring the CKM phase* γ, Phys. Rev. D **63** (2001) 036005.

[62] D. Atwood, I. Dunietz, and A. Soni, *Enhanced CP violation with* $B \to KD^0(\overline{D}^0)$ *modes and extraction of the Cabibbo-Kobayashi-Maskawa angle* γ, Phys. Rev. Lett. **78** (1997) 3257.

[63] D. Atwood and A. Soni, *Role of a charm factory in extracting CKM-phase information via* $B \to DK$, Phys. Rev. D **68** (2003) 033003.

[64] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining* γ *using* $B^\pm \to \mathrm{DK}^\pm$

with multibody D decays, Phys. Rev. D **68** (2003) 054018.

[65] CLEO, Y. Kubota *et al.*, *The CLEO-II detector*, Nucl. Instrum. Meth. **A320** (1992) 66.

[66] D. M. Asner *et al.*, *Physics at BES-III*, Int. J. Mod. Phys. **A24** (2009) S1, `arXiv:0809.1869`.

[67] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ using $B^\pm \to DK^\pm$ with $D \to K^0_S\pi^+\pi^-$, $K^0_S K^+K^-$ decays*, JHEP **10** (2014) 097, `arXiv:1408.2748`.

[68] BaBar collaboration, B. Aubert *et al.*, *Measurement of γ in $B^\mp \to D^{(*)}K^\mp$ decays with a Dalitz analysis of $D \to K^0_S\pi^-\pi^+$*, Phys. Rev. Lett. **95** (2005) 121802, `arXiv:hep-ex/0504039`.

[69] LHCb collaboration, R. Aaij *et al.*, *Measurement of B_c^+ production in proton-proton collisions at $\sqrt{s}$ =8 TeV*, Phys. Rev. Lett. **114** (2015) 132001, `arXiv:1411.2943`.

[70] LHCb collaboration, R. Aaij *et al.*, *Measurements of B_c^+ production and mass with the $B_c^+ \to J/\psi\pi^+$ decay*, Phys. Rev. Lett. **109** (2012) 232001, `arXiv:1209.5634`.

[71] D. Ebert, R. N. Faustov, and V. O. Galkin, *Weak decays of the B_c meson to charmonium and D mesons in the relativistic quark model*, Phys. Rev. D **68** (2003) 094020.

[72] C.-F. Qiao, P. Sun, D. Yang, and R.-L. Zhu, *B_c exclusive decays to charmonium and a light meson at next-to-leading order accuracy*, Phys. Rev. D **89** (2014) 034008.

[73] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016, 2014.

[74] LHCb collaboration, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, `arXiv:1808.08865`.

[75] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP observables in $B^\pm \to D^{(*)}K^\pm$ and $B^\pm \to D^{(*)}\pi^\pm$ decays*, Phys. Lett. **B777** (2017) 16, `arXiv:1708.06370`.

[76] L. Bel *et al.*, *Anatomy of $B \to D\overline{D}$ decays*, JHEP **07** (2015) 108, `arXiv:1505.01361`.

[77] D. Sahoo *et al.*, *Prediction of the CP asymmetry C_{00} in $B^0 \to D^0\overline{D^0}$ decay*, JHEP

11 (2017) 087, `arXiv:1709.08301`.

[78] R.-H. Li, X.-X. Wang, A. I. Sanda, and C.-D. Lu, *Decays of B meson to two charmed mesons*, Phys. Rev. **D81** (2010) 034006, `arXiv:0910.1424`.

[79] H.-F. Fu, G.-L. Wang, Z.-H. Wang, and X.-J. Chen, *Semi-leptonic and non-leptonic B meson decays to charmed mesons*, Chin. Phys. Lett. **28** (2011) 121301, `arXiv:1202.1221`.

[80] L.-X. Lu, Z.-J. Xiao, S.-W. Wang, and W.-J. Li, *The double charm decays of B mesons in the mSUGRA model*, Commun. Theor. Phys. **56** (2011) 125, `arXiv:1008.4987`.

[81] M. Jung and S. Schacht, *Standard model predictions and new physics sensitivity in $B \to DD$ decays*, Phys. Rev. **D91** (2015) 034027, `arXiv:1410.8396`.

