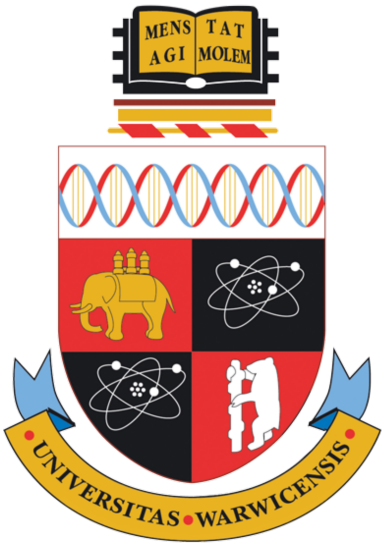

Measurement of the Relative Branching Fractions of Charmless Three-Body B^+ Decays at LHCb

by

Cayo Costa Sobral

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

September 2019

Contents

> *I got home and I thought I should stop leading so aimless an existence. It is harder than you might think to stop leading an existence, & if you can't do that the only thing you can do is try to introduce an element of purposefulness.... and though I might have to wait another 30 or 40 years for my body to join the non-sentient things in the world at least in the meantime it would be a less absolutely senseless sentience.*

Helen DeWitt, *The Last Samurai*

Acknowledgments

I would first like to thank my supervisors Tim Gershon and Tom Latham, for four years of guidance, wisdom, and inexhaustible patience without which this work would not have been possible. Additionally, this thesis has been much improved due to your insightful comments. I would also like to thank Michal Kreps for providing my first opportunity to work with LHCb as a summer student, and Tom Blake for always been happy to answer my questions.

Abhijit, thanks for helping me out on countless occasions and the frequent Q&A sessions late in the day; Ed, for often being on the same wavelength as me in so many aspects of the PhD life; and other members of the Warwick LHCb and EPP groups, for making this corner of the department fun to work in.

Mum and Renato, there will never be enough thanks for all the sacrifices you have made to get me to this point. Vó, obrigado por estabelecer o meu gosto pelo conhecimento e pelo contínuo encorajamento. To all my family, thanks for always cheering me on.

To the old Llewellyn Road gang and associates, and also Bethan, thanks for providing plenty of good reasons to leave the Warwick bubble. Many thanks to the pub Wednesday lot for the often needed midweek breaks. To those of you lucky enough to be part of both groups, I can only apologise for having to deal with me for those extra years. Dan and Ciaran, thanks for the many lovely meals shared all around Leamington. Jess and Sian, thanks for always lending an ear and dealing with my ramblings and annoyances, often on a daily basis; I can only hope I have been as effective a listener.

Cayo Costa Sobral

Declarations

The work presented in this thesis is all of my own work, unless it is specifically referenced to the contrary. This thesis has not been submitted, in any form, to this or any other university for another qualification.

Cayo Costa Sobral

December 6, 2019

Abstract

This thesis documents a measurement of the relative branching fractions of the charmless three-body decays $B^+ \to K^+K^+K^-$, $B^+ \to K^+\pi^+K^-$, $B^+ \to K^+\pi^+\pi^-$, and $B^+ \to \pi^+\pi^+\pi^-$, performed using the dataset collected by the LHCb experiment in the years 2011 and 2012 and corresponding to a total integrated luminosity of $3\,\text{fb}^{-1}$.

The branching fractions relative to that of the $B^+ \to K^+K^+K^-$ decay are determined to be

$$\begin{aligned}
\mathcal{B}(B^+ \to K^+\pi^+K^-) \,/\, \mathcal{B}(B^+ \to K^+K^+K^-) &= 0.1521 \pm 0.0040\,(\text{stat}) \pm 0.0087\,(\text{syst})\,, \\
\mathcal{B}(B^+ \to K^+\pi^+\pi^-) \,/\, \mathcal{B}(B^+ \to K^+K^+K^-) &= 1.714 \;\; \pm 0.011\,(\text{stat}) \;\; \pm 0.061\,(\text{syst})\,, \\
\mathcal{B}(B^+ \to \pi^+\pi^+\pi^-) \,/\, \mathcal{B}(B^+ \to K^+K^+K^-) &= 0.501 \;\; \pm 0.005\,(\text{stat}) \;\; \pm 0.015\,(\text{syst})\,.
\end{aligned}$$

The branching fractions relative to the other decay channels are also determined. These measurements show a significant improvement in precision over the current world average values.

CHAPTER 1

Introduction

> *When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science, whatever the matter may be.*
>
> Lord Kelvin, *Electrical Units of Measurement*

The development of particle physics as a field of research is undoubtedly coupled with the rapid technological developments of the 20th century. The discovery of new particles has repeatedly followed an increase in the energies accessible by particle accelerators. Similarly, it is difficult to imagine how the ever-larger datasets required to observe rare processes could be analysed if all that were available were photographic plates, such as the ones used by Anderson in the discovery of the positron [1].

Although particle physics is heavily associated with modern technology, the fundamental questions it attempts to answer have a long history. The ancient Greeks are recognised as the first people to put forward theories of matter and its behaviour without invoking supernatural causes [2]. It is a testament to their collective ingenuity that one can look back, more than two thousand years ago, and see resemblance in some of their writings to the basic ideas of modern theory. Leucippus and his more famous student Democritus, in the fifth century BCE, introduced the idea of the World being composed of indivisible atoms and the void, in which atoms moved.

Their atoms were supposedly infinite in number and type, immutable and impenetrable, and in constant random motion. Interestingly, the idea that the atoms must necessarily be microscopic in size does not seem to originate with Democritus, but instead with Epicurus, in a later refinement of the theory [3].

Philosophical atomism would resurface occasionally through history but it would not be until the 19th century, with developments in chemistry and kinetic theory, that the idea of the atom would gain scientific significance. It would still take a century of theory-building and experimental observation for atomic theory to cement itself, with Einstein's theory of Brownian motion and its experimental verification by Jean Perrin [4, 5] ultimately proving its validity.

Democritus's victory over his contemporaries would never prove itself absolute: J. J. Thomson's discovery of the electron [6] revealed the existence of particles with masses much smaller than the lightest atom, Rutherford's scattering experiment [7] proved the atom had substructure, and experiments into radioactive decay showed atoms are not immutable. However, not all is lost for the Atomists as not only does the electron appear to be an indivisible, fundamental particle but so do other particles, such as quarks.

The current-best theory that describes the behaviour of electrons, quarks, and the other fundamental particles of the Universe is called the Standard Model (SM) of particle physics. The SM is the culmination of almost a century of theoretical development and has shown itself to be both fantastically predictive and accurate throughout the years. This is best exemplified by achievements such as the SM prediction for the electron anomalous magnetic moment, which matches experiment to ten significant figures [8, 9], and the discoveries of many particles that were first predicted by the SM, such as the discovery of the Higgs boson by the ATLAS [10] and CMS experiments [11].

However, it is accepted that the Standard Model is not a complete theory. A description of gravity is not included in the Standard Model; fundamental differences exist between the accepted theory of gravity, the theory of General Relativity, and the SM. The exclusion of gravity in the SM framework is not its only shortcoming, and this is particularly evident in areas where the theory clashes with cosmological evidence. A number of astronomical observations in the last hundred years have shown the particles described by the Standard Model constitute but a small fraction of the total mass content of the Universe. The current estimate for the fraction of the total mass that is composed of 'ordinary' matter is $\sim 16\%$ [12]; the remaining mass is in the form of 'dark matter', of which very little is known.

One natural assumption of the standard cosmological model, known as the Λ-CDM model, is that matter and antimatter should have been produced in equal amounts at the time of the Big Bang. Nonetheless, the present-day Universe is predominantly matter-dominated, which raises the question of what processes in the early Universe generated the current imbalance

observed. The precise mechanism that generates this imbalance, known as baryogenesis, is still unknown. However, three requirements, the 'Sakharov conditions', are known to be necessary in any theory of baryogenesis [13]: particle interactions that violate baryon number conservation must exist, and also interactions that violate CP symmetry, *i.e.* the symmetry between particles and antiparticles; futhermore, these interactions must have occurred at a time when the Universe was out of thermal equilibrium. The Standard Model can technically satisfy these three conditions, although at a level orders of magnitude lower than what is necessary to account for the observed matter-antimatter asymmetry [14].

The nature of CP symmetry and the processes which violate it in the Standard Model are discussed in Chapter 2. The study of known CP-violating processes and searches for new sources of CP violation are a significant part of the physics programme of the LHCb experiment. The decay channels studied in this thesis have already shown themselves as rich laboratory for CP-violating effects [15–18]. The branching fraction measurements presented in this work represent an additional step towards a better understanding of these decay channels.

Chapter 2 provides a summary of the Standard Model, its symmetries, and mechanisms through which CP violation occurs. The theoretical motivation for the measurement and existing predictions are also discussed. In Chapter 3, a description of the LHCb detector is given, and its performance is discussed. Chapter 4 summarises the strategy used to perform the branching fraction measurement, along with detail of the samples used in the analysis. The strategies used to identify signal decays, against the different background sources considered, are detailed in Chapter 5. In Chapter 6, the different components of the selection efficiency are presented, and their variation across the phase-space is considered. The fitting strategy used to determine the yields of the four signal channels is presented in Chapter 7. The systematic uncertainties associated with the measurement are listed in Chapter 8. Finally, the conclusion in Chapter 9 summarises the final results obtained, and their relation to other studies performed.

The work presented in this thesis is currently under review within the LHCb collaboration. The inclusion of charge-conjugate processes is implied throughout this thesis, unless otherwise specified.

CHAPTER 2

Theory Overview

> *Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which are there.*
>
> Richard Feynman, *The Character of Physical Law*

In this chapter, the relevant theoretical background and motivation for the work presented in this thesis is described. The framework of the Standard Model is described first, in Section 2.1, followed by a short description of the manifestations of CP violation in quarks in Section 2.2. A summary of the properties of decays involving three final-state particles and how they can be described is given in Section 2.3. Finally, an overview of previous studies of the decays explored in this thesis is presented in Section 2.4.

2.1 The Standard Model

The Standard Model is a relativistic quantum field theory that respects the symmetries of the gauge group $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_W \otimes \mathrm{U}(1)_Y$, where the subscripts denote the associated charges conserved by each group: C standing for *colour*, W for *weak isospin*, and Y for *weak hypercharge*. The result is a theory that describes the interactions of matter under the strong, weak, and electromagnetic forces. Gravity, the only other fundamental force, is unformulated as a quantum theory and therefore remains outside the framework of the SM.

While the fundamental objects in the theory are the quantum fields themselves, we generally must learn their properties via the study of the excitations of these fields, *i.e.* the particles associated with each field. There are currently twelve different types of particles in the SM that can be said to constitute the matter in the Universe; another four types, the gauge bosons, serve as mediators of the three fundamental forces being described. A final particle, the Higgs boson, is associated with the mechanism responsible for generating particle masses. The SM particles are listed in Figure 2.1 alongside their mass, electric charge, and spin.

The twelve matter particles, or fermions, can be subdivided into quarks, which feel the strong force, and leptons, which do not. The leptons can be further divided into charged leptons and neutrinos, the latter of which only interact via the weak force. The columns in Figure 2.1 also show another possible division of the fermions, into generations. In fact, all the stable matter in the Universe is made up of generation I fermions only, with the particles in the other generations having identical quantum numbers as their lightest counterpart. Finally, each fermion has an associated anti-matter partner with the same mass and lifetime but opposite quantum numbers.

The six quarks experience the strong nuclear force because they carry the colour charge, of which three states $\{$*red, blue, green*$\}$ exist. The symmetry in the SM associated with the strong force is $\mathrm{SU}(3)_C$ and the theory that models its interactions is called Quantum Chromodynamics (QCD). In QCD, strong interactions are mediated by gluons, which are spin-1 massless bosons and which also carry colour charge themselves. The gluon self-interaction leads to a particular feature of the strong interaction: the value of the strong coupling constant α_s, *i.e.* the strength of the force, is low at short distances (high energies) but high at long distances (low energies). As with electromagnetism, one can think of the lines of force of the gluon field around two quarks interacting. The difference however is due to the gluons being exchanged also interacting with each other, narrowing the flux lines into a 'flux tube'. The potential increases as the two particles are separated and it rapidly becomes more favourable to create a new pair of quarks, rather than to 'stretch the tube'. The final result is *colour confinement* – neither quarks nor gluons can propagate individually but only within bound states, called hadrons, which must have no overall colour charge. In most cases this is achieved in one of two different ways: mesons are quark-antiquark ($q\bar{q}$) states and baryons are three-quark (qqq) states; exotic states with four or five quarks/antiquarks are possible and have been observed [20, 21]. A related consequence of gluon self-interaction is that around the confinement energy scale, $\Lambda_{\mathrm{QCD}} \sim 200\,\mathrm{MeV}$, the coupling strength becomes greater than unity and QCD becomes non-perturbative. This is a problem that will be expanded upon later in Section 2.4.

In the SM the electromagnetic and weak forces are unified into a common 'electroweak' (EW) framework [22–24], the Glashow-Weinberg-Salam (GWS) theory, which is associated with the symmetry $\mathrm{SU}(2)_W \otimes \mathrm{U}(1)_Y$. The electromagnetic force is mediated by the photon (γ),

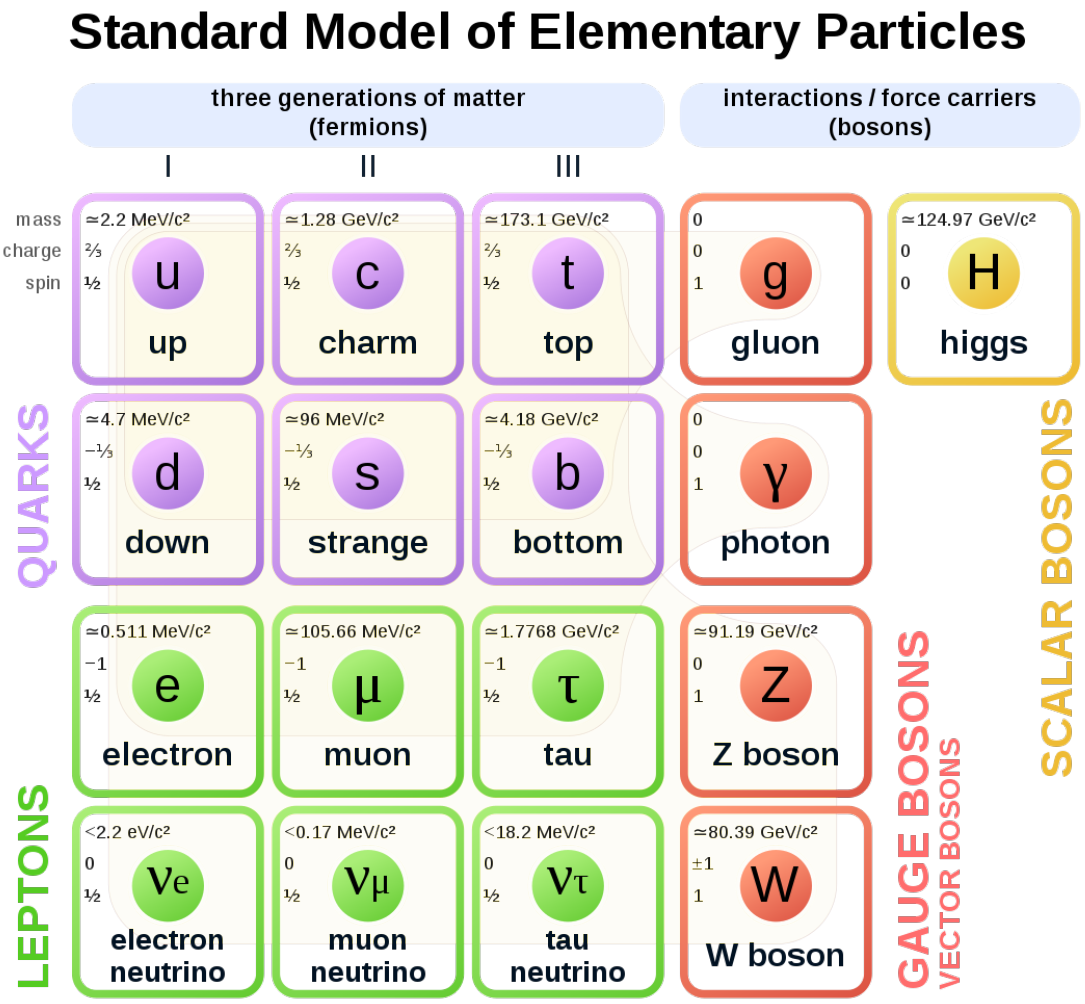

Figure 2.1: The fundamental particles of the Standard Model and their main properties [19].

while the weak force is mediated by three massive spin-1 bosons ($W^{\pm}$ and Z^0). The fact that in nature we observe the three weak bosons as having mass, or equivalently that the weak force is short-range, is *a priori* a problem for any gauge theory describing the weak interaction, since the inclusion of any boson mass terms is forbidden by gauge symmetry. A further complication for a gauge theory of weak interactions is the observation that the weak force is parity-violating (see Section 2.1.1), as it only interacts with the left-handed components of particle fields, which leads to the fermion mass terms also not being gauge invariant. An elegant mechanism to solve both of these issues was found near-simultaneously by Brout and Englert [25], Higgs [26], and

Guralnik, Hagen, and Kibble [27] in which all particles are massless at high energies and their masses below the electroweak scale are generated dynamically by the spontaneous breaking of the electroweak symmetry $\mathrm{SU}(2)_W \otimes \mathrm{U}(1)_Y$.

The Higgs mechanism works via the introduction of a new doublet of complex scalar fields which have a non-zero vacuum expectation value. Three of the four degrees of freedom in the Higgs doublet become the longitudinal components of the now-massive $W^{\pm}$ and Z^0 bosons. The remaining degree of freedom is a real scalar field, which we call the Higgs field, with its associated particle the Higgs boson. Fermions are allowed to interact with the Higgs field via Yukawa interactions that, in the quark sector, take the form

$$\mathcal{L}_{\text{Yukawa}} = -Y^d_{ij}\bar{Q}_{Li}\phi d_{Rj} - Y^u_{ij}\bar{Q}_{Li}\phi_c u_{Rj} + h.c., \tag{2.1}$$

where the $Y^{d,u}$ are 3×3 complex coupling matrices, the $\bar{Q}_L = (u_L, d_L)$ are the left-handed $\mathrm{SU}(2)$ quark doublets, ϕ and ϕ_c the Higgs doublet and its charge conjugate, and d_R and u_R the right-handed $\mathrm{SU}(2)$ quark singlets. The indices i, j run over the number of quark generations. The two Yukawa matrices $Y^{d,u}$ need not be diagonal and indeed they cannot be diagonalised simultaneously. The consequence of this result is a mismatch between quark mass and weak eigenstates and mixing between up- and down-type quarks of different generations, in charged-current weak interactions. The mixing between generations is encapsulated in the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which will be further described in Section 2.1.2.

2.1.1 CPT Symmetries

In addition to its gauge symmetries, the SM can also display three discrete symmetries: parity inversion (P), charge conjugation (C), and time reversal (T). These three can be combined together, with the most important combinations being CP and CPT. The SM must be invariant under the combination CPT as this can be shown to be equivalent with Lorentz invariance for a relativistic quantum field theory. At one time, all three discrete symmetries were thought to be individually conserved, however we now know that both P and C are maximally violated by the weak interaction and their combination CP is also not conserved.

The parity operator inverts the sign of all three spatial coordinates, *i.e.* a vector $\mathbf{a}$ becomes $P(\mathbf{a}) = -\mathbf{a}$. Parity is a good symmetry of both electromagnetic and strong interactions, however it was shown by Wu [28], using the beta decay of cobalt-60, to be (maximally) violated by the weak interaction. The charge conjugation operation flips the sign of all quantum numbers of a particle, *i.e.* it transforms a particle into its antiparticle. It can be seen to be violated by the weak force by noting that the application of C to left-handed neutrinos yield left-handed antineutrinos, which have never been seen in experiment and are not present in the SM.

It was expected that the combination CP would instead be a good symmetry of weak interactions – a left-handed neutrino transforms into a right-handed antineutrino, both being able to participate in weak interactions. However CP was also shown to be violated, first in the decays of $K^0_{\rm L}$ mesons in 1964 [29] and later in the B^0 system [30, 31]. Since then, CP violation has also been seen to manifest in the B^+ and B^0_s systems [32, 33], and most recently in charm decays [34].

2.1.2 The CKM Matrix

The CKM matrix is a 3×3 unitary matrix that describes the possible mixings between the six different quarks. It was, however, first introduced by Kobayashi and Maskawa [35] as a potential model to accommodate CP violation in electroweak theory, at a time when half of the quarks required to make the model work remained undiscovered. In a general form, the mixing between the down-type quarks' mass and weak eigenstates can be written as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}_{\rm CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.2}$$

where the primed states are the weak eigenstates.