[82] A. Abashian *et al.*, *The Belle detector*, Nucl. Instrum. Meth. **A479** (2002) 117.

[83] BaBar, D. Boutigny *et al.*, *BaBar Technical Design Report*, in *BaBar Technical Design Report EPAC Meeting Stanford, California, March 17-18, 1995*, 1995.

[84] Belle collaboration, I. Adachi *et al.*, *Measurement of the branching fraction and charge asymmetry of the decay $B^+ \to D^+\overline{D}^0$ and search for $B^0 \to D^0\overline{D}^0$*, Phys. Rev. **D77** (2008) 091101, `arXiv:0802.2988`.

[85] BaBar collaboration, B. Aubert *et al.*, *Measurement of branching fractions and CP-violating charge asymmetries for B-meson decays to $D^{(*)}\overline{D}^{(*)}$, and implications for the Cabibbo-Kobayashi-Maskawa angle γ*, Phys. Rev. **D73** (2006) 112004, `arXiv:hep-ex/0604037`.

[86] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[87] *Convention for the establishment of a European organization for nuclear research: Paris, 1st July, 1953 : as amended. Convention pour l'établissement d'une Organisation européenne pour la Recherche nucléaire. Paris, le 1er juillet 1953 : telle qu'elle a été modifiée*, CERN, Geneva, 1971.

[88] *CERN annual personnel statistics*, 2017.

[89] F. J. Hasert *et al.*, *Search for elastic ν_μ electron scattering*, Phys. Lett. **B46** (1973) 121.

[90] F. J. Hasert *et al.*, *Observation of neutrino like interactions without muon or electron in the Gargamelle neutrino experiment*, Phys. Lett. **B46** (1973) 138.

[91] UA1 collaboration, G. Arnison *et al.*, *Experimental observation of isolated large transverse energy electrons with associated missing energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. **B122** (1983) 103.

[92] UA2 collaboration, M. Banner *et al.*, *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN* $\bar{p}p$ *collider*, Phys. Lett. **B122** (1983) 476.

[93] UA1 collaboration, G. Arnison *et al.*, *Experimental observation of lepton pairs of invariant mass around 95 GeV/*c^2 *at the CERN SPS collider*, Phys. Lett. **B126** (1983) 398.

[94] UA2 collaboration, P. Bagnaia *et al.*, *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\bar{p}p$ *collider*, Phys. Lett. **B129** (1983) 130.

[95] ALEPH collaboration, D. Decamp *et al.*, *Determination of the number of light neutrino species*, Phys. Lett. **B231** (1989) 519.

[96] DELPHI collaboration, P. Aarnio *et al.*, *Measurement of the mass and width of the* Z^0 *particle from multihadronic final states produced in* e^+e^- *annihilations*, Phys. Lett. **B231** (1989) 539.

[97] L3 collaboration, B. Adeva *et al.*, *A determination of the properties of the neutral intermediate vector boson* Z^0, Phys. Lett. **B231** (1989) 509.

[98] OPAL collaboration, M. Z. Akrawy *et al.*, *Measurement of the* Z^0 *mass and width with the OPAL detector at LEP*, Phys. Lett. **B231** (1989) 530.

[99] H. Burkhardt *et al.*, *First evidence for direct CP violation*, Physics Letters B **206** (1988) 169 .

[100] NA48 collaboration, V. Fanti *et al.*, *A new measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. **B465** (1999) 335, `arXiv:hep-ex/9909022`.

[101] LHCb collaboration, R. Aaij *et al.*, *Observation of* $J/\psi\, p$ *resonances consistent with pentaquark states in* $\Lambda_b^0 \to J/\psi\, pK^-$ *decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[102] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004.

[103] S. Holmes, R. S. Moore, and V. Shiltsev, *Overview of the Tevatron collider complex: goals, operations and performance*, JINST **6** (2011) T08001, `arXiv:1106.0909`.

[104] ATLAS collaboration, G. Aad *et al.*, *The ATLAS experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[105] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[106] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[107] W. Herr and T. Pieloni, *Beam-beam effects*, in *CAS - CERN Accelerator School: Advanced Accelerator Physics Course: Trondheim, Norway, August 18-29, 2013*, 2014, `arXiv:1601.05235`. doi: 10.5170/CERN-2014-009.431.