Out of the nine independent components of a general 3×3 unitary matrix, five components of the CKM matrix are unphysical phases that can be absorbed into the phases of the quark fields. We are left with four parameters: three real mixing angles and the KM complex phase, the only source of CP violation in the SM. The CKM matrix can be rewritten in terms of these parameters only, the 'standard parameterisation':

$$V_{\rm CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.3}$$

where $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$, and δ is the KM phase.

Another useful parameterisation was first obtained by Wolfenstein [36], by noticing that $s_{13} \ll s_{23} \ll s_{12} < 1$ and expanding in terms of $s_{12} \equiv \lambda \approx 0.23$. The CKM matrix can then be described using the parameters λ, A, ρ, and η, which are commonly defined through

$$s_{23} \equiv A\lambda^2, \quad s_{13}e^{i\delta} \equiv A\lambda^3(\rho - i\eta). \tag{2.4}$$

These definitions ensure that this parameterisation remains unitary to all orders in λ. The CKM

matrix in the Wolfenstein parameterisation is given, to $\mathcal{O}(\lambda^3)$, by

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \tag{2.5}$$

In this parameterisation, the hierarchy of the CKM matrix can easily be seen: the intra-generational couplings are strong, $\mathcal{O}(1)$, while the couplings between generations I and III are very small.

The unitarity of the CKM matrix allows the construction of six triangle relations between its different components, given by

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}, \quad \sum_j V_{ij}V_{kj}^* = \delta_{ik}, \tag{2.6}$$

in the cases where $\delta_{jk}, \delta_{ik} = 0$. The six triangles have the same area, which is a measure of the total amount of *CP* violation in the SM [37].

It is conventional to take one of these relations in which all terms are of the same order in the expansion parameter λ:

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

and dividing it by the best-determined term, $V_{cd}V_{cb}^*$, such that one of the sides is of unit length and the apex is the point $\overline{\rho} + i\overline{\eta} \equiv -V_{ud}V_{ub}^*/V_{cd}V_{cb}^*$. The parameters $\overline{\rho}$ and $\overline{\eta}$ are defined in terms of the other Wolfenstein parameters via

$$\rho + i\eta = \frac{\sqrt{1-A^2\lambda^4}(\overline{\rho}+i\overline{\eta})}{\sqrt{1-\lambda^2}[1-A^2\lambda^4(\overline{\rho}+i\overline{\eta})]}, \tag{2.8}$$

from which the approximations $\overline{\rho} \approx \rho(1-\lambda^2/2)$ and $\overline{\eta} \approx \eta(1-\lambda^2/2)$ can be derived.

The relation in Equation 2.7, normalised by $V_{cd}V_{cb}^*$, is often referred to as the 'unitarity triangle', an illustration of which is shown in Figure 2.2. The three angles of the unitarity triangle are also accessible to experimental determination and are defined as:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.9}$$

The parameters of the CKM matrix cannot be determined from first principles in the SM, but must be determined *a posteriori*. This leaves their experimental determination a crucial goal in modern particle physics. The same CKM parameters can be evaluated by a number of different

measurements; this overconstraining helps both to improve the precision on these parameters and to 'stress test' the SM, with the hope of uncovering beyond-Standard-Model effects.

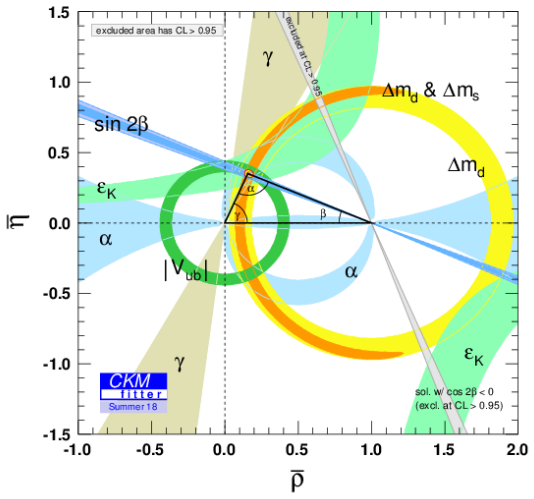

Figure 2.2: Current experimental constraints on the different unitarity triangle parameters [38]. The inner (outer) shaded areas on the $\sin 2\beta$ and the triangle apex constraints correspond to the 68% (95%) confidence levels. The constraints given for all other quantities are their 95% confidence levels.

2.2 *CP* Violation in Quarks

The complex phase in the CKM matrix is the sole source of *CP* violation in the SM. However a non-zero CKM phase, or 'weak phase', is not a sufficient condition for observable *CP*-violating effects in the quark sector. Additionally, *CP* violation can only manifest itself in processes to which multiple amplitudes contribute. This further requirement can be shown by considering the decay of a particle and its *CP* conjugate.

CP violation in decay corresponds to an asymmetry in the rate of a particular decay and its *CP* conjugate. The amplitudes for the two processes can be generally written as

$$A_f = \sum_k |a_k| e^{i(\delta_k + \phi_k)}, \quad \bar{A}_{\bar{f}} = \sum_k |a_k| e^{i(\delta_k - \phi_k)}, \tag{2.10}$$

where the ϕ_k are weak phases, which change sign under *CP*, the δ_k are 'strong phases', the phase components which do not change sign under *CP* (generally produced by strong-force effects), and k labels the possible contributing processes. In the case where the decay can only proceed

via a single amplitude, it is easily shown that there can be no CP-violating decay rate difference:

$$\begin{aligned} |A_f|^2 - |\bar{A}_{\bar{f}}|^2 &= |a_1|^2 e^{i(\delta_1+\phi_1)} e^{-i(\delta_1+\phi_1)} - |a_1|^2 e^{i(\delta_1-\phi_1)} e^{-i(\delta_1-\phi_1)} \\ &= |a_1|^2 - |a_1|^2 \equiv 0. \end{aligned} \tag{2.11}$$

CP violation emerges when multiple amplitudes can interfere with each other, *e.g.* tree- and loop-level contributions:

$$\begin{aligned} \mathcal{A}_{CP} &= \frac{|A_f|^2 - |\bar{A}_{\bar{f}}|^2}{|A_f|^2 + |\bar{A}_{\bar{f}}|^2} \\ &= \frac{2|a_1||a_2|\sin(\delta_1 - \delta_2)\sin(\phi_1 - \phi_2)}{|a_1|^2|a_2|^2 + 2|a_1||a_2|\cos(\delta_1 - \delta_2)\cos(\phi_1 - \phi_2)}. \end{aligned} \tag{2.12}$$

We can see that the $\mathcal{A}_{CP}$ observable depends on both weak and strong phase differences, meaning that extracting the weak phase generally requires knowledge of the strong contribution – often difficult due to hadronic non-perturbative effects.

The CP violation mechanism illustrated above relates to the weak decays of hadrons. However, the phenomenon of neutral meson mixing can induce further forms of CP violation. Mixing is dependent on the mass eigenstates of the neutral meson not being equal to the flavour eigenstates. Using (B^0, $\overline{B}^0$) as an example, the lowest-order Feynman diagrams contributing to mixing are shown in Figure 2.3. The two corresponding mass eigenstates can be written as

$$|B_L\rangle = pB^0\rangle + q|\overline{B}^0\rangle, \quad |B_H\rangle = pB^0\rangle - q|\overline{B}^0\rangle, \tag{2.13}$$

where $B_L(B_H)$ denotes the lighter (heavier) mass state, and p, q are complex numbers requiring $|p|^2 + |q|^2 = 1$. CP violation in mixing equates to an asymmetry in the $B^0 \to \overline{B}^0$ *vs.* $\overline{B}^0 \to B^0$ oscillation rates. This can be shown to correspond to the requirement that $|q/p| \neq 1$ and, equivalently, that mass eigenstates are not CP eigenstates.

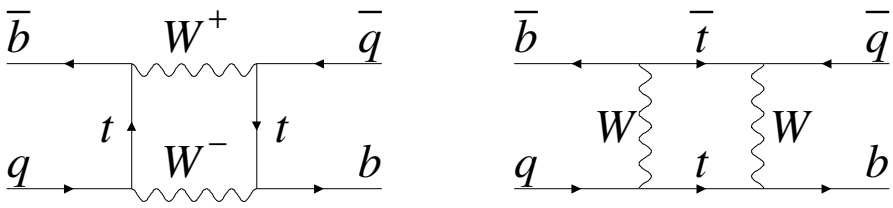

Figure 2.3: $B^0 \leftrightarrow \overline{B}^0$ oscillations via FCNC box diagrams [32].

A third mechanism for CP violation is also available to neutral mesons, in cases where both particle and antiparticle can decay to the same final state. Two decay chains that can

interfere are then available: $B^0 \to f$ and $B^0 \to \overline{B}^0 \to f$. The condition for CP violation in this mechanism is defined as

$$\arg(\lambda_f) + \arg(\lambda_{\bar{f}}) \neq 0, \tag{2.14}$$

where the parameter describing the interference is defined as $\lambda_f \equiv q\bar{A}_f/pA_f$ (and equivalently for the CP conjugate final state $\bar{f}$). In the subset of cases where the final state is a CP eigenstate, the CP asymmetry condition reduces to $\mathcal{I}m(\lambda_f) \neq 0$. A particular detail of note is that the asymmetry from the interference in mixing and decay does not require either component to violate CP individually, *i.e.* one can have $|q/p| = 1$ and $\left|\bar{A}_f/A_f\right| = 1$ and still satisfy the condition $\mathcal{I}m(\lambda_f) \neq 0$.

2.3 Three-body Kinematics

In two-body decays of spin-0 particles, conservation of momentum ensures that the momenta of the decay products are equal and opposite in the rest frame of the initial particle. The only degree of freedom available is the arbitrary choice of decay axis, due to the isotropy of the problem. In contrast, three-body decays (or generally any multi-body decay) have additional degrees of freedom and each decay product can carry away different amounts of the total energy available.

A generic set of three spin-0 particles contains a total of 12 degrees of freedom – all the components of the three four-momenta. Knowledge of the three particle masses removes three degrees of freedom, while requiring conservation of four-momentum in a decay removes another four. In the specific case where both initial particle and all final-state particles are spin-0, there is no angular dependence in the decay – this removes an additional three degrees of freedom. Ultimately, a spinless three-body decay can be fully described by a pair of variables.

Richard Dalitz pioneered the use of the scatter plot of these two variables to describe and study the decay $K^+ \to \pi^+\pi^+\pi^-$ [39]. In modern times, these Dalitz plots (DP) are constructed using pairs of two-body invariant mass combinations $m^2_{ij} = \left(p^\mu_i + p^\mu_j\right)^2$, where $i, j \in \{1, 2, 3\}$ are the indices of the final-state particles. Three different combinations are possible (m^2_{13}, m^2_{23}, m^2_{12}), with the choice of the pair used to represent the DP depending on the particular decay being studied. The boundaries of the DP can be determined from four-momentum conservation, as shown in Figure 2.4.

The decay rate of a three-body decay is given by

$$d\Gamma = \frac{1}{32(2\pi)^3 M^3}|\mathcal{M}|^2 dm^2_{13} dm^2_{23}, \tag{2.15}$$

where M is the mass of the initial-state particle and $\mathcal{M}$ is the matrix element that describes

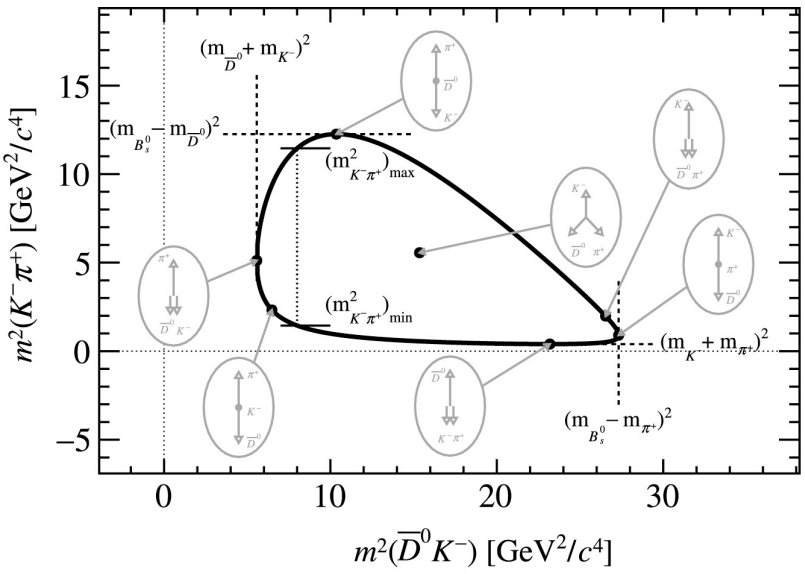

Figure 2.4: The Dalitz plot kinematic boundaries and the corresponding momentum configurations of the three final-state particles in $B_s^0 \to \overline{D}^0 K^- \pi^+$ decays at points of interest in the DP [40].

the dynamics of this particular decay. If $\mathcal{M}$ is constant, the allowed phase space is uniformly populated with events. In most cases, however, a three-body decay will proceed via a number of intermediate quasi-two-body decays, *e.g.* $B^+ \to X\pi^+ \to \pi^+\pi^-\pi^+$ where X is a resonant state such as $\rho^0(770)$ or $f_2(1270)$.

The contributions of the different resonant states manifest themselves in the DP as bands at the value of the resonance mass squared in the corresponding m^2_{ij}variable, extending across the other DP axis. Furthermore, the spin of the resonance will be reflected in the structure of the resonant band: a scalar will generate a uniformly populated band, a vector generates a two-lobe structure, and a tensor generates three-lobe bands. Finally, interference effects between different intermediate states are also reflected in the Dalitz plot. An example of a Dalitz plot with multiple contributing resonances is shown in Figure 2.5.

In decays of B mesons, events typically fall near the kinematic boundaries of the Dalitz plot. It is convenient in such cases to apply a coordinate transformation that extends the edge regions such that variations due to resonance interference and/or experimental effects can be studied more easily. The 'square Dalitz plot' (sqDP) is one such transformation, which also reshapes the kinematically-allowed region into a square with sides of unit length [41]. The

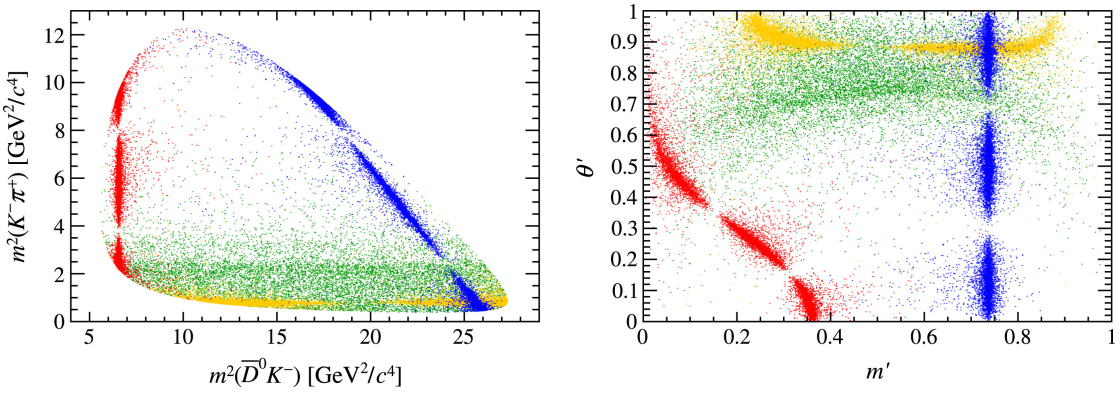

Figure 2.5: Illustration of the Dalitz plot of the decay $B_s^0 \to \overline{D}^0 K^- \pi^+$ with intermediate scalar (green), vector (yellow), and tensor (red and blue) contributions highlighted. Interference effects are not included in this illustration. Both conventional (left), and square (right) representations are shown. An unphysical $\overline{D}^0\pi^+$ resonance (blue) has been added to help visualise the transformation properties of the square Dalitz plot [40].

coordinates of this new representation are given by

$$m' \equiv \frac{1}{\pi} \arccos\left(2\frac{m_{ij} - m_{ij}^{\min}}{m_{ij}^{\max} - m_{ij}^{\min}} - 1\right), \tag{2.16}$$

$$\theta' \equiv \frac{1}{\pi}\theta_{ij}, \tag{2.17}$$

where $m_{ij}^{\min} = m_i + m_j$, $m_{ij}^{\max} = M_B - m_k$, and θ_{ij} is the angle between i and k in the rest frame of ij (*i.e.* the resonance rest frame). The effect of the sqDP transformation on the different regions of the conventional DP can be seen in Figure 2.5.

In order to extract information about the different intermediate states, a model of the decay dynamics is created and its parameters fitted to data. The isobar model is a useful approximation to describe the full three-body decay amplitude. Here the decay amplitude is treated as a coherent sum of a number of amplitudes describing the different intermediate states:

$$\mathcal{A}(m_{ij}^2, m_{jk}^2) = \sum_R c_R F_R(m_{ij}^2, m_{jk}^2), \tag{2.18}$$

where F_R encapsulates the dynamics (such as the lineshape and spin-dependence) of the intermediate state R, and c_R describes the relative magnitude and phase of the different states. The decay amplitude of the CP-conjugate state, $\bar{\mathcal{A}}$, can be constructed in terms of $\bar{c}_R$ and $\bar{F}_R$. Much can be said about how to model F_R and extract the coefficients c_R; a good description of the methods

used is found in Ref. [40]. For the purposes of this thesis, an important set of parameters that can be extracted from an amplitude analysis are the fit fractions of each resonant state:

$$FF_R \equiv \frac{\int_{DP} \left(\left| \bar{c}_R \bar{F}_R \right|^2 + \left| c_R F_R \right|^2 \right) dm_{ij}^2 dm_{jk}^2}{\int_{DP} \left(\left| \mathcal{A} \right|^2 + \left| \bar{\mathcal{A}} \right|^2 \right) dm_{ij}^2 dm_{jk}^2}, \tag{2.19}$$

which can be used to calculate the branching fraction for the quasi-two-body decay, *e.g.* $B^+ \to \rho^0(770)\pi^+$, given (in this example) by:

$$\mathcal{B}(B^+ \to \rho^0(770)\pi^+) = \mathcal{B}(B^+ \to \pi^+\pi^-\pi^+) \cdot FF_{\rho^0}. \tag{2.20}$$

2.4 The $B^+ \to h^+h'^+h''^-$ Decay Channels

The charmless three-body decays of a charged B meson to charged pions and kaons have collectively been an active area of experimental study for the last thirty years. The previous section introduced the notion of many intermediate states contributing to a particular three-body final state. The main Feynman diagrams contributing to the decays studied in this thesis are shown in Figures 2.6, 2.7, 2.8, and 2.9 for the decays $B^+ \to K^+K^+K^-$, $B^+ \to \pi^+\pi^+\pi^-$, $B^+ \to K^+\pi^+K^-$, and $B^+ \to K^+\pi^+\pi^-$ respectively.

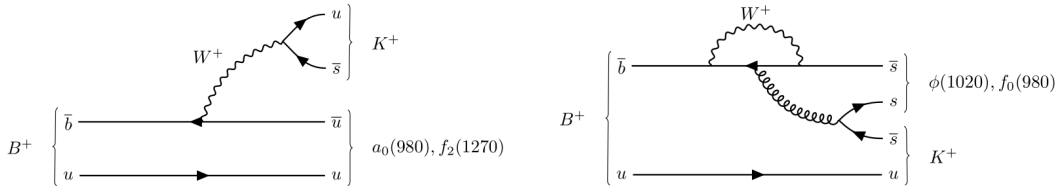

Figure 2.6: Dominant processes contributing to the decay $B^+ \to K^+K^+K^-$. The specific resonant states are given for illustrative purposes, *e.g.* other a, f and ϕ states will also contribute to this decay.

The decay $B^+ \to \pi^+\pi^+\pi^-$ was first studied by the CLEO [42] and ARGUS [43] collaborations in the context of searches for $b \to u$ transitions. Soon after, the decays $B^+ \to K^+K^+K^-$ and $B^+ \to K^+\pi^+\pi^-$ were also studied by the ARGUS collaboration [44]. These studies, and later ones by the DELPHI experiment [45, 46], were not able to claim observations of any of these modes, setting branching fraction upper limits (at the 90% confidence level) of 5–10 times the current observed values.

The large datasets collected by the BaBar and Belle experiments allowed for not only the branching fractions of all four modes ($B^+ \to K^+K^+K^-$, $B^+ \to K^+\pi^+K^-$, $B^+ \to K^+\pi^+\pi^-$,

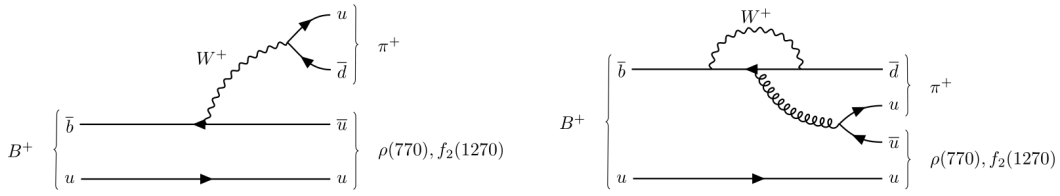

Figure 2.7: Primary quark-level Feynman diagrams contributing to the decay $B^+ \to \pi^+\pi^+\pi^-$. Spectroscopic states other than those mentioned also contribute to the decay.

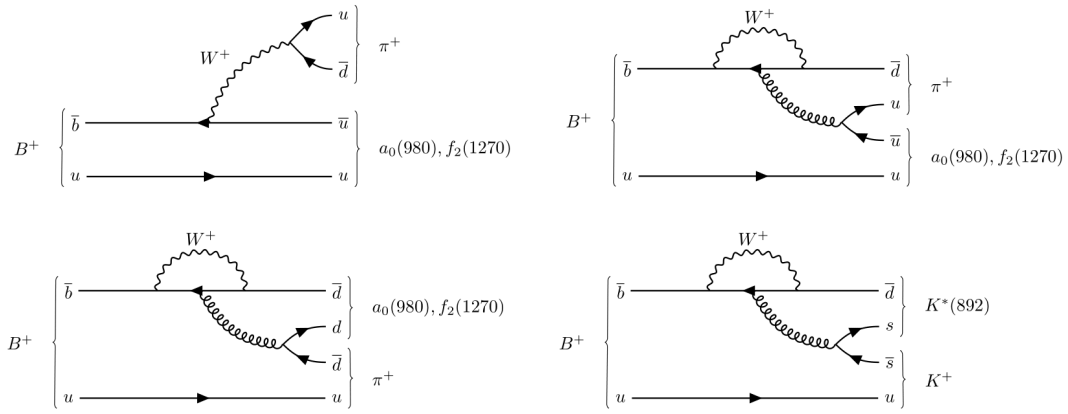

Figure 2.8: Diagrams of the dominant processes that contribute to the decay $B^+ \to K^+\pi^+K^-$. The specific resonant states included are given as representative cases.

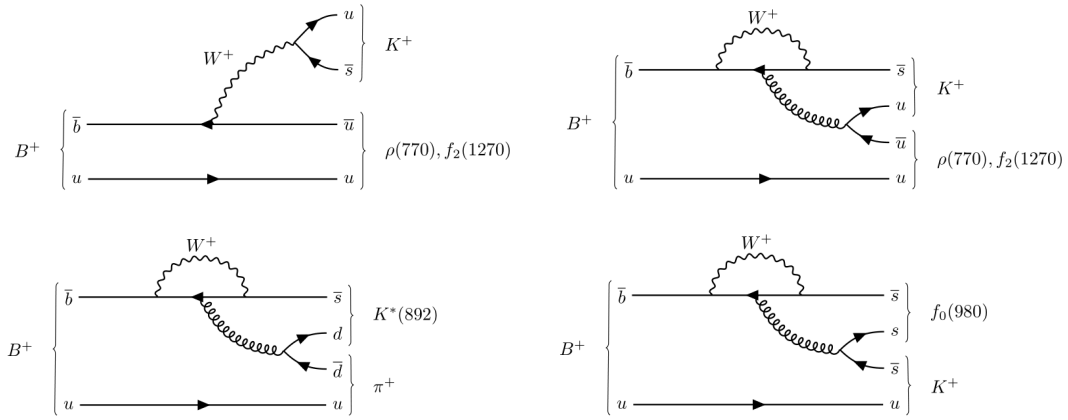

Figure 2.9: Feynman diagrams of the main quark-level processes that contribute to the decay $B^+ \to K^+\pi^+K^-$.

and $B^+ \to \pi^+\pi^+\pi^-$) to be determined, but also for the amplitude analyses of $B^+ \to K^+K^+K^-$, $B^+ \to K^+\pi^+\pi^-$, and $B^+ \to \pi^+\pi^+\pi^-$ to be performed. Table 2.1 lists the current world averages for the branching fractions of the four $B^+ \to h^+h'^+h''^-$ decay channels, which are still dominated by the B factories results.

In lockstep with these experimental developments, the theoretical framework to study these decays was also being developed. A short summary of the main approaches that have been used to provide predictions related to the $B^+ \to h^+h'^+h''^-$ channels is given below.

Table 2.1: Current-best averages of the branching fractions of $B^+ \to h^+h'^+h''^-$ decays [32].

Mode	PDG branching fraction (10^{-6})	References
$B^+ \to K^+K^+K^-$	34.0 ± 1.4	[47, 48]
$B^+ \to K^+\pi^+K^-$	5.2 ± 0.4	[49, 50]
$B^+ \to K^+\pi^+\pi^-$	51.0 ± 2.9	[51, 52]
$B^+ \to \pi^+\pi^+\pi^-$	$15.2 \pm 0.6\,^{+1.3}_{-1.2}$	[53]

The non-perturbative nature of QCD at low energies introduces intractable corrections to the calculable weak decay. Much of the theory effort in the area is dedicated to the development of rigorous techniques for handling such corrections.

The weak decays of B mesons involve three distinct energy scales ($m_W \gg m_b \gg \Lambda_{\rm QCD}$), which allow the problem to be split into a high-energy perturbative and a low-energy non-perturbative regime. The framework used to achieve this separation is the operator product expansion (OPE) [54, 55], in which the Hamiltonian of the full theory being considered can be written as a sum of local operators, each weighted by an associated Wilson coefficient.

The amplitude for the tree-level $b \to u\overline{u}d$ transition shown in Figure 2.10 (a) is given by

$$A = i\frac{G_F}{\sqrt{2}}V_{ub}^*V_{ud}\frac{m_W^2}{k^2 - m_W^2}(\overline{u}b)_{V-A}(\overline{u}d)_{V-A}, \tag{2.21}$$

where G_F is the Fermi constant, $(\overline{u}b)_{V-A} = \left(u^\dagger\gamma_\mu(1-\gamma_5)b\right)$, and k is the momentum transferred via the W propagator. The momentum transfer is of order m_b and so the propagator term can be expanded in terms of $\frac{k^2}{m_W^2}$:

$$\begin{aligned} A &= -i\frac{G_F}{\sqrt{2}}V_{ub}^*V_{ud}(\overline{u}b)_{V-A}(\overline{u}d)_{V-A}\left[1 + \frac{k^2}{m_W^2} + \ldots\right] \\ &= -i\frac{G_F}{\sqrt{2}}V_{ub}^*V_{ud}Q_2 + \mathcal{O}\left(\frac{k^2}{m_W^2} + \ldots\right), \end{aligned} \tag{2.22}$$

allowing everything other than the leading term to be discarded, and a new operator $Q_2 = (\overline{u}b)_{V-A}(\overline{u}d)_{V-A}$ to be defined. It is said that the bosonic degree of freedom has been 'integrated out' and what is left is effectively the Fermi theory (plus CKM factors), with the weak interaction modelled as a four-fermion contact interaction, as shown in Figure 2.10 (b).

Short-distance QCD corrections can now be introduced, the simplest of which corresponds to a hard-gluon exchange between the b quark and any of the other three quark lines participating in the local interaction, as illustrated in Figure 2.10 (c,d). The new operator describing this correction is identical to the one above in Equation 2.22 with the exception of allowing for colour exchange between the two quark currents. The new operator Q_1 is given, alongside the original operator Q_2, by

$$Q_1 = (\overline{u}_i b_j)_{V-A}(\overline{u}_j d_i)_{V-A}, \tag{2.23}$$

$$Q_2 = (\overline{u}_i b_i)_{V-A}(\overline{u}_j d_j)_{V-A}, \tag{2.24}$$

where i, j are colour indices and a summation over repeated indices is implied. Further operators corresponding to different possible transitions can be defined, *e.g.* the gluonic penguin operators Q_{3-6} [55].

Both the operators and their Wilson coefficients are non-trivial functions of the energy scale μ, which essentially separates the high/low energy regimes. The coefficients can be calculated with perturbative methods at $\mu \sim m_W$ to match the effective theory to the full theory and subsequently evolved down to the scale $\mu \sim m_b$.

Different techniques exist to deal with the low-energy QCD phenomenology embedded in the local operators given above, such as QCD factorisation (QCDF) [57], perturbative QCD (pQCD) [58, 59], and Soft Collinear Efffective Theory (SCET) [60, 61]. These three methods build upon an earlier framework, dubbed 'naive factorisation', in which the operators of the effective theory can be expressed as

$$\langle M_1 M_2 | Q_i | B \rangle = \langle M_2 | J_2 | 0 \rangle \langle M_1 | J_1 | B \rangle. \tag{2.25}$$

The matrix element corresponding to the decay $B \to M_1 M_2$ is factorised as a product of a $B \to M_1$ form factor and the M_2 decay constant, with M_1 being defined as the final-state particle that picks up the spectator quark from the B meson. Form factors and decay constants are still non-perturbative quantities but are nonetheless simpler to predict, using lattice QCD or QCD sum rules, or to determine experimentally from (semi)leptonic decays. This approach is limited in that gluon exchanges between M_2 and the (BM_1) system are not considered below $\mu \sim m_b$.

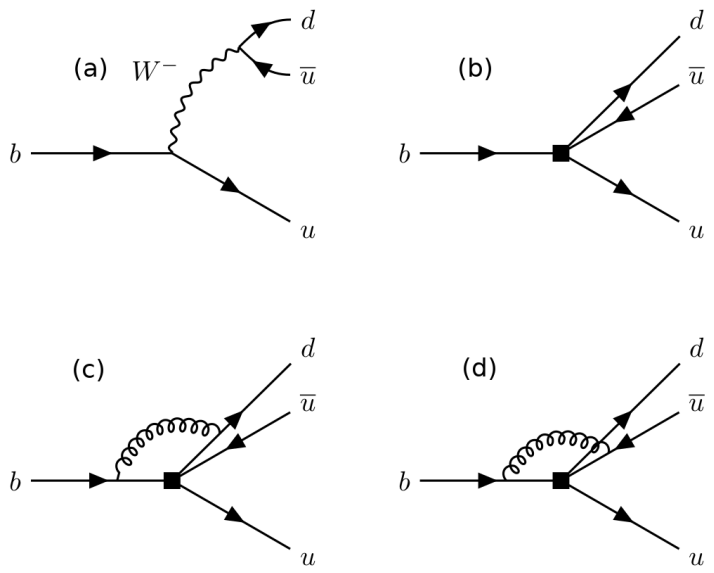

Figure 2.10: Leading-order Feynman diagram for the decay $b \to u\overline{u}d$ in the full theory (a). Corresponding diagram in the OPE, in which the square represents the four-fermion operator Q_2 (b). One-loop QCD corrections contributing to the operator Q_1 (c)(d). An additional diagram connecting $b - u$ has been omitted [56].

The QCDF approach is based on the heavy-quark limit $m_b \gg \Lambda_{\rm QCD}$ and the principle of 'colour transparency' – the quark pair that hadronises into M_2 is emitted in a compact state which behaves as a small colour dipole, which only interacts with soft gluons via terms suppressed by $\Lambda_{\rm QCD}/m_b$. Hence, to leading order in the expansion parameter, only hard gluonic interactions connect M_2 to (BM_1) [62].

In the pQCD approach, the decay amplitude is written as a convolution of the B, M_1, and M_2 wavefunctions and a hard-scattering component. The meson wavefunctions encapsulate the non-perturbative contributions but are universal – once determined they can be used in any transitions involving that meson. The hard component is a six-quark operator corresponding to diagrams with hard-gluon exchanges involving the spectator quark.

In SCET, one identifies an intermediate energy scale $\mu \sim \sqrt{m_b \Lambda_{\rm QCD}}$ corresponding to energies at which quarks can emit gluons collinear to their momenta. The effective theory decomposes quarks and gluons into distinct collinear/soft/ultrasoft fields, which can interact according to different Feynman rules.

Flavour $\mathrm{SU}(3)$ and isospin are also powerful tools when considering multiple decays that are related via these symmetries. An alternative approach to those outlined above is found by considering observables from multiple decay amplitudes, related to each other via $\mathrm{SU}(3)$ flavour

symmetries, and performing global fits to extract theory parameters that can then be used to make predictions [63, 64]. Symmetry-breaking parameters can also be added to the fits in order to test the validity of flavour $\mathrm{SU}(3)$ symmetry in the presence of unequal u, d, s quark masses and possible electromagnetic corrections. This approach is non-perturbative in the sense that the parameters extracted from data encompass all orders of strong interactions, which allows for interesting comparisons with predictions given by the methods above.

The methods given above have been used to obtain predictions for a number of quasi-two-body decays. Some examples are given in Table 2.2. It can be seen that there is a fair agreement between theory and experiment, though large uncertainties are associated with some predictions.

It is not trivial to move from theoretical methods for two- and quasi-two-body decays towards inclusive predictions for three-body decays. Indeed, there are still no established theories for such decays. Though many quasi-two-body intermediate states contribute to the inclusive $B^+ \to h^+h'^+h''^-$ branching fractions, the non-resonant fractions are large (*e.g.* $\sim 65\%$ in $B^+ \to K^+K^+K^-$ [48]). There has been, however, some progress in this area. One approach uses Heavy Meson Chiral Perturbation Theory (HMChPT) to model the non-resonant component at the centre of the Dalitz plot, with an exponentially-decaying momentum dependence to extend the amplitude past the region where HMChPT is valid [65]. The non-resonant amplitude is then included alongside predicted resonant amplitudes in an isobar amplitude to give inclusive three-body results, given in Table 2.2.

The LHCb collaboration has recently performed amplitude analyses of the $B^+ \to K^+\pi^+K^-$ [16] and $B^+ \to \pi^+\pi^+\pi^-$ [17, 18] modes, which use data samples significantly larger than those of the B factories. These results, and corresponding amplitude analyses of $B^+ \to K^+K^+K^-$ and $B^+ \to K^+\pi^+\pi^-$, will improve our understanding of the multiple intermediate states that contribute to these decay channels. The fit fractions reported by the amplitude analyses must be converted into branching fractions of the different quasi-two-body states, using Equation 2.20, to allow for comparison with theoretical predictions and to serve as input to theory models. It is clear then that to profit fully from the larger LHCb dataset, the inclusive $B^+ \to h^+h'^+h''^-$ branching fractions must also be determined.

Table 2.2: Branching fraction predictions ($\times 10^{-6}$) for intermediate resonant contributions to $B^+ \to h^+h'^+h''^-$. The values given are for $B^+ \to Rh^+$, corrected for secondary $R \to hh$ branching fractions. For conciseness, only the sums in quadrature of the errors in the theory predictions are shown. The results of the latest LHCb amplitude analyses of $B^+ \to K^+\pi^+K^-$ [16] and $B^+ \to \pi^+\pi^+\pi^-$ [17, 18] are not included in Ref. [32].

Mode	Experiment [32]	QCDF [65]	pQCD	SCET * [66]
$B^+ \to K^+K^+K^-$	34.0 ± 1.4	$28.8\,^{+7.9}_{-6.4}$	–	–
$\phi(1020)K^+$	$8.8\,^{+0.7}_{-0.6}$	$8.9\,^{+1.6}_{-1.4}$ **	$7.8\,^{+5.9}_{-1.8}$ [67]	$9.7\,^{+5.2}_{-4.2}$ $8.6\,^{+3.4}_{-2.9}$
$B^+ \to K^+\pi^+K^-$	5.2 ± 0.4	$5.2\,^{+1.3}_{-1.1}$	–	–
$K^*(892)^0K^+$	< 1.1 †	$0.32\,^{+0.6}_{-0.6}$ ††	$0.50\,^{+0.18}_{-0.14}$ [68]	$0.49\,^{+0.28}_{-0.22}$ $0.51\,^{+0.19}_{-0.17}$
$B^+ \to K^+\pi^+\pi^-$	51.0 ± 2.9	$42.2\,^{16.1}_{-10.7}$	–	–
$\rho^0(770)K^+$	3.7 ± 0.5	$2.9\,^{+0.7}_{-0.2}$	$4.04\,^{+0.86}_{-0.69}$ [69]	$6.7\,^{+2.9}_{-2.4}$ $4.6\,^{+1.9}_{-1.6}$
$f_2(1270)^0K^+$	1.07 ± 0.27	$3.8\,^{+7.8}_{-3.0}$ [70]	$1.976\,^{+1.251}_{-1.065}$ [71]	–
$K^*(892)^0\pi^+$	10.1 ± 0.8	$12.6\,^{+3.2}_{-2.8}$ †	$7.12\,^{+2.77}_{-2.08}$ [68]	$8.5\,^{+5.0}_{-3.9}$ $9.9\,^{+3.7}_{-3.2}$
$B^+ \to \pi^+\pi^+\pi^-$	$15.2 \pm 0.6\,^{+1.3}_{-1.2}$	$17.0\,^{+2.2}_{-2.4}$	–	–
$\rho^0(770)\pi^+$	8.3 ± 1.2	$7.3\,^{+0.4}_{-0.4}$	$8.84\,^{+1.91}_{-1.69}$ [69]	$10.7\,^{+1.2}_{-1.1}$ $7.9\,^{+0.8}_{-0.8}$
$f_2(1270)^0\pi^+$	$1.6\,^{+0.7}_{-0.4}$	$2.7\,^{+1.4}_{-1.2}$ [70]	$1.879\,^{+1.129}_{-0.891}$ [71]	–

* The authors find two solutions for their fit to the non-perturbative inputs to the analysis. Both solutions are used to obtain predictions, hence the two values given for each resonant contribution.

** Corrected for $\mathcal{B}\left(\phi \to K^+K^-\right) = 49.2 \pm 0.5\%$.

† The LHCb amplitude analysis of $B^+ \to K^+\pi^+K^-$ observed $B^+ \to K^*(892)^0K^+$ with a fit fraction of $7.5 \pm 0.8\%$. This corresponds to a branching fraction of $B^+ \to K^*(892)^0K^+ = 0.59 \pm 0.08$, when accounting for $\mathcal{B}\left(K^*(892)^0 \to K^+\pi^-\right) = 2/3$.

†† Corrected for $\mathcal{B}\left(K^*(892)^0 \to K^+\pi^-\right) = 2/3$.

CHAPTER 3

The LHCb Experiment

> *It doesn't work to build half an accelerator. The particles need to go all the way around.*
>
> Steven Weinberg, *On the Shoulders of Giants*

The work described in this thesis uses data collected by the Large Hadron Collider beauty (LHCb) experiment, located at one of the collision points of the Large Hadron Collider (LHC) at CERN. This chapter describes relevant aspects of the accelerator system in Section 3.1, and details of the LHCb detector and its sub-detectors in Section 3.2. Information on the LHC is mainly derived from Ref. [72], while information on LHCb comes primarily from Refs. [73, 74]; other sources are indicated throughout.

3.1 The Large Hadron Collider

The Large Hadron Collider is the world's largest particle accelerator, a $27\,\mathrm{km}$ circular collider located under the Franco-Swiss border at the European Organization for Nuclear Research, CERN. It is primarily designed to collide a pair of proton beams at centre-of-mass energies of up to $14\,\mathrm{TeV}$, but it is also able to accelerate and collide lead ion beams, and recently xenon beams, up to an energy of $2.8\,\mathrm{TeV}$ per nucleon.

The LHC is the final step in a long chain of particle accelerators required to increase particle energies up to the TeV range. An illustration of the full accelerator complex is shown in Figure 3.1. Protons are produced by ionising hydrogen stored in a gas canister. In the first step in

the chain, these protons are accelerated in a linear accelerator, LINAC2, to an energy of 50 MeV. From here, the beam is injected into a sequence of successively larger circular accelerators: the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS), which raise the beam energy to 1.4 GeV, 25 GeV, and 450 GeV respectively. After exiting the SPS, the beam is split into two; the two resulting beams are transferred to the LHC beam pipes, in which they circle in opposite directions. The two beams are then accelerated up to their collision energy: in 2011, this energy was 3.5 TeV per beam; in 2012, 4 TeV; finally, in 2015-2018, the beam energy reached 6.5 TeV.