[108] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, 2013. General Photo.

[109] D. Boussard and T. P. R. Linnecar, *The LHC superconducting RF system*, Tech. Rep. LHC-Project-Report-316. CERN-LHC-Project-Report-316, CERN, Geneva, 1999.

[110] M. Benedikt *et al.*, *Future Circular Collider*, Tech. Rep. CERN-ACC-2018-0058, CERN, Geneva, 2018. Submitted for publication to Eur. Phys. J. ST.

[111] M. Bajko *et al.*, *Report of the task force on the incident of 19th September 2008 at the LHC*, Tech. Rep. LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168, CERN, Geneva, 2009.

[112] *Luminosity: Taking a close look at LHC*, `https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.luminosity`. Accessed January 2019.

[113] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, in *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pp. 183–187, 2014, `arXiv:1410.3667`. doi: 10.5170/CERN-2014-004.183.

[114] *CERN accelerator performance and statistics*, `http://acc-stats.web.cern.ch/`

acc-stats/#lhc/overview-panel. Accessed January 2019.

[115] *LHCb operations plots webpage*, https://lbgroups.cern.ch/online/OperationsPlots/index.htm. Accessed January 2019.

[116] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, 2003.

[117] LHCb collaboration, *LHCb: letter of intent*, Tech. Rep. CERN-LHCC-95-5. LHCC-I-8, CERN, Geneva, 1995.

[118] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[119] LHCb collaboration, B. Adeva *et al.*, *Roadmap for selected key measurements of LHCb*, arXiv:0912.4179.

[120] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[121] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[122] *Tracking and alignment plots for conferences*, https://twiki.cern.ch/twiki/bin/view/LHCb/ConferencePlots. Accessed Feburary 2019.

[123] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007, 2000.

[124] M. Vesterinen, *Considerations on the LHCb dipole magnet polarity reversal*, Tech. Rep. LHCb-PUB-2014-006. CERN-LHCb-PUB-2014-006, CERN, Geneva, 2014.

[125] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002.

[126] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001.

[127] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[128] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **9** (2017) P11016, arXiv:1708.00819.

[129] *LHCb Silicon Tracker - material for publications*, `https://www.physik.uzh.ch/groups/lhcb/public/material/`. Accessed Feburary 2019.

[130] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, `arXiv:physics/0503191`.

[131] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[132] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037, 2000.

[133] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[134] A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, Nucl. Instrum. Meth. **A876** (2017) 221, `arXiv:1703.08152`.

[135] P. A. Cherenkov, *Visible emission of clean liquids by action of γ radiation*, Doklady Akademii Nauk SSSR **2** (1934) 451.

[136] R. W. Forty and O. Schneider, *RICH pattern recognition*, Tech. Rep. LHCb-98-040, CERN, Geneva, 1998.

[137] *PID plots for conference*, `https://twiki.cern.ch/twiki/bin/view/LHCb/PIDConferencePlots`. Accessed Feburary 2019.

[138] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000.

[139] C. Lippmann, *Particle identification*, Nucl. Instrum. Meth. **A666** (2012) 148, `arXiv:1101.3276`.

[140] A. Arefev *et al.*, *Beam test results of the LHCb Electromagnetic Calorimeter.*, Tech. Rep. LHCb-2007-149. CERN-LHCb-2007-149, CERN, Geneva, 2008.

[141] C. Coca *et al.*, *The Hadron Calorimeter prototype beam-test results*, Tech. Rep. LHCb-2000-036, CERN, Geneva, 2000.

[142] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001.

[143] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-

2003-031, 2003.

[144] T. Head, *The LHCb trigger system*, JINST **9** (2014) C09015.

[145] J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, `arXiv:1310.8544`.

[146] LHCb collaboration, *Trigger schemes*, `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html`. Accessed Feburary 2019.

[147] V. V. Gligorov, *A single track HLT1 trigger*, Tech. Rep. LHCb-PUB-2011-003. CERN-LHCb-PUB-2011-003. LHCb-INT-2010-053, CERN, Geneva, 2011.

[148] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016. LHCb-INT-2011-030, CERN, Geneva, 2011. LHCb-INT-2011-030.