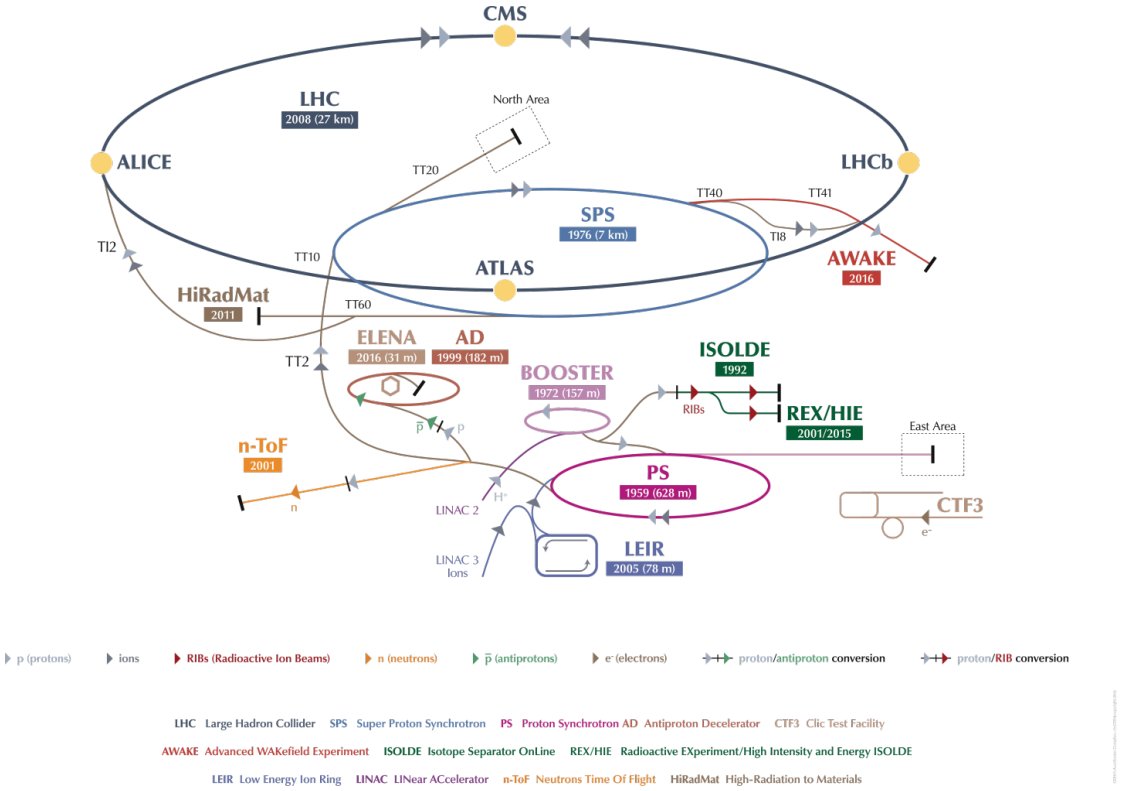

Figure 3.1: The LHC accelerator complex and experiments [75].

The LHC beams are not continuous beams but are composed of a number of proton bunches, with a designed maximum of 2808 bunches per beam and with each bunch containing $\mathcal{O}(10^{11})$ protons. During 2017-2018, the LHC was able to reach a total of 2556 bunches per beam delivered to interaction points 1 and 5, where ATLAS and CMS are respectively located. The number of colliding bunches at interaction point 8, where LHCb is located, is slightly smaller due to the structure of the beams. The minimum spacing between bunches is 25 ns, corresponding

to a bunch crossing frequency of 40 MHz. The maximum design instantaneous luminosity is $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, however operations to date have seen the machine achieve peak instantaneous luminosities of approximately double this number. The LHCb detector, however, operates at a much lower instantaneous luminosity, for reasons discussed in the next section.

The main data-taking periods to date have been 2011-2012, known as 'Run 1', and 2015-2018, known as 'Run 2'. The long shutdown between the runs was required in order to prepare the machine to operate at higher energy and for overall renovation work. A second shutdown, 'Long Shutdown 2', has started in 2019 and will continue until 2021.

The two proton beams interact simultaneously at four points around the LHC ring. It is at these points where each of the four major LHC experiments are located. ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid) are 'general-purpose' detectors with wide-ranging physics programs, including Higgs, top, and electroweak physics, but also direct searches for beyond-Standard-Model particles. ALICE (A Large Ion Collider Experiment) is dedicated to quark-gluon plasma physics, with these plasmas being produced in heavy-ion collisions. The LHCb experiment is focused on the study of the decays of heavy-flavoured hadrons and searching for indirect evidence of non-Standard-Model physics via precision measurements of observables such as *CP* asymmetries, angular observables, and unitarity triangle parameters.

3.2 LHCb Detector Overview

The LHCb detector is a single-arm forward spectrometer specifically designed for the study of hadrons containing b or c quarks. To this purpose, it is composed of a number of sub-detectors that broadly fall into one of two categories: tracking detectors, and particle identification detectors. The LHCb coordinate system is defined as a right-handed Cartesian system, with its origin at the nominal interaction point and the positive z-axis parallel to the beampipe, in the direction of the various sub-detectors. A side-view schematic of the detector and its subsystems is shown in Figure 3.2.

The geometry of the detector has been chosen to exploit the physics of b-quark production at the LHC [76, 77]. At LHC energies, parton-parton interactions are asymmetric in the relative momentum of the two partons, resulting in a momentum transfer to the produced $b\bar{b}$ pair in the direction of the higher-momentum parton. The relatively low-mass pair is therefore significantly boosted, with both quark and anti-quark travelling in the same direction. The LHCb detector has an angular coverage extending from 10 to 300 mrad in the horizontal plane and from 10 to 250 mrad in the vertical plane. This corresponds to a pseudorapidity (defined as $\eta = -\ln(\tan\theta/2)$, where θ is the polar angle with respect to the z axis) range of $1.6 < \eta < 4.9$. The expected distribution of $b\bar{b}$ pairs, as a function of the pseudorapidities of the two in the pair,

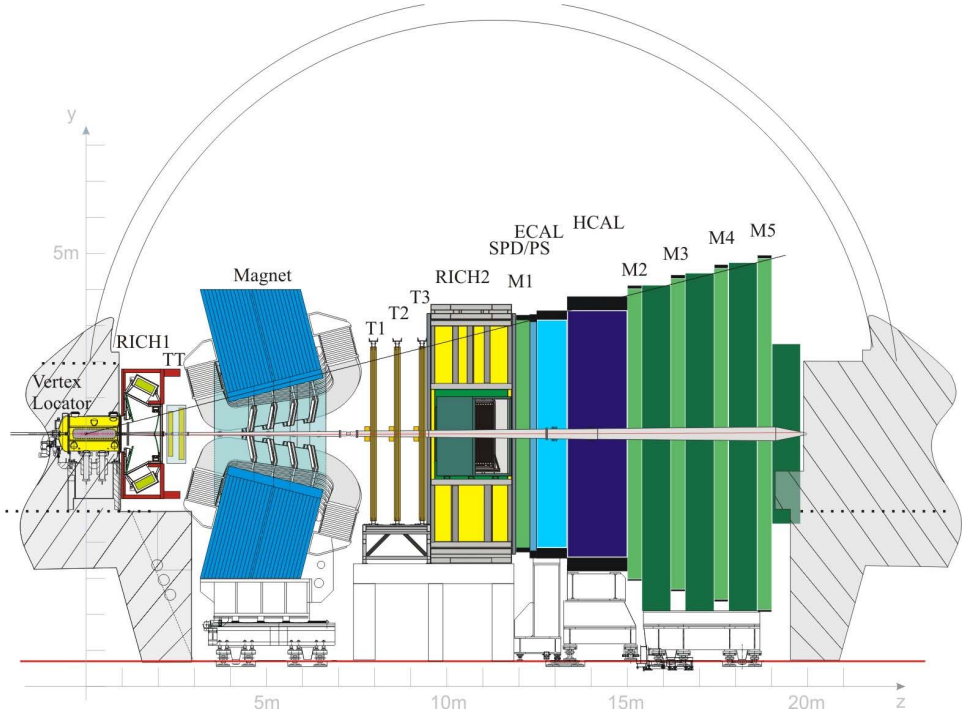

Figure 3.2: A cross-section of the LHCb detector and its sub-detectors [73].

is shown in Figure 3.3; approximately 25% of pairs produced fall within the LHCb detector acceptance.

The LHCb detector was designed to operate at an instantaneous luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, or fifty times lower than the maximum design LHC luminosity. Nonetheless, LHCb has mainly been operated at twice this design luminosity. At this lower luminosity, the probability of a single pp interaction per bunch crossing is maximised, which reduces the overall number of tracks that must be reconstructed. At the start of a fill, the beams at the LHCb collision point start with the minimal amount of overlap required to reach the target luminosity. As the number of protons in the bunches decrease, the beam overlap is increased to compensate. The technique, known as 'luminosity levelling', allows LHCb to maintain a constant luminosity throughout most of a fill, as shown in Figure 3.4. It has the additional benefits of reducing the amount of radiation damage to the Vertex Locator sub-detector and allowing the same trigger configuration to be used throughout a fill.

The detector's unique configuration among the four major LHC experiments also allows the LHCb experiment to position itself as a general forward experiment, able to perform measurements in an acceptance region that is complementary to ATLAS and CMS. Examples include vector boson [78], top [79], and charmonia [80, 81] production cross-sections, dark matter

searches [82], and Higgs-like boson searches [83, 84].

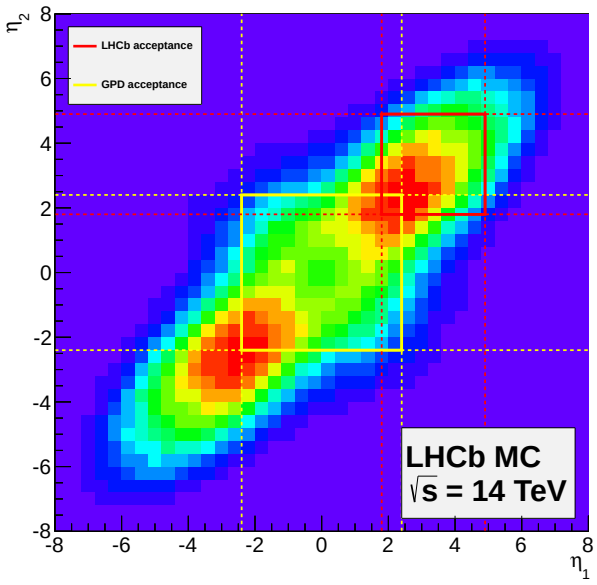

Figure 3.3: Pseudorapidity distribution of $b\bar{b}$ pairs produced in simulated pp collisions [85]. The central yellow square corresponds to the angular acceptance of ATLAS and CMS, while the acceptance of the LHCb detector is denoted by the red square in the top-right quadrant.

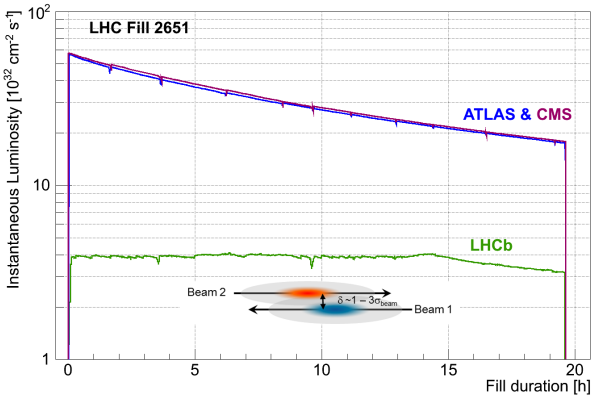

Figure 3.4: Instantaneous luminosity at LHCb, ATLAS, and CMS during a Run 1 fill [74].

3.3 Tracking

The LHCb tracking system consists of the Vertex Locator (VELO), a warm dipole magnet, and four tracking stations: one, the Tracker Turicensis (TT), located before ('upstream' of) the magnet, and three (T1, T2, T3) located after ('downstream' of) the magnet. Each of the three downstream stations are divided into two sections: the inner, silicon-based, parts of the three stations are collectively referred to as the Inner Tracker (IT); the outer parts, which use straw tubes instead, are known as the Outer Tracker (OT).

3.3.1 Vertex Locator

The VELO is the first LHCb sub-detector encountered by particles produced in the pp collision, surrounding the interaction point to allow for precise primary and secondary vertex reconstruction. It is made up of two halves, each containing 21 semi-circular silicon strip modules providing hit information in the r- and ϕ-directions. An overview of the VELO layout is seen in Figure 3.5.

To achieve its goal, the VELO must be able to gather measurements from the closest possible region to the collision point. During data-taking conditions, taking into consideration detector safety reasons, the active area of the modules starts at a distance of $8.2\,\mathrm{mm}$ away from the beamline. During beam injection and ramping, however, the aperture required by the LHC is larger than this distance, which requires the VELO halves to be made retractable by a distance of $30\,\mathrm{mm}$. A further mechanism allows the VELO to recentre around the real beam position, which can change from fill to fill, during the closing procedure.

Each VELO half is placed in a $300\,\mu\mathrm{m}$ thick aluminium shielding box, the RF foil, designed to prevent RF interference induced by the beam. The RF foil also protects the LHC vacuum from possible outgassing from the detector modules.

The silicon modules are positioned within the VELO such that any track within the $300\,\mathrm{mrad}$ acceptance must cross at least four modules. This means most modules populate the region closest to the nominal interaction point, with a minimum distance in the z direction between modules of $3.5\,\mathrm{cm}$. To fully cover the acceptance in the ϕ, or azimuthal direction, the two halves must overlap; this is achieved by shifting one VELO half by $1.5\,\mathrm{cm}$ in the z direction with respect to the other half.

Each of the 21 standard VELO modules is composed of two sensors: the R-sensor provides information on the radial distance from the beam axis, and the ϕ-sensor provides information on the azimuthal direction. In each VELO half there are two additional upstream stations, containing only an R-sensor each, that collectively form the pile-up system. These stations are also referred to as "Veto stations", such as in Figure 3.5. Knowledge of a hit's position

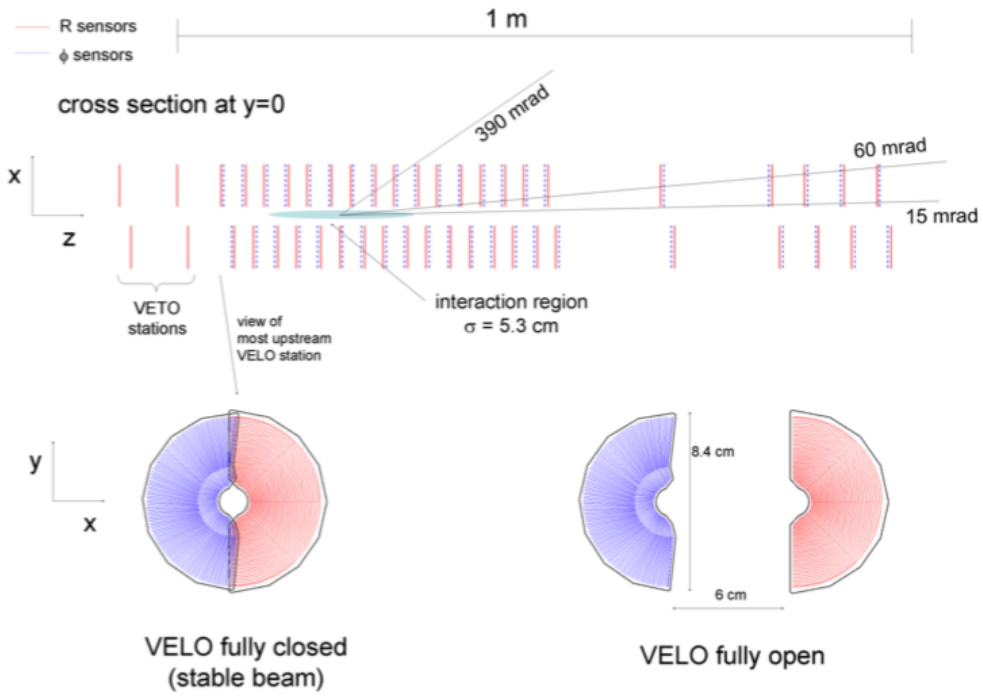

Figure 3.5: Layout of the VELO modules along the beamline (top). Front view of a module in both closed and open configurations (bottom). Notice the small module overlap in the closed position [73].

along the z axis is given by the position of the sensor itself in the LHCb coordinate system.

The R-sensors are segmented into concentric strips with a linearly increasing pitch, from 38 μm at the inner edge to an outer pitch of 102 μm. The strips are also divided into four 45° regions to reduce the occupancy and capacitance in each strip. In the ϕ-sensors, the strips are split into two regions in order to reduce the pitch at the outer edge of the sensor. In the inner region, the strip pitch increases from 38 μm to 79 μm at a radius of 17.25 mm. The pitch in the outer region starts at 39 μm and increases to 97 μm at the outer edge. The strips are not orthogonal to those in the R-sensors but are laid out at a skew of 20° (10°) away from the radial direction in the inner (outer) region. This improves pattern recognition as it introduces a stereo angle when adjacent modules are placed so that they have opposite skew. The geometries of both types of sensor are shown in Figure 3.6.

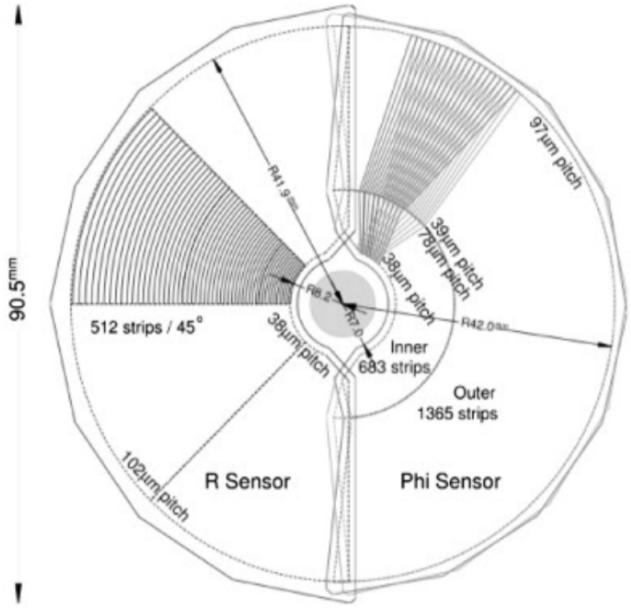

Figure 3.6: Sketch of the sensor geometry. On the ϕ-sensor, the strips from two, adjacent, modules are shown to highlight the stereo angle [73].

3.3.2 Tracker Turicensis

The TT is a silicon microstrip detector located between RICH1 and the LHCb magnet, covering the full acceptance of the detector. The layout and the TT's physical dimensions are shown in Figure 3.7. It is composed of four layers arranged in a (x-u-v-x) formation: the first and last layers have vertical strips; the second and third have strips rotated by -5° and 5° respectively. The first two layers are separated from the last two by approximately $27\,\mathrm{cm}$ in the z direction.

Each layer is built up of half-modules, each a column of seven silicon sensors. The regions directly above and below the beampipe are covered by one half-module each. To either side of the beampipe, there are a further seven (eight) half-modules in the first (last) two layers.

The purpose of the TT is two-fold: it is used to reconstruct the trajectories of low-momentum particles that are swept by the magnet and do not reach the T-stations (T1–T3), and to reconstruct long-lived particles such as Λ and K^0_S which decay outside of the VELO.

3.3.3 Inner Tracker

The IT covers the high-occupancy region of the three downstream tracking stations, with silicon trackers arranged in a cross shape around the beam pipe. Although the IT covers only 1.3% of the area in each tracking station, approximately 20% of all charged particles produced will pass

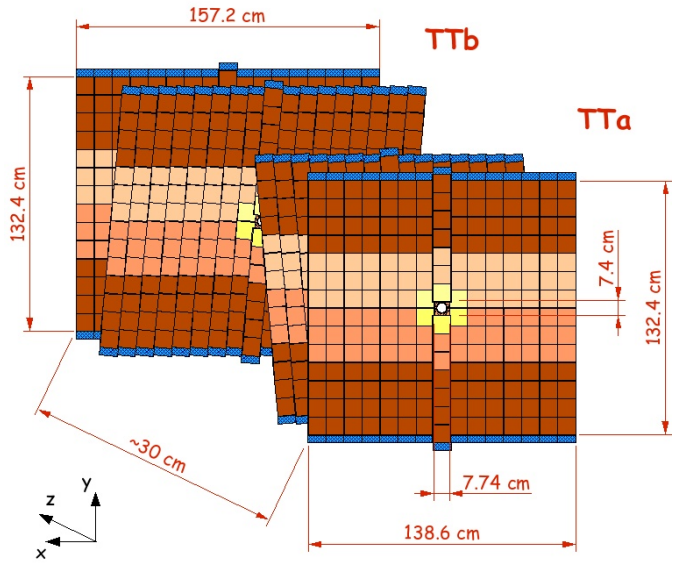

Figure 3.7: Layout and dimensions of the Tracker Turicensis, highlighting the skewed formation of the layers [86].

through it. It follows a very similar design as the TT, with each station being composed of four layers in the same (x-u-v-x) arrangement. Unlike the TT, each layer is built up of modules with one or two silicon sensors, rather than seven, as shown in Figure 3.8.

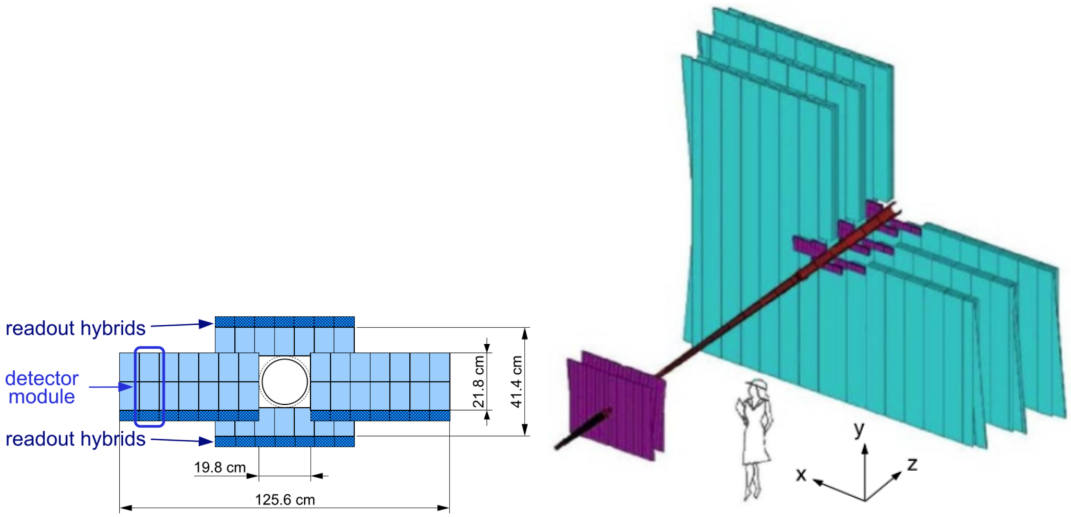

Figure 3.8: Schematic of one Inner Tracker station (left). Overview layout of the tracking stations, highlighting the difference in size between IT and OT (right) [73].

3.3.4 Outer Tracker

The OT is a drift-time detector, composed of individual gas-tight straw-tube modules, and it covers most of the acceptance of the three T-stations. In each station, the OT consists of four layers of modules also matching the (x-u-v-x) formation found in the TT and IT. Each module is made up of two rows of straw tubes, as shown in Figure 3.9.

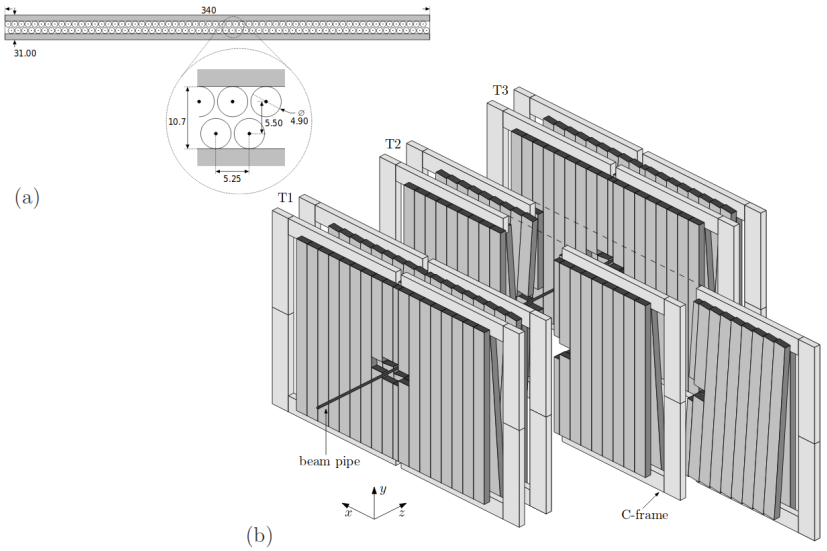

Figure 3.9: Cross section of one OT layer, showing the dimensions and packing of the two layers of straw tubes (a). Layout of the OT system; each station is separated into two retractable halves (b) [87].

In a drift-time detector charged particles ionise the gas inside the tubes, with the ions then travelling to an anode wire at the centre of the tube. Reaching the design momentum resolution of $\delta p/p \sim 0.4\%$ requires a fast drift time of under 50 ns and good resolution of the drift coordinate (200 μm). The gas mixture chosen is a 70/28.5/1.5 mix of Argon/CO_2/O_2.

3.3.5 Magnet

In order to achieve the required precision on the momentum measurements of charged particles travelling through the detector, a well-understood magnetic field is needed to curve their path. The LHCb magnet is a warm dipole magnet consisting of two identical saddle-shape coils, each composed of fifteen aluminium layers, enclosed in an iron yoke, as shown in Figure 3.10. The coils are placed symmetrically about the z-axis, at a distance such that the window between them covers the full LHCb angular acceptance.

The magnetic field is parallel to the y-axis and the magnet is designed to allow its polarity

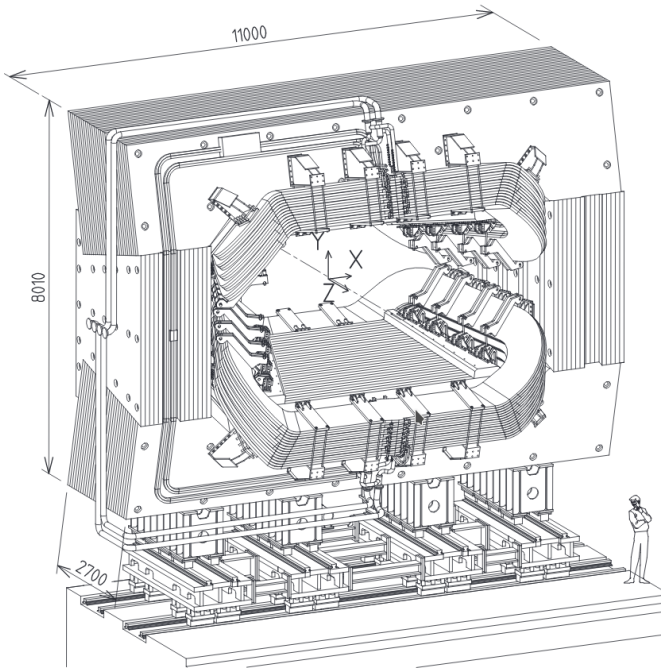

Figure 3.10: Schematic view of the dipole magnet (units in mm). In this projection, the positive z direction points out of the page [73].

to be reversed regularly throughout data-taking periods. This allows the study of systematic effects related to detector asymmetries, which affect CP asymmetry measurements. The integrated magnetic field delivered over a $10\,\mathrm{m}$ track length is $4\,\mathrm{Tm}$. In order to achieve the required momentum precision of $\delta p/p < 5 \times 10^{-3}$ for particles with momentum lower than $100\,\mathrm{GeV}/c$, the magnetic field integral must be known to a relative precision of $\sim 10^{-4}$. This was achieved with the use of an array of Hall probes that measured all components of the magnetic field within the detector volume up to the RICH2 detector. The distribution of the magnetic field in the tracking region of the detector is shown in Figure 3.11.

3.3.6 Vertexing and Tracking Performance

Determining the precise location of primary and secondary vertices in an event is crucial in identifying heavy-flavour decays. Primary vertex (PV) resolution depends strongly on the number of tracks used to reconstruct the vertex. At LHCb, the resolution is measured by randomly splitting the set of tracks into two and reconstructing the PVs using both sets. The PV resolution is the width of the distribution of the difference in vertex position, multiplied by a factor of $\sqrt{2}$. A nominal event with 25 reconstructed tracks associated with a PV has a PV resolution of $13\,\mu\mathrm{m}$ in x, y and $71\,\mu\mathrm{m}$ in z. The resolution as a function of the number of tracks is shown in Figure 3.12.

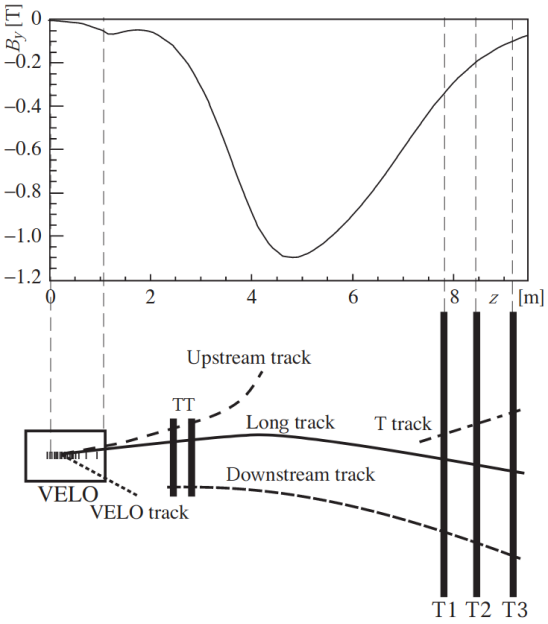

Figure 3.11: The magnetic field strength as a function of z, from the interaction point to the third tracking station (top). Track types at LHCb and the tracking sub-detectors they travel through (bottom) [88].

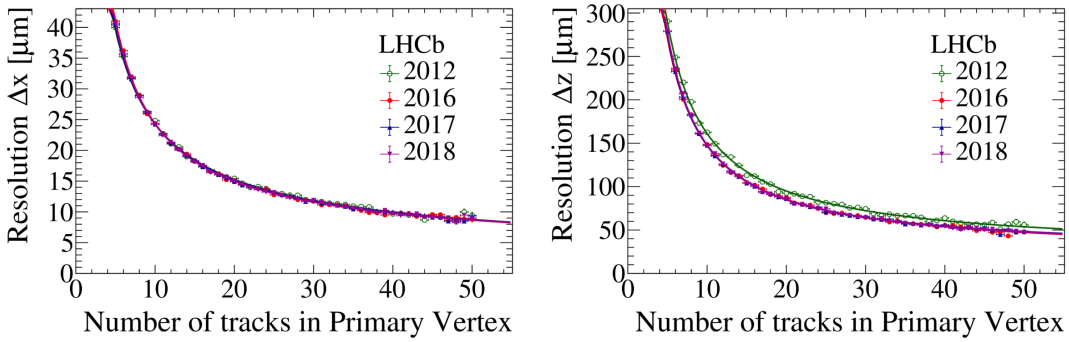

Figure 3.12: Primary vertex resolution as a function of number of tracks in x (left) and z (right), for both Run 1 and Run 2 [89].

The impact parameter (IP) of a track is defined as the distance-of-closest-approach between the track and a PV. Tracks made by the decay products of a B meson will generally have large IPs, as the B decay vertex is sufficiently displaced from the PV. It is therefore an important quantity for event selection in B physics analysis. The IP resolution as a function of p_{T}, for a number of data-taking periods, is given in Figure 3.13.

Tracks at LHCb are classified by which tracking sub-detectors they cross through, as

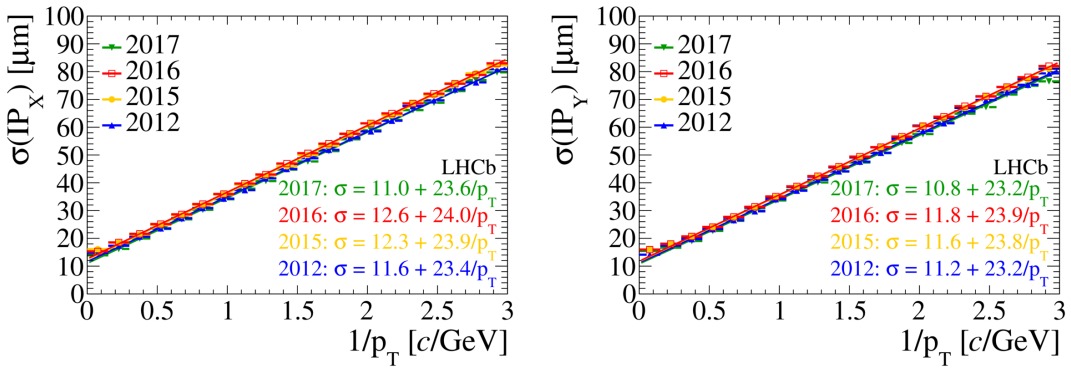

Figure 3.13: Impact parameter resolution as a function of p_{T} in x (left) and y (right), for both Run 1 and Run 2 [89].

shown in Figure 3.11. The work presented in this thesis uses only 'long' tracks, those which cross the entire tracking system. The tracking efficiency of these are measured using a 'tag-and-probe' method. A J/ψ meson decaying into two muons is identified, with one of the muons being fully reconstructed (the 'tag') while the other (the 'probe') is only partially reconstructed. In this case the probe muon is reconstructed using only the information from the muon system, which is described in Section 3.4.3. The average track reconstruction efficiency is above 96% in most of the phase space, as shown in Figure 3.14.

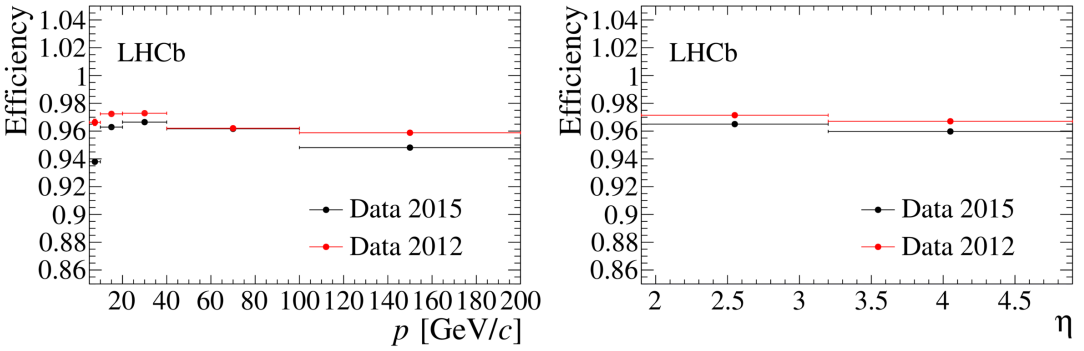

Figure 3.14: Track reconstruction efficiency as a function of momentum (left) and pseudorapidity (right), for both Run 1 and Run 2 [89].

3.4 Particle Identification

The majority of decays studied by the LHCb experiment rely on the accurate identification of a small selection of final-state particles, namely: charged hadrons (pions, kaons, protons), muons,

electrons, and photons. The LHCb detector employs two Ring Imaging Cherenkov (RICH) detectors to distinguish between species of charged hadrons; a calorimetry system to measure the energy and identify electrons, photons, and hadrons; a muon detector system to identify and measure the momentum of muons. The positioning of these subsystems within the detector can be seen in Figure 3.2.

3.4.1 Ring Imaging Cherenkov Detectors

The two RICH detectors have the crucial goal of identifying pions, kaons, and protons. The decay channels studied in this work are all exclusively composed of pions and kaons in the final state and share the same topology; the information extracted from these sub-detectors is vital to the study of these decays.

Both detectors rely on the phenomenon of Cherenkov radiation: a charged particle travelling through an optically transparent, dielectric medium at a speed greater than the phase velocity of light in the medium will emit radiation coherently in a cone of angle θ along the path of the particle. The relation between the Cherenkov angle and the particle's speed is given by

$$\cos(\theta) = \frac{c}{nv_p}, \tag{3.1}$$

where n is the refractive index of the material, c is the speed of light in vacuum, and v_p is the speed of the charged particle. The Cherenkov angles as a function of momentum for the different charged final-state particles in LHCb can be seen in Figure 3.15. The mass of the particle can then be inferred from its momentum and the resolved Cherenkov angle.

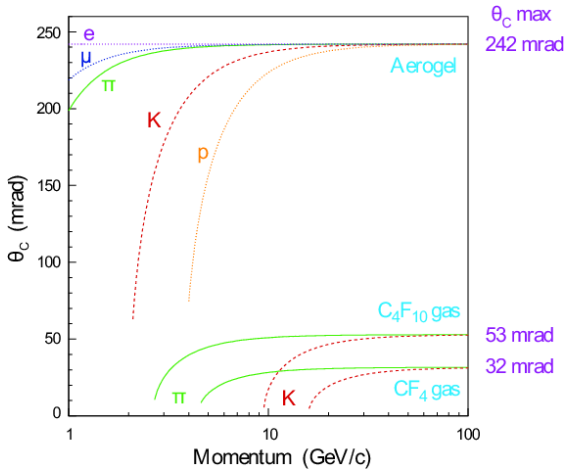

Figure 3.15: Cherenkov angle as a function of particle momentum for the different media used in the LHCb RICH detectors [73].

In both RICH detectors, the photons produced in the material are focused onto Hybrid Photon Detectors (HPDs), located outside of the detector's acceptance, by a combination of flat and spherical mirrors. These HPDs must be shielded from the fringe field of the dipole magnet to protect their performance. This is achieved by enclosing them in external iron shields. The layout of both RICH detectors is shown in Figure 3.16.

RICH1 is located immediately after the VELO and covers the full detector acceptance. It was designed to use both decafluorobutane (C_4F_{10}) and aerogel as Cherenkov radiators: the C_4F_{10} providing coverage a momentum range of $\sim 10 - 60\,\mathrm{GeV}/c$, with the aerogel covering the range below the C_4F_{10} Cherenkov threshold, $\sim 2 - 10\,\mathrm{GeV}/c$. During Run 1, due to the detector operating at higher luminosity than designed, it was found that the large number of photons produced in the aerogel deteriorated the overall particle identification (PID) in RICH1. A decision was made to remove the aerogel for the start of Run 2, resulting in a speed-up in the RICH reconstruction with no loss of PID performance [90].

The RICH2 detector is located between the tracking stations and the calorimetry system, covering a reduced acceptance range: 15-120 mrad horizontally and 15-100 mrad vertically. It is designed to cover a higher momentum range than RICH1, from $15 - 100\,\mathrm{GeV}/c$, which also justifies the reduced acceptance as low-momentum particles are swept away by the magnet. The radiator material chosen is tetrafluoromethane (CF_4).

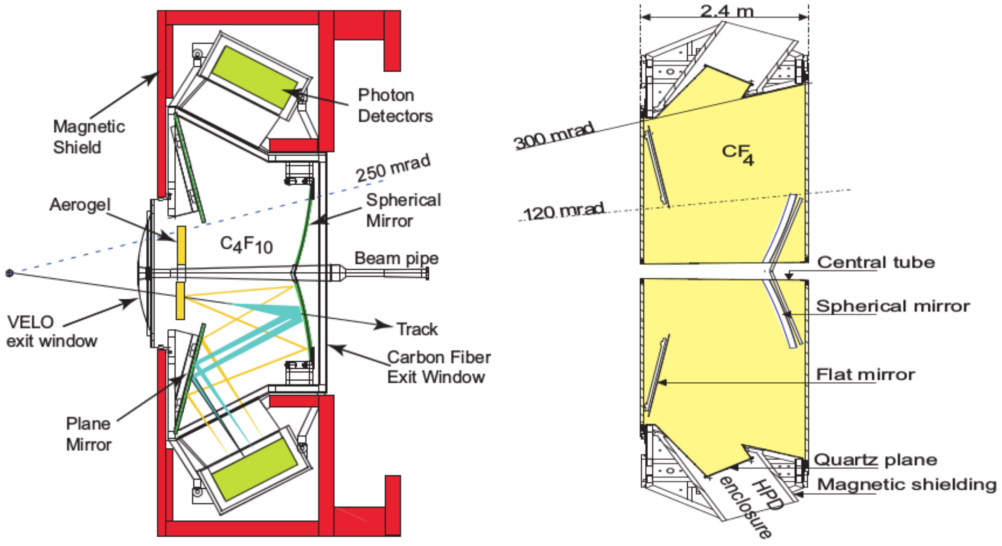

Figure 3.16: Cross-sections of the RICH1 (left) and RICH2 (right) detectors. Note that RICH1 is shown in the y-z plane, while RICH2 in the x-z plane [74].

3.4.2 Calorimetry

The calorimetry system provides identification of photons, electrons, and hadrons, through the measurement of their energies. It also performs a crucial role in the hardware stage of the trigger, allowing events of interest to be identified; the trigger is further discussed in Section 3.5. It consists of four sub-detectors located downstream of RICH2: a Scintillator Pad Detector (SPD), a Preshower Detector (PS), an Electromagnetic Calorimeter (ECAL), and a Hadronic Calorimeter (HCAL). As particles travel through the calorimeter material, energy is lost due to interactions with the detector material and a particle shower is produced. All four sub-detectors use scintillator material (doped polystyrene) as their active medium; the scintillation light produced is transported to photomultiplier tubes in the periphery via wavelength-shifting fibres. Layers of absorber material are interlaced with the scintillator in order to contain the showering particles. The hit density decreases by two orders of magnitude from the inner edge of each calorimeter to the outer edge – each calorimeter sub-detector has variable segmentation to accommodate this, as shown in Figure 3.17.