[149] M. Williams *et al.*, *The HLT2 topological lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, 2011.

[150] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019, 2005.

[151] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.

[152] C. Eck *et al.*, *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, Technical Design Report LCG, CERN, Geneva, 2005.

[153] *The LHCb data flow*, `https://lhcb.github.io/starterkit-lessons/first-analysis-steps/dataflow.html`. Accessed Feburary 2019.

[154] I. Belyaev *et al.*, *Simulation application for the LHCb experiment*, eConf C **0303241** (2003) TUMT003. 6 p.

[155] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[156] C.-H. Chang, C. Driouichi, P. Eerola, and X. G. Wu, *BCVEGPY: An event generator for hadronic production of the B_c meson*, Comput. Phys. Commun. **159** (2004) 192, `arXiv:hep-ph/0309120`.

[157] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth.

A462 (2001) 152.

[158] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, The European Physical Journal C - Particles and Fields **45** (2006) 97.

[159] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[160] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[161] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, Journal of Instrumentation **8** (2013) P02013.

[162] LHCb collaboration, R. Aaij *et al.*, *Measurements of the Λ_b^0, Ξ_b^-, and Ω_b^- baryon masses*, Phys. Rev. Lett. **110** (2013) 182001, `arXiv:1302.1072`.

[163] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, `arXiv:1803.00824`.

[164] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[165] LHCb collaboration, R. Aaij *et al.*, *Study of beauty hadron decays into pairs of charm hadrons*, Phys. Rev. Lett. **112** (2014) 202001, `arXiv:1403.3606`.

[166] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040.

[167] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: Bounds for K-fold and progressive cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), pp. 203–208, ACM, 1999. doi: 10.1145/307400.307439.

[168] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[169] M. Karson, *Handbook of methods of applied statistics. Volume I: techniques of computation descriptive methods, and statistical inference. Volume II: planning of surveys and experiments. I. M. Chakravarti, R. G. Laha, and J. Roy, New York, John Wiley*, Journal of the American Statistical Association **63** (1968) 1047.

[170] G. Punzi, *Sensitivity of searches for new signals and its optimization*, eConf **C030908** (2003) MODT002, `arXiv:physics/0308063`.

[171] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[172] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, `arXiv:physics/0306116`.

[173] F. James and M. Roos, *MINUIT: A system for function minimization and analysis of the parameter errors and corrections*, Comput. Phys. Commun. **10** (1975) 343.

[174] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

[175] LHCb collaboration, R. Aaij *et al.*, *Observation of $B_c^+ \to J/\psi\, D_s^+$ and $B_c^+ \to J/\psi\, D_s^{*+}$ decays*, Phys. Rev. **D87** (2013) 112012, `arXiv:1304.4530`.

[176] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, `arXiv:1007.1727`, [Erratum: Eur. Phys. J. C73, 2501 (2013)].

[177] L. Moneta *et al.*, *The RooStats project*, PoS **ACAT2010** (2010) 057, `arXiv:1009.1003`.

[178] LHCb collaboration, R. Aaij *et al.*, *Measurement of $D_s^{\pm}$ production asymmetry in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2018) 008, `arXiv:1805.09869`.

[179] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^{\pm}$ production asymmetry and the CP asymmetry in $B^{\pm} \to J/\psi\, K^{\pm}$ decays*, Phys. Rev. **D95** (2017) 052005, `arXiv:1701.05501`.

[180] A. Martin Sanchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 calorimeter trigger*, Tech. Rep. LHCb-PUB-2011-026. CERN-LHCb-PUB-2011-026, CERN, Geneva, 2012.

[181] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in B_s^0–$\overline{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, `arXiv:1605.09768`.

[182] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, 2016.

[183] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, JHEP **07** (2014) 041, `arXiv:1405.2797`.

[184] Belle II collaboration, T. Abe *et al.*, *Belle II Technical Design Report*, `arXiv:1011.0352`.

[185] Belle II collaboration, W. Altmannshofer *et al.*, *The Belle II Physics Book*, `arXiv:1808.10567`.

[186] J. Albrecht *et al.*, *Future prospects for exploring present day anomalies in flavour physics measurements with Belle II and LHCb*, `arXiv:1709.10308`.

List of figures

List of tables