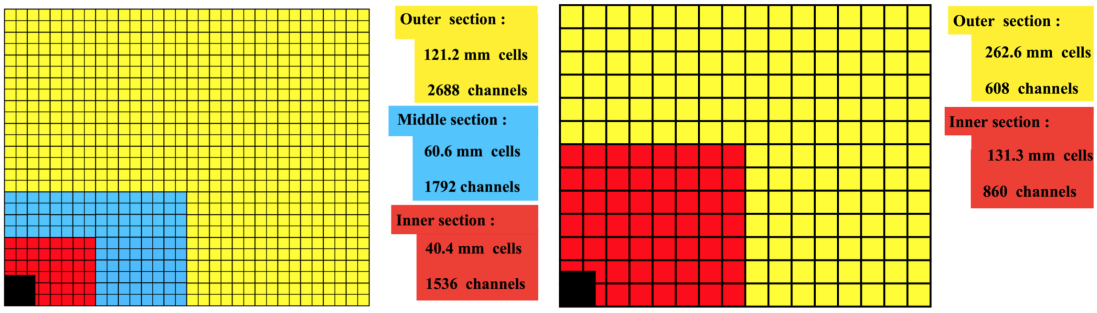

Figure 3.17: Illustration of the SPD, PS and ECAL (left) and HCAL scintillating pad regions (right). The beam pipe is shown in black [73].

The SPD/PS detectors are two high-granularity rectangular scintillator pads with a 15 mm layer of lead absorber between them. Both are segmented such that they projectively match the ECAL, this requires the SPD layer to be $\sim 0.45\%$ smaller than the PS layer. They are used in conjunction with the ECAL to provide discrimination between photons, electrons and pions. Photons will not interact and create a signal travelling through the SPD but will initiate a shower in the lead absorber layer, generating a signal in the PS. The presence of a signal in the PS but a lack of one in the SPD identifies the particle as a photon. The thickness of the lead absorber has been optimised to differentiate between electrons and charged hadrons: the radiation length of lead is much shorter than its interaction length, therefore electrons are more likely to start

showering in the PS than pions. Information from the PS alone provides a pion-rejection rate of $\sim 92\%$ for an electron-acceptance rate of 95%; including ECAL and tracking information increases the pion-rejection rate to $\sim 99\%$.

The ECAL is composed of alternating layers of $2\,\mathrm{mm}$-thick lead and $4\,\mathrm{mm}$-thick scintillator tiles, with a total of 66 layers each. The ECAL is therefore $42\,\mathrm{cm}$ deep, corresponding to 25 radiation lengths, which is enough to fully contain the electromagnetic showers produced by high-energy photons and electrons. It is designed to achieve an energy resolution of $\sigma_E/E = 10\%/\sqrt{E/\,\mathrm{GeV}} \oplus 1\%$ and a π^0 mass resolution of $\sim 8\,\mathrm{MeV}/c^2$.

The HCAL uses an iron absorber instead of lead, while using scintillator tiles as the active material. While in the ECAL the lead and scintillator layers are arranged transversely, in the HCAL the tiles run parallel to the beam axis. The primary goal of the HCAL is to be used in the hardware trigger, therefore it is not required to fully contain the produced hadronic showers. Due to the limited space available in the LHCb cavern the HCAL has a depth of $1.65\,\mathrm{m}$, corresponding to only 5.6 interaction lengths in steel. Hadronic showers tend to be larger than electromagnetic ones and so the HCAL is segmented into only two zones, as seen in Figure 3.17. The design energy resolution is $\sigma_E/E = 69\%/\sqrt{E/\,\mathrm{GeV}} \oplus 0.9\%$.

3.4.3 Muon System

Many decays of interest at LHCb involve muons in the final state, *e.g.* $B_s^0 \to \mu^+\mu^-$ [91] and $B^0 \to K^{*0}\mu^+\mu^-$ [92]. The muon system must provide fast p_T measurements for use in the hardware trigger and excellent muon identification offline. A view of the muon system is shown in Figure 3.18.

Muon identification and measurement at LHCb is performed by a system composed of five stations, M1 to M5. Muons are highly-penetrating particles and, as a result, most of the muon stations (M2–M4) are located at the end of the detector, past the calorimeters, where background from other particles is greatly reduced. These four stations are interspaced with $80\,\mathrm{cm}$ thick iron shielding plates to further decrease background. The first muon station, M1, is positioned upstream of the calorimetry system to improve the p_T measurement used in the hardware trigger.

Each station is divided into four concentric regions (R1–R4) such that the occupancy remains roughly constant between regions; the linear dimensions of each station scale projectively with its distance to the interaction point. The muon system, with the exception of the inner region in M1, is composed of Multi-Wire Proportional Counters (MWPCs). In M1R1, where the particle flux is highest, triple-Gas Electron Multipliers (GEM) are used instead.

The first three stations are more granular than M4–M5 and are used to determine the muon track direction, providing a p_T resolution of 20% in the bending plane. M4 and M5 have

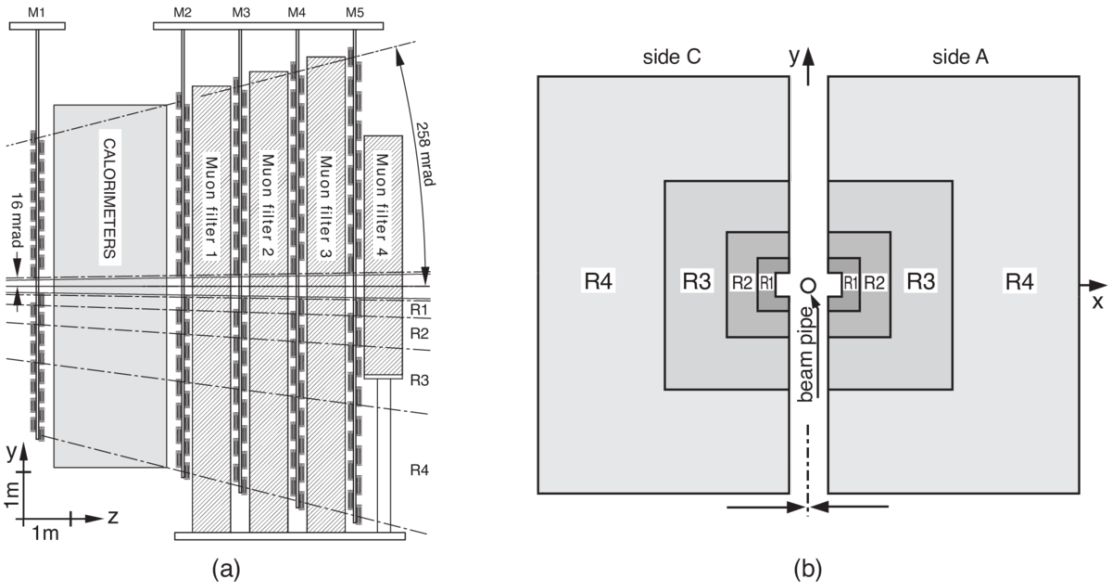

Figure 3.18: Side view of the LHCb muon system (left) and layout of the four concentric regions in one of the stations (right) [93].

limited spatial resolution and are used to identify penetrating particles. The minimum momentum required for a muon to travel through the whole detector up to M5 is 6 GeV/c.

3.4.4 PID Performance

The information gathered by the RICH, the calorimeters, and the muon system can be combined in different ways to provide variables that can discriminate between the different species of final-state particle in LHCb. The reconstruction algorithms of each sub-detector, though very different, can all be used to construct likelihood ratios for different particle hypotheses for each track. These are combined to give an overall likelihood of a track being of a particular type, relative to its likelihood of being a pion. A second, newer, approach uses detector information as inputs to neural networks to give a classifier related to the probability of a track being a certain hypothesis and not any other hypothesis. This approach uses information not available in the likelihood calculations and takes into account correlations between the different sub-detectors.

For the purposes of this work, variables related to the (mis-)identification of kaons and pions are crucial. The performance of the particle identification at LHCb is measured using data calibration samples, from decays in which final-state particles can be identified using only kinematic information. For kaons / pions, the most appropriate calibration decay channel is

$D^{*+} \to D^0(K^-\pi^+)\pi^+$. A comparison of the kaon identification and $\pi \to K$ mis-identification efficiencies between the 2012 and 2016 data-taking periods can be seen in Figure 3.19.

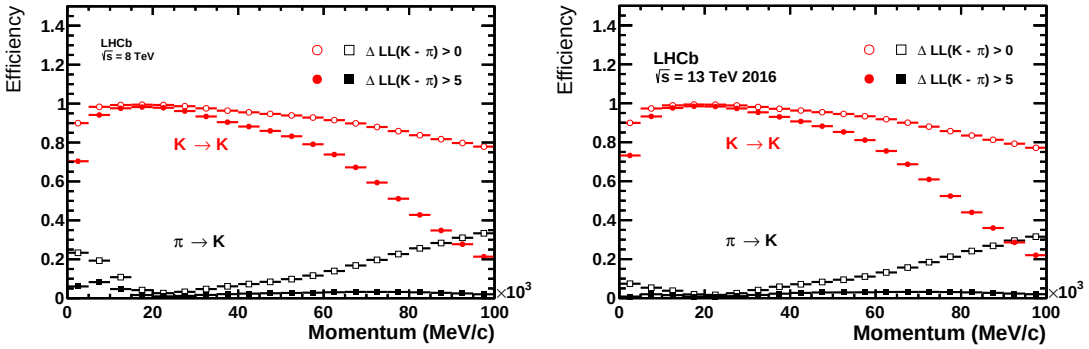

Figure 3.19: Kaon ID efficiency (red) and $\pi \to K$ mis-ID efficiency (black) as a function of track momentum, in 2012 (left) and 2016 (right). Two different PID requirements are shown. An overall improvement in PID efficiency can be seen in 2016 versus 2012 [85].

Muon identification is also considered, as vetoing muon candidates is an important step in reducing backgrounds for the decay channels studied. A binary selection can be performed on muon candidates, based on the penetration of the candidate in the muon system [94]. This low-level requirement is sufficient to provide good discrimination between muons and charged hadrons, as shown in Figure 3.20.

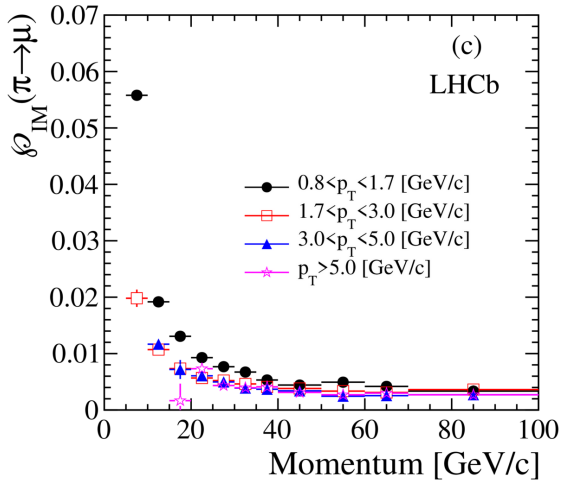

Figure 3.20: $\pi \to \mu$ mis-ID efficiency as a function of track momentum, for different p_{T} requirements [94].

3.5 Trigger

The nominal LHC collision rate of 40 MHz translates to an effective frequency of interactions at LHCb of about 10 MHz, due to gaps in the beams and also accounting for the fraction of pp collisions that are inelastic. However, this rate is orders of magnitude higher than what can stored for physics analysis. A trigger system is therefore crucial to reduce the event rate, while ensuring that events of interest are kept. The current LHCb trigger system is formed of a hardware level-0 (L0) stage, and two software High Level Trigger stages (HLT1, HLT2) [95]. An overview of the LHCb trigger and the differences in its running between Run 1 and Run 2 are shown in Figure 3.21. The output rate of the LHCb trigger was 3 kHz in 2011, 5 kHz in 2012, and 12.5 kHz in Run 2 [96].

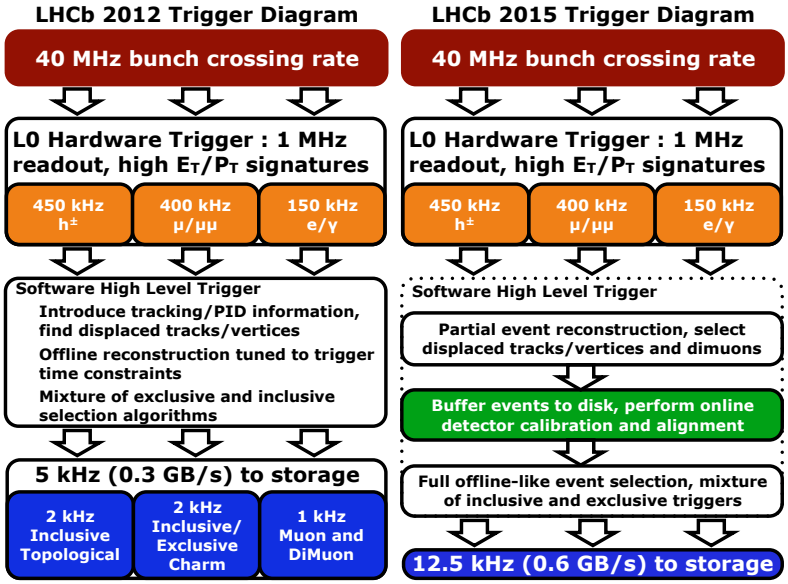

Figure 3.21: Overview of the LHCb trigger in Run 1 (left) and Run 2 (right) [85].

The L0 trigger is implemented with the use of custom electronics, which operate synchronously with the 40 MHz bunch crossing frequency. Its decisions are based on information from the calorimeter and muon systems and are used to reduce the event rate to 1 MHz, the point at which the full detector can be read out. The calorimeter trigger searches for deposits with large transverse energy (E_{T}): hadron candidates required a HCAL cluster (including an associated ECAL cluster) with minimum E_{T} of 3.5 (3.7) GeV in 2011 (2012); photon candidates required an ECAL cluster with $E_{\mathrm{T}} > 2.5$ (3.0) in 2011 (2012), with one or two associated PS hits but no associated SPD hits; electron candidates have the same requirements as photon candidates, but

with at least one associated SPD hit. The muon trigger searches for the two muon candidates with the largest p_{T} in each quadrant. A single muon candidate is selected if any single track had $p_{\mathrm{T}} > 1.48\ (1.76)\,\mathrm{GeV}/c$ in 2011 (2012); a dimuon candidate is selected if the product of the p_{T} of any two tracks is higher than 1.3 (1.6) GeV^2/c^4. The threshold values during the Run 2 data-taking periods can be found in Ref. [89]. An additional requirement is also set on the total number of SPD hits in an event, to remove events that would take too long to process in the HLT.

Events accepted by the L0 are transferred to a computer system called the Event Filter Farm (EFF), consisting of roughly 1000 nodes during the Run 1 period with an extra 800 nodes added for Run 2. The EFF can run $\mathcal{O}(10^4)$ instances of the HLT algorithm in parallel.

At the first software stage, HLT1, partial event reconstruction is performed to reduce the event rate to $50-110\,\mathrm{kHz}$. Tracks are reconstructed in the VELO and are used to identify primary vertices. Vertices are reconstructed from a minimum of five VELO tracks and those vertices within 300 $\mu\mathrm{m}$ of the average position of the pp interaction are considered primary vertices [97]. The VELO tracks are then extrapolated to the TT to form upstream tracks and to the T-stations to form long tracks. The IP and p_{T} of tracks are then calculated; the inclusive b-hadron trigger line selects events containing at least one track with high p_{T} and sufficiently displaced from any PV.

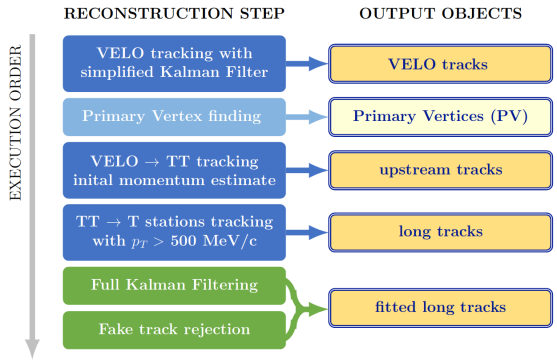

Figure 3.22: The HLT1 reconstruction algorithm chain [89].

HLT2 performs full event reconstruction on all events passing HLT1. The lower input rate allows the trigger selection to be more flexible than the two previous stages. Two types of 'trigger line' are defined at this level: exclusive trigger lines are optimised for specific final states, while inclusive lines have generic topological requirements on the final-state particles. The work presented in this thesis relies on the 'topological' trigger lines: these are designed to trigger on partially reconstructed b-hadron decays, with a displaced decay vertex and at least two charged particles in the final state. Tracks are selected due to their fit quality and impact parameter; exclusive trigger lines containing electrons or muons also particle identification information. The topological lines use a multivariate algorithm, based on the kinematics of the tracks associated

with the proto-candidate [98, 99].

The HLT has gone through significant changes since 2011. During that year, the EFF was only in use during periods of active data-taking, *i.e.* during fills, corresponding to only 30% of the available time. In order to use the idle time of the EFF during the 2012 data-taking period, roughly 20% of events passing the L0 trigger were stored on local disks, to be further processed by the HLT at a later time. This strategy led to an effective increase of approximately 25% in the HLT processing rate. For Run 2, this idea was expanded further and the two HLT stages were made fully asynchronous. HLT1 still operates synchronously with the L0 stage, but all events that pass this stage are sent to a temporary 10-petabyte buffer to be later processed by HLT2. This has allowed dedicated HLT1 lines to be used to calibrate the detector in real time, while previously this could only be performed after the HLT2 stage. The new strategy allows the HLT2 reconstruction to perform as effectively as the 'offline' reconstruction. An additional system, the Turbo stream [100], was created in Run 2 to select events directly from the output of the trigger.

CHAPTER 4

Relative Branching Fractions of $B^+ \to h^+h'^+h''^-$ decays

“ It will never be possible by pure reason to arrive at some absolute truth. ”

Werner Heisenberg, *Physics and Philosophy*

The following chapters detail a measurement of the ratios of branching fractions of the four $B^+ \to h^+h'^+h''^-$ decay channels, *i.e.* the decay channels $B^+ \to K^+K^+K^-$, $B^+ \to K^+\pi^+K^-$, $B^+ \to K^+\pi^+\pi^-$, and $B^+ \to \pi^+\pi^+\pi^-$. The ratios are determined for each channel relative to the other three $h^+h'^+h''^-$ modes. The LHCb experiment has set the most stringent limits on the branching fractions of the decays $B^+ \to K^+K^+\pi^-$ and $B^+ \to \pi^+\pi^+K^-$ [101]; these heavily-suppressed decays are not included in this analysis.

The measurement is performed with the combined Run 1 dataset collected by LHCb, corresponding to a total integrated luminosity of $3\,\text{fb}^{-1}$. As previously stated, the inclusion of charge-conjugate processes is implied throughout, unless otherwise specified.

4.1 Measurement Strategy

The collision environment of the B factories allowed for the total number of B mesons produced to be measured precisely [102]. In contrast, the hadronic environment at the LHC limits the precision available on luminosity and B^+ cross-section measurements. This motivates the choice of measuring relative, as opposed to absolute, branching fractions at LHCb. Furthermore, there

are benefits to measuring branching fraction ratios; for example, a number of potential systematic effects are expected to cancel to first order.

The branching fraction ratios (*e.g.* relative to the KKK channel) are given by

$$\frac{\mathcal{B}(B^+ \to h^+ h'^+ h''^-)}{\mathcal{B}(B^+ \to K^+ K^+ K^-)} = \frac{\mathcal{N}_{hhh}}{\mathcal{N}_{KKK}} \times \frac{\bar{\epsilon}_{KKK}}{\bar{\epsilon}_{hhh}}, \tag{4.1}$$

where $\mathcal{N}$ is the observed signal yield, $\bar{\epsilon}$ is the phase-space-averaged selection efficiency, and hhh labels any of the other three decay modes.

The signal yields of the four channels are extracted from an unbinned extended maximum likelihood fit performed simultaneously to the four hhh invariant mass distributions. Each signal channel is a potential source of background for the others, due to candidates misdentified by the PID system. Performing the invariant mass fits simultaneously allows for the yields of such backgrounds to be constrained. The fit model is constructed by considering this and other sources of possible backgrounds, detailed in Chapter 5. Monte Carlo (MC) simulated data samples are used to obtain the probability density functions (PDFs) that compose the fit model.

Prior to the fit procedure, the dataset goes through a number of selection steps aimed at reducing the total number of background candidates present, while minimising the loss of signal candidates. The total efficiency of the selection procedure is determined with a combination of MC simulated signal samples and calibration data. Once the signal yields, the total selection efficiencies, and their distribution as a function of the phase space are determined, the branching fraction ratios can be evaluated using Equation 4.1.

One final point in this discussion of the measurement strategy concerns the inclusion / exclusion of intermediate charmonium states in the definition of signal. Amplitude analyses of charmless decays generally veto contributions via intermediate J/ψ and $\psi(2S)$ states, since their inclusion would require modelling resolution effects that could be otherwise ignored. However, other charmonium states, such as the χ_{c0}, have not been consistently removed. In this analysis, a decision was made to include all charmonium resonances in the definition of signal, since their lifetimes are negligible and the resolution effects do not need to be modelled. Contributions from intermediate D states are still considered as a background source.

4.2 Data and Monte Carlo Samples

The Run 1 dataset corresponds to data taken in 2011 and 2012 at centre-of-mass energies of $7\,\text{TeV}$ and $8\,\text{TeV}$, respectively. The B meson production cross-section is assumed to increase linearly due to the higher centre-of-mass energy. Furthermore, the 2011 data-taking period was not evenly split between the two magnet polarisation states. The fractions of the total luminosity

taken under the four running conditions are given in Table 4.1.

Table 4.1: Total integrated luminosity as a function of the running conditions during data-taking.

Year	Magnet Polarisation	Integrated Luminosity (fb^{-1})
2011	MagDown	564 ± 10
	MagUp	422 ± 7
2012	MagDown	993 ± 12
	MagUp	1001 ± 12

The selection efficiency is mostly determined using simulated $B^+ \to h^+h'^+h''^-$ events, as detailed in Chapter 6. Roughly one million signal MC candidates were generated for each of the running conditions, so that the potential variations in efficiency can be studied. As the number of generated MC candidates per subsample are roughly equal, the efficiency of the total Run 1 dataset is taken as a weighted average, with the integrated luminosity ratios used to weigh each subsample.

The signal MC samples have been generated with a model that is distributed uniformly across the square Dalitz plot. This ensures that the edges and corners of the conventional DP are well-populated, while minimising the computational load required to generate large samples. However, this approach can cause some regions of the phase-space to be overrepresented, an issue that is further discussed in Section 7.1.2.

CHAPTER 5

Backgrounds and Candidate Selection

> *Data becomes useful knowledge of something that matters when it builds a bridge between a question and an answer. This connection is the signal.*

Stephen Few, *Signal: Understanding What Matters in a World of Noise*

The raw dataset used in this analysis must go through several steps of refinement before the signal yields can be extracted. For each candidate corresponding to a true $B^+ \to h^+h'^+h''^-$ decay there are many others that approximate their signature, originating from a number of background sources. The different sources of background considered in this analysis, and the strategies and tools employed to optimise their reduction, are described in this chapter.

5.1 Background Sources

The backgrounds considered can be roughly divided into the following classes:

- **Cross-Feed Background**
  True $B^+ \to h^+h'^+h''^-$ candidates in which one or more final-state particles are misidentified appear as backgrounds in the invariant mass distributions corresponding to the other signal modes. In addition to cross-feed from other signal modes, decay channels with protons or muons in the final state, misdentified as kaons or pions, are potential sources of background.

The primary effect of the misidentification (misID) on the invariant mass distribution of cross-feed candidates is to shift it by an amount approximately proportional to the mass difference between the correctly-identified and misidentified particle species and the number of misdentified particles. The cross-feed distribution is shifted towards higher mass values in cases where the particle is misidentified as a higher-mass particle, *e.g.* $\pi \to K$, and towards low masses in the opposite case, *i.e.* $K \to \pi$. Additionally, the misidentification also smears the cross-feed distribution. This is particularly relevant to the $B^+ \to K^+\pi^+K^-$ and $B^+ \to K^+\pi^+\pi^-$ final states; the distribution of candidates with both like-sign particles misidentified is significantly different to the signal shape, even though it peaks at the correct B mass due to the mass-shift effects cancelling each other.

Cross-feed background can be very effectively reduced with the use of PID variables, constructed out of information from the LHCb PID system described in Section 3.4.4. Decay channels with final-state muons misidentified as pions are suppressed with a veto on a low-level muon PID variable. The strategy for optimising the PID selection is detailed in Section 5.3.

- **Combinatorial Background**

  Random combinations of tracks in an event can happen to form candidates that approximate signal properties well enough to pass the selection criteria. These candidates do not peak at any particular mass value but instead follow a monotonic distribution, generally decreasing in the region of the B mass. The majority of tracks in any given event originate from the primary vertex, therefore combinatorial background can be very effectively reduced by exploiting topological features of B decays, such as a displaced decay vertex, with the use of multivariate analysis (MVA). The description of the MVA used and the optimisation of its output is given in Section 5.4. Furthermore, combinatorial tracks are predominantly pion tracks and so, for the signal modes with kaons in the final state, PID requirements can also be effective at removing this type of background.

  A related background category originates from true $B^0/B_s^0 \to h^+h'^-$ decays, to which an additional combinatorial track is added during reconstruction. These 'partially combinatorial' candidates peak in the two-body invariant mass distribution at the parent particles's mass value; the candidate's three-body invariant mass is pushed towards the high-mass region of the spectrum, due to the added energy, and the whole distribution is heavily smeared. Small yields will be indistinguishable from the bulk combinatorial background but an observable excess of candidates could affect the modelling of the combinatorial components of the invariant mass fits.

- **Partially Reconstructed Background**

  Partially reconstructed backgrounds originate from four-body decays with three final-state particles in common with the signal modes, in which the fourth final-state particle has not been included in the candidate.

  The overall shape of each partially reconstructed background depends on the kinematics of the missing particle; however, a universal feature of such backgrounds is a maximum bound on the three-body invariant mass distribution, given by the mass difference between the parent particle and the missing particle, *e.g.* $m_B - m_{\text{miss}}$. Background channels originating from B^+, B^0, or B_s^0 will therefore populate the low-mass region in the vicinity of the signal peak, with decay channels where the missing particle is a pion or a photon being of most concern.

  These backgrounds can be roughly categorised as a function of the masses of the parent particle and the missing particle, and the topology of the decay. The (B^+, B^0) mass difference is negligible and, in this context, so is the difference in (π^+, π^0) masses. Therefore it is reasonable to group these together under a generic $B \to$ `4-body` category. The mass difference between (B^+, B^0) and the B_s^0 is large enough to identify $B_s^0 \to$ `4-body` decays as a separate category. The distribution of partially reconstructed decays containing an intermediate charm particle is assumed to differ sufficiently from that of charmless decays to justify their inclusion as a separate category. Finally, decays with a missing photon form a distinct category since the mass threshold extends to the mass of the parent particle.

- **Intermediate Charm Backgrounds**

  This category of background consists of $B^+ \to Dh^+$ decays, *i.e.* charmed decays, where the D decays to $h'^+h''^-$. Candidates originating from decays via an intermediate charmonium state, *e.g.* $B^+ \to J/\psi(\to h'^+h''^-)h^+$, are counted as signal candidates in this analysis, as discussed in Chapter 4. The distribution of charmed candidates with the same final-state particles as a signal channel will peak directly underneath the signal peak. These candidates are suppressed by including goodness-of-fit information on the B decay vertex, which is constructed from all three tracks, in the MVA training. The most effective approach for removing the remaining candidates is an explicit veto on the two-body invariant mass region around the D peak.

  Charmed backgrounds can also suffer from misidentified final-state particles. In cases where the misidentified particle stems from the D meson, the resulting shift in the mass distribution causes the bulk of these candidates to evade the veto. The PID requirements, optimised for $B^+ \to h^+h'^+h''^-$ cross-feed, can potentially be less effective against this

type of background, given the different topology of $B^+ \to Dh^+$ decays and their larger branching fractions.

5.2 Pre-Selection Steps

Prior to the optimisation, a number of general selection cuts are applied to all signal candidates to ensure that good-quality, relevant events are selected. The general goal at this stage is to select events with well-reconstructed tracks and vertices, that are likely to contain a B meson that decays hadronically.

5.2.1 Stripping Requirements

The total dataset available from the output of the trigger system is too large to be analysed efficiently by every LHCb member working on physics analyses. This problem is overcome by running centrally a number of user-defined sets of loose selection criteria, known as *stripping lines*. The output of similar stripping lines are gathered into *streams*, such that events shared between lines in a particular stream are only stored once. The name of the stripping line used in this analysis is `StrippingBu2hhh_KKK_inclLine`, in the `Bhadron` stream; a summary of the cuts applied at this level is given in Table 5.1. Only information related to the signal candidate and some select information about the event as a whole is saved in this stream, reducing storage requirements.

Tracks are required to have traversed the entire tracking system, *i.e.* tracks must be 'long' (different track types have been defined in Section 3.3.6), with some minimum momentum cuts applied. These tracks must be of good quality, defined by a low χ^2 value resulting from the track-fitting procedure. The probability of the track being made of partial trajectories from more than one charged particle, *i.e.* the probability of a 'ghost' track, must also be low. Furthermore, the total number of long tracks in an event is limited to under 200 tracks.

At the stripping level, all final-state particles are reconstructed as kaons, with the reconstruction of the other mass hypotheses being done offline. Consequently, no particle identification requirements are applied at this level. Only very loose cuts on the invariant mass of the B candidates are applied, allowing the other signal final states to be reconstructed without loss of efficiency.

Three tracks are combined by requiring that the sum of their four-momenta is loosely in the region of the B invariant mass. The minimum distance between their trajectories must also be small, *i.e.* the three tracks must roughly have the same origin. These two criteria are applied to reduce the number of candidates that undergo the vertex fit. Following this fit, additional cuts are

applied to ensure the goodness-of-fit of the B candidate decay vertex and its separation from the primary vertex (PV) with which it is most closely associated. The associated PV is defined as the primary vertex that returns the smallest χ^2_{IP} value for a given B candidate.

Table 5.1: Selection requirements in the `StrippingBu2hhh_KKK_inclLine` stripping line.

Particle	Stripping line requirements	Description
hhh tracks	$p_{\text{T}} > 100\,\text{MeV}/c$	Track transverse momentum
	$p > 1500\,\text{MeV}/c$	Track momentum
	$\chi^2_{\text{IP}} > 1$	Minimum χ^2 distance of particle's trajectory to PV
	χ^2/NDOF < 4	Maximum χ^2 per degree of freedom of the track fit
	GHOSTPROB < 0.5	Max. ghost track probability
hhh combination cuts	$5.05 < m(KKK) < 6.3\,\text{GeV}/c^2$	Total invariant mass of the final-state particles
	DOCA $< 0.2\,\text{mm}$	Distance-of-closest-approach between any two child tracks
Parent B cuts	Lead $p_{\text{T}} > 1500\,\text{MeV}/c$	p_{T} of the track with the highest p_{T}
	DIRA $(\theta) > 0.99998$	Cosine of angle between B candidate momentum and direction from the associated PV to the decay vertex
	Flight $\chi^2 > 500$	χ^2 distance of B candidate from associated PV
	Flight distance $> 3\,\text{mm}$	Separation between the decay vertex and closest PV
	Vertex $\chi^2 < 12$	χ^2 of the B decay vertex
	$\chi^2_{\text{IP}} < 10$	IP χ^2 with respect to the associated PV
	$p_{\text{T}} > 1000\,\text{MeV}/c$	p_{T} of the B candidate
	$\sum p_{\text{T}} > 4500\,\text{MeV}/c$	Total p_{T} of the B children
	$\sum p > 20000\,\text{MeV}/c$	Total p of the B children
	$\sum \chi^2_{\text{IP}} > 500$	Total IP χ^2 of tracks with respect to the associated PV
	$4 <$ BPVCORRM $< 7\,\text{GeV}/c^2$	Corrected mass [†] under KKK mass hypothesis

[†] The corrected mass is defined as BPVCORRM $= \sqrt{M^2 + |P_{\text{T}}^{\text{miss}}|^2} + |P_{\text{T}}^{\text{miss}}|$, where M is the B candidate mass and $P_{\text{T}}^{\text{miss}}$ is the missing transverse momentum.

Further decay tree fits are performed offline using the DecayTreeFitter (DTF) package [103], which parameterises the whole signal decay tree as a function of the final-state momenta and vertex properties and simultaneously fits them. It is at this point that the final-state particles are reconstructed under all four $B^+ \to h^+h'^+h''^-$ hypotheses, with all DTF parameters being recalculated. Another fit is also performed, in which the masses of the B and the final-state particles are constrained to their current world-average values and the B is constrained to originate from the associated PV. The Dalitz plot variables (both conventional and square) are extracted from this fit to ensure all candidates share the same DP definition and are constrained within its boundaries.

5.2.2 Trigger Requirements

Post reconstruction, tracks can be associated with the energy deposits that have caused the L0 trigger to fire. In this manner, the particle(s) that have fired the trigger are identified and can be clas-

sified under the particular L0 trigger fired (`L0Hadron`, `L0Muon`, `L0DiMuon`, `L0Electron`, `L0Photon`). If the track that fired the trigger originates from the signal candidate, the event is classified as *trigger on signal* (TOS). Otherwise, if the trigger decision originates from the rest of the event, it is classified as *trigger independent of signal* (TIS). The requirements imposed in this analysis are that signal candidates have fired the L0 hadron trigger (`L0Hadron_TOS`) or that any L0 trigger is fired by the rest of the event (`L0Global_TIS`). Events where activity related to both the signal and the rest of the event was required to cause the trigger are rejected.

At the HLT1 level, signal candidates must pass the `Hlt1TrackAllL0Decision` line, which requires one good-quality, high p_{T} track that is significantly displaced from the primary vertex. Finally, at the HLT2 level, the B candidate is required to pass either the 2-body or 3-body topological trigger `Hlt2Topo{2,3}BodyBBDTDecision`, discussed in Section 3.5.

5.2.3 Fiducial Cuts

The reconstruction in the RICH detectors becomes less efficient for tracks outside a specific kinematic range. The PID efficiency plots in Section 3.4.4 show that for track momenta above $100\,\mathrm{GeV}/c$, the efficiency drops significantly while the misidentification rate steadly rises. Similar behaviour is seen for tracks with very low/high pseudorapidity. Furthermore, the ring reconstruction in the RICH depends on the total number of charged tracks in a given event; the efficiency drops in events with large occupancies.

A number of 'fiducial cuts' are applied as a part of the selection, to ensure that the data+MC samples used have good quality PID information. Each track is required to be within the momentum range $1.5 < p < 100\,\mathrm{GeV}/c$ and the pseudorapidity range $2.0 < \eta < 5.0$. The event occupancy is parameterised by using the `BestTracks` variable, defined as the total number of long, upstream, and downstream tracks in the event (with only upstream/downstream tracks not used to reconstruct long tracks being counted). All events are required to have an occupancy lower than $\texttt{BestTracks} < 500$.

The event occupancy is known to not be well modelled in MC; it is generally lower than in real data. This issue is overcome by resampling the MC `BestTracks` distribution using the corresponding distribution in *sWeighted* $B^+ \to K^+K^+K^-$ data. The cumulative distribution function (CDF) of the `BestTracks` variable is taken for both MC and *sWeighted* data. Each value in the MC `BestTracks` distribution is associated with the value in the data distribution that shares the same value in the CDF. It is that data `BestTracks` value that is given to all MC events that have the original value in the MC distribution. The effect of the resampling on the `BestTracks` distribution in $B^+ \to K^+K^+K^-$ MC is given in Figure 5.1.

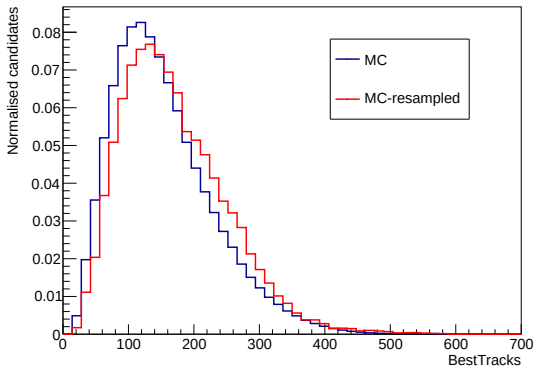

Figure 5.1: Distribution of `BestTracks` before (blue) and after (red) resampling for the 2012 MagDown $B^+ \to K^+K^+K^-$ MC sample.

5.3 PID Selection

The information extracted from the LHCb PID subsystems is used as input to neural networks whose outputs are used, in an one-against-all approach, to classify tracks as being a certain particle hypothesis. These PID variables are named `PROBNNx` in LHCb and throughout this text, with `x` $\in$ {`K`, `pi`, `mu`, *etc.* } depending on the particular hypothesis. Each `PROBNN` variable ranges from 0 to 1, but these are not strictly probabilities, in particular the values of the different `PROBNN` for any given track need not add up to one. It is possible to combine `PROBNN` variables in order to simultaneously provide both positive discrimination in favour of a particular particle type and negative discrimination against another.

The four signal channels considered are composed of only kaons and pions in the final state, with $\pi \leftrightarrow K$ misidentification being the most common type of misID. Therefore it is beneficial to devise a PID selection, such that for kaons (pions) it selects tracks that have high probability of being kaons (pions) and, at the same time, low probability of being pions (kaons). A reasonable proposal is to investigate the distribution of kaons and pions in the `PROBNN` parameter space and exploit any useful features. To first order it is sufficient to look at the 2D (`PROBNNpi`, `PROBNNK`) plane, shown in Figure 5.2.

As expected, the kaon distribution peaks around (0,1), where `PROBNNK` is high and `PROBNNpi` is low, and the pion distribution peaks around (1,0), where the opposite is true. However, the shape of the two distributions are quite different: kaons are preferentially distributed along the edges of the parameter space, while pions roughly follow a radial distribution around (1,0) with a long tail stretching along the x axis. The use of circular cuts, centred around the point (1,0), was seen as a natural choice in this scenario. One circle is of course enough to

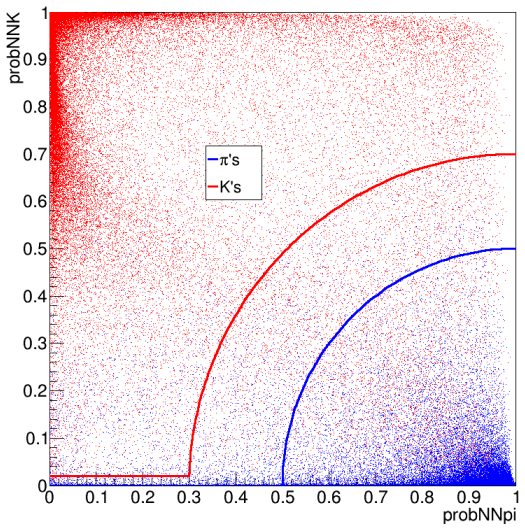

Figure 5.2: Distribution of kaons (red) and pions (blue) in the PROBNN parameter space. Both samples originate from truth-matched $B^+ \to K^+\pi^+K^-$ MC candidates. An illustration of the chosen topology for the PID cuts used is overlaid.

define a boundary between the two populations, however it has been chosen to use two circles, defined by their radii r_K and r_π. The two circles allow for more freedom in the optimisation procedure, *e.g.* naturally permitting the exclusion of regions in the centre of Figure 5.2 where both hypotheses are roughly equally likely. The form of the cuts is therefore

$$\text{PID}_K : (\texttt{PROBNNpi} - 1)^2 + \texttt{PROBNNK}^2 > r_K^2, \tag{5.1}$$

$$\text{PID}_\pi : (\texttt{PROBNNpi} - 1)^2 + \texttt{PROBNNK}^2 < r_\pi^2, \tag{5.2}$$

with the only constraint on the radii being $r_K > r_\pi$, to avoid overlapping definitions. An additional cut is required to exclude the long pion tail from the kaon acceptance region, which is implemented as a horizontal cut at $\texttt{PROBNNK} > 0.02$.

The PID response in MC is known to not replicate what is seen in data, in part due to the occupancy not being well modelled (as described in the previous section). This affects the efficiency of the PID variables and therefore MC cannot be used to evaluate this efficiency accurately. An alternative, data-driven approach is the use of the PIDCalib package [104]. It makes use of large, clean samples of final-state particles that are reconstructed without the use of the PID system, *i.e.* with kinematic information only. Both the samples of kaons and pions originate from the CKM-favoured decay $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$. The charge of the slow pion

from the D^{*+} decay tags the flavour of the $\overline{D}$ meson, with the decay products of that $\overline{D}$ making up the `PIDCalib` samples.

The efficiency of any cut on the available PID variables can be determined for these calibration samples and parameterised in terms of the kinematics of the tracks. In this analysis, the PID response is parameterised in terms of track momentum and pseudorapidity and event multiplicity (using the `BestTracks` variable), with a user-defined binning scheme given in Table 5.2. The efficiency of the bin in (p, η, `BestTracks`) space in which a MC track falls in is assigned to it as a weight, with a per-candidate efficiency weight being defined as the product of the efficiencies of its three children.

Table 5.2: Binning scheme used in the `PIDCalib` procedure.

Variable	Bin boundaries
p	$[0; 5500; 11500; 13400; 16500; 19200; 22000; 24800;$ $27800; 31000; 34000; 38000; 42000; 46000; 50000;$ $54000; 59000; 64000; 70000; 76000; 83000; 90000; 100000]$
η	$[2.0; 2.40; 2.55; 2.65; 2.75; 2.85; 2.95;$ $3.05; 3.15; 3.30; 3.45; 3.75; 4.10; 5.0]$
`BestTracks`	$[0; 100; 200; 300; 500]$

It is these per-candidate efficiencies that are used to optimise the PID selection. The optimisation is performed by maximising a figure of merit (FoM) that is a function of signal and cross-feed efficiencies and takes the form

$$\mathrm{FoM}_{\mathrm{PID}} = \frac{\epsilon_{\mathrm{sig}}}{\sqrt{\epsilon_{\mathrm{sig}} + \sum a_{\mathrm{cf}}\epsilon_{\mathrm{cf}}}}, \tag{5.3}$$

where ϵ_{sig} and ϵ_{cf} are the efficiencies of signal and cross-feed channels for a given PID cut, and $a_{\mathrm{cf}} = \mathcal{B}_{\mathrm{cf}}/\mathcal{B}_{\mathrm{sig}}$ is the ratio of branching fractions. The current world-average values are used here for the branching fractions, given in Table 2.1. Cross-feed backgrounds with one or two misidentified tracks are included in the optimisation.

The figure of merit is slightly modified for the $B^+ \to K^+\pi^+K^-$ channel. It has the smallest branching fraction out of the four signal modes and it sees cross-feed contributions on both sides of the signal peak. Partially reconstructed backgrounds with a $K\pi\pi X$ final state are of concern as they would usually populate the low-mass sideband but a $\pi \to K$ misID shifts the distribution towards the signal region. The smearing caused by the misidentification makes backgrounds of this sort difficult to model. Instead, a factor of 2 is introduced in the cross-feed term of the FoM to loosely account to the existence of both $K\pi\pi$ and $K\pi\pi X$ misID,

i.e. $\sum a_{\rm cf}\epsilon_{\rm cf} \to 2\sum a_{\rm cf}\epsilon_{\rm cf}$.

The figure of merit is optimised independently for each signal mode, by scanning through PID_K and/or PID_π. The optimisation is performed independently for each 2011/2012, MagDown/MagUp subsample as a check of the stability of the result. It should be noted that the FoM is a function of only PID_K (PID_π) in the $B^+ \to K^+K^+K^-$ ($B^+ \to \pi^+\pi^+\pi^-$) case, but it is a function of both variables in the other two cases. Figure 5.3 shows the PID FoMs obtained for the 2012 MagDown subsamples; little variation is found between different year+polarisation subsamples of the sample signal MC.

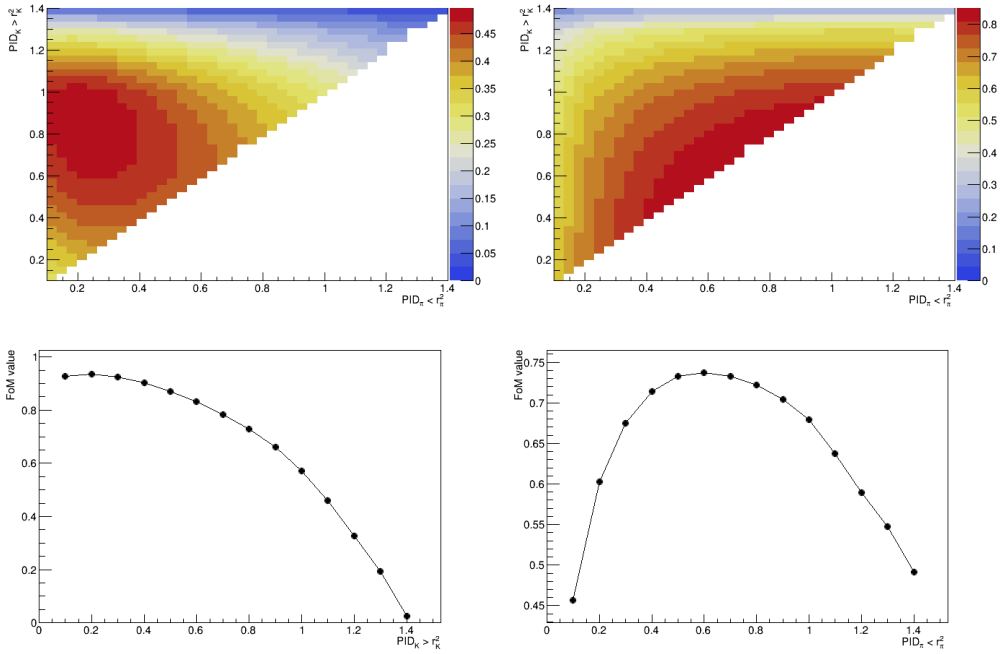

Figure 5.3: PID figures of merit for the 2012 MagDown subsamples of $B^+ \to K^+\pi^+K^-$ (top-left), $B^+ \to K^+\pi^+\pi^-$ (top-right), $B^+ \to K^+K^+K^-$ (mid-left), and $B^+ \to \pi^+\pi^+\pi^-$ (mid-right) signal MC.

The optimal PID cuts given by the peak of the figure of merit vary greatly between signal modes, as shown in Table 5.3. However, one additional constraint is required to ensure that the definition of kaons and pions are also mutually-exclusive across final states. A pair of baseline cut values is optimised by summing the individual FoMs, weighted by the square root of the respective signal branching fraction, and taking the point at which the combined FoM is highest.

The combined FoM takes the form

$$\mathrm{FoM}_{\mathrm{PID}}^{\mathrm{combined}} = \sum_{i \in \{hhh\}} \sqrt{\mathcal{B}_i}\ \mathrm{FoM}_i. \tag{5.4}$$

The baseline cuts obtained serve as maximum boundaries for the kaon/pion acceptance regions, *i.e.* the baseline replaces the individually optimised cuts whenever that cut would be looser than the baseline. The values obtained for the baseline cuts are $(r_\pi, r_K) = (0.5, 0.5)$. The baseline cuts and the final, adjusted cuts for each signal mode are also given in Table 5.3.

Table 5.3: List of PID cuts, both individually optimised and adjusted for mutual-exclusivity.

Signal mode	Optimal PID cut (r_π, r_K)	Adjusted PID cut (r_π, r_K)
$B^+ \to K^+K^+K^-$	(–, 0.2)	(–, 0.5)
$B^+ \to K^+\pi^+K^-$	(0.2, 0.9)	(0.2, 0.9)
$B^+ \to K^+\pi^+\pi^-$	(0.7, 0.7)	(0.5, 0.7)
$B^+ \to \pi^+\pi^+\pi^-$	(0.6, –)	(0.5, –)
Combined	(0.5, 0.5)	(0.5, 0.5)

5.3.1 Muon Misidentification

A low-level muon PID variable, `IsMuon`, is constructed by checking whether a particular track can be associated with any hits in the muon chambers [94]. Backgrounds with muons in the final state that could possibly be misidentified as pions or kaons by the other PID systems are effectively suppressed with a veto on the `IsMuon` variable on each final-state particle. The decay $B^+ \to J/\psi(\mu^+\mu^-)K^+$ is identified as a background of concern, since its branching fraction (6×10^{-5}) is slightly larger than the branching fraction of $B^+ \to K^+\pi^+\pi^-$. The efficiency of the muon veto is investigated using MC and it is found that fewer than 0.1% of $B^+ \to J/\psi(\mu^+\mu^-)K^+$ candidates survive the `IsMuon` veto, while roughly 68% of $B^+ \to K^+\pi^+\pi^-$ candidates pass this requirement. Therefore, the contribution from misdentified muons is assumed to be negligible after the `IsMuon` veto is applied.

5.4 MVA Selection

A multivariate algorithm is, in almost all cases, much more efficient in reducing combinatorial background than a simple cut-based selection, since it can take advantage of correlations between

its input variables. This analysis makes use of the NEUROBAYES neural network (NN) algorithm for this purpose [105].

Among the current branching fraction measurements of the $B^+ \to h^+h'^+h''^-$ channels, that of the $B^+ \to K^+\pi^+K^-$ decay has the largest uncertainty. Hence, the MVA selection is optimised with respect to this channel. The shared topology of the $B^+ \to h^+h'^+h''^-$ channels should allow this MVA to be applied to the other channels with almost no loss in efficiency. This assumption is validated later by comparing the efficiency of MVAs trained with the other signal channels.

The neural network is trained using the $B^+ \to K^+\pi^+K^-$ MC sample as the signal proxy and the high-mass sideband in the $B^+ \to K^+\pi^+K^-$ data spectrum as the background proxy, in both cases combining the 2011 and 2012 samples. The sideband is defined here as the candidates in the region $5600 < m_{K\pi K} < 6300\,\mathrm{MeV}/c^2$, a region where no significant $K\pi\pi$ cross-feed is expected. The selection requirements described above have been applied to the input samples to guarantee that the distributions of the input variables are as accurate as possible at this point.

Seven variables are chosen as inputs to the neural network; these are listed in Table 5.4 in order of their discriminating power. The choice of variables is made following a number of criteria:

- They must show good discriminating power. At first, this is checked simply by looking at the distribution of the variable in both the signal and background samples, given in Figures 5.4 and 5.5. Some variables pass this first requirement but are found to not add any further power when combined with others and therefore have been dropped prior to the definitive MVA training.

- The distribution of the input variables in MC should be in agreement with the corresponding distributions in *sWeighted* data. The MVA should use differences in the topology of signal and background candidates to achieve separation, not artefacts caused by mismodelling in the MC. The *sWeighted* distributions are also shown in Figures 5.4 and 5.5. Two variables, `Bu_PTASYM_2_0` and `Bu_MINIPCHI2`, do show some mismatch between MC and *sWeighted* data. The mismatch will be taken as source of systematic uncertainty, in order not to lose the separation power of these variables.

- The input variables should not be correlated with the hhh invariant mass nor the Dalitz plot variables.

The topology of the network is of one hidden layer, with 7 nodes in it. No significant change in performance is seen by varying the size of the hidden layer. The neural network is found to perform better when the numbers of signal and background candidates are of the same

Table 5.4: Input variables to the MVA in order of their importance.

Rank	Input variables	Description
1	`Bu_POINTING`	$\lvert p_B\rvert \sin(\theta_{\mathrm{DIRA}}) / \left(\lvert p_B\rvert \sin(\theta_{\mathrm{DIRA}}) + \sum_{\mathrm{tracks}} p_{\mathrm{T}}\right)$
2	`Bu_PTASYM_2_0`	p_{T} asymmetry for a cone of radius 2.0 rad in the η–ϕ plane
3	`Bu_MINVDCHI2`	χ^2 of the distance between primary and secondary vertices
4	`Bu_ENDVERTEX_CHI2NDOF`	Decay vertex χ^2 per degree of freedom
5	`lead_MINIPCHI2`	IP χ^2 of the track with the highest p_{T}
6	`lead_P`	Momentum of the track with the highest p_{T}
7	`Bu_MINIPCHI2`	IP χ^2 of the B candidate

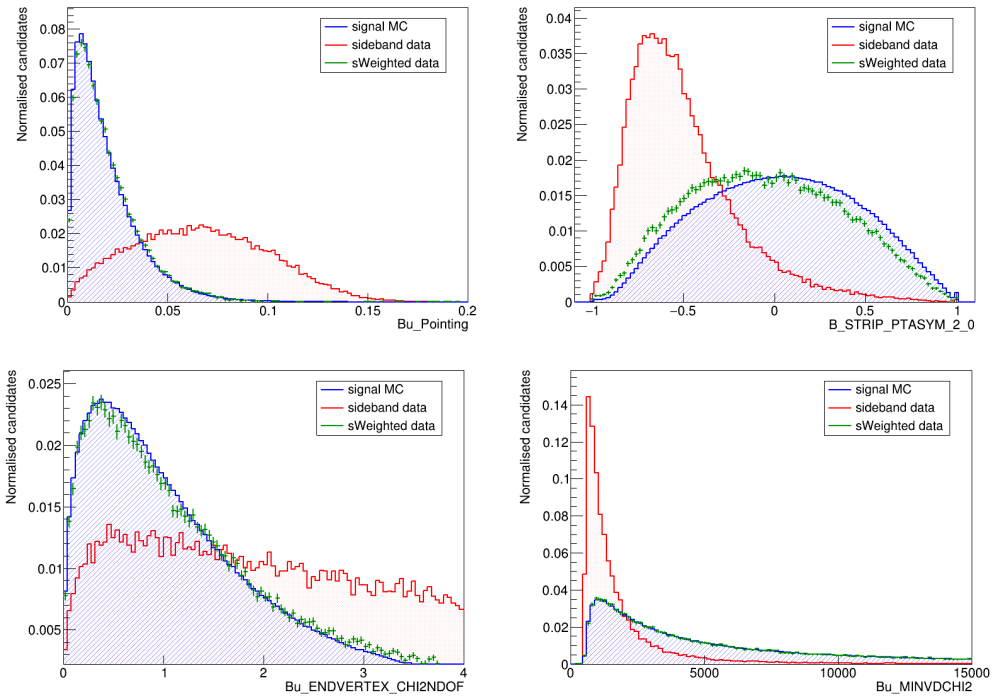

Figure 5.4: Distribution of the MVA input variables for MC (blue), *sWeighted* data (green), and sideband data (red).

order of magnitude, therefore a limited, but still significant, percentage of the available MC is used. The exact numbers used for both samples are given in Figure 5.6. Out of the total signal+background sample, 70% of candidates are used in the training stage and the remaining 30% in the test stage. The output of the classifier as applied to the test sample is shown in Figure 5.6 and good separation between signal and background can be seen.

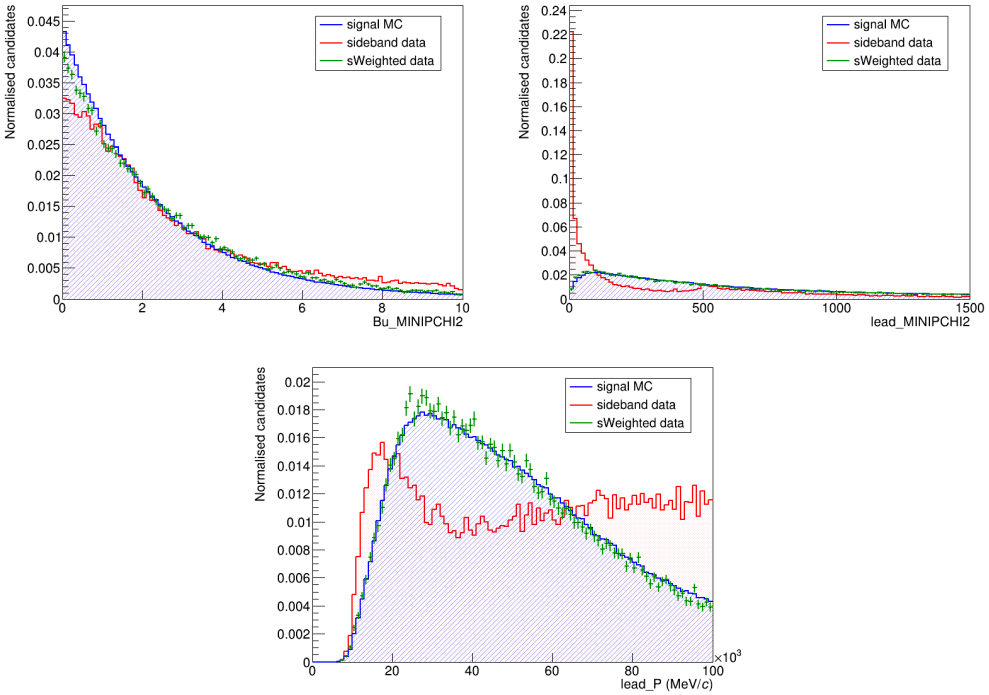

Figure 5.5: Distribution of the MVA input variables for MC (blue), *sWeighted* data (green), and sideband data (red).

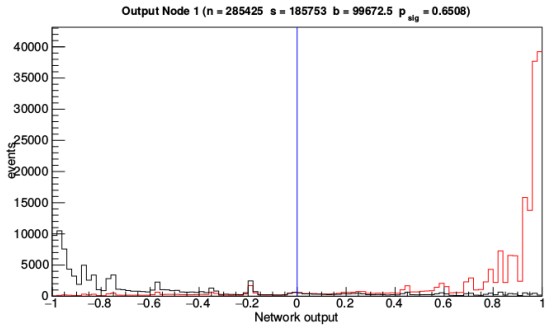

Figure 5.6: Distribution of the NEUROBAYES output variable. The histogram in red (black) corresponds to the distribution of signal (background) events.

The figure of merit to optimise in this case is

$$\mathrm{FoM}_{\mathrm{MVA}} = \frac{N_S}{\sqrt{N_S + N_B}}, \tag{5.5}$$

where N_S and N_B are estimated signal and background yields in the signal region, defined as

the range $[5240, 5320]\,\mathrm{MeV}/c^2$. This FoM approximates the signal significance in a counting experiment. The yield of combinatorial background, N_B, is obtained by performing a fit to the previously defined high-mass sideband region and extrapolating it to the signal region. The signal yield, N_S, is taken from the efficiency of the cut on the MVA output in signal MC, multiplied by a baseline yield taken from a fit to the $B^+ \to K^+K^+K^-$ data sample with no MVA cut applied. An extra factor, accounting for the ratios of branching fractions and selection efficiencies, is applied in the case of the other signal modes. The distribution of the figure of merit as a function of MVA cut value is shown in Figure 5.7. The cut value which maximises the $B^+ \to K^+\pi^+K^-$ figure of merit, $\mathrm{MVA} > 0.5$, is taken as the cut for all four modes.

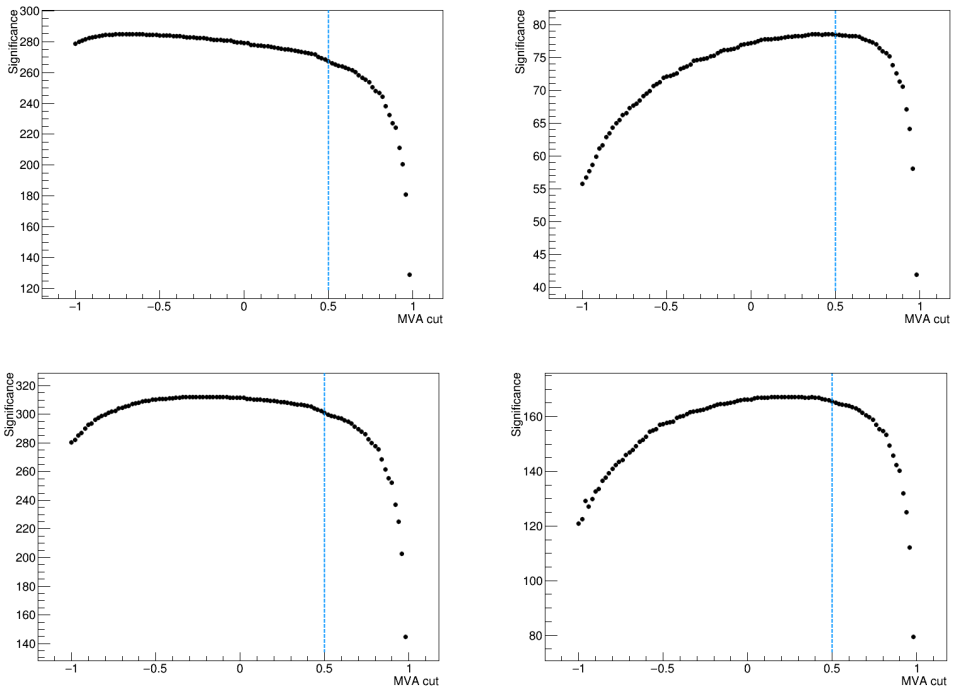

Figure 5.7: MVA figures of merit for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

5.5 Charm Vetoes

Candidates originating from $B^+ \to Dh^+$ decays can pass through the selection steps described in significant amounts. An explicit veto on the two-body invariant mass variables m_{13} and m_{23} around the D mass is used to remove $B^+ \to Dh^+$ candidates with the correct PID hypothesis. The width of the veto region is determined by fitting samples of $B^+ \to Dh^+$ MC that have the

above selection requirements applied. The veto region is defined as $\pm 30\,\mathrm{MeV}/c^2$, roughly $\pm 5\sigma$, around the fitted D mass, which corresponds to the range $[1830, 1890]\,\mathrm{MeV}/c^2$.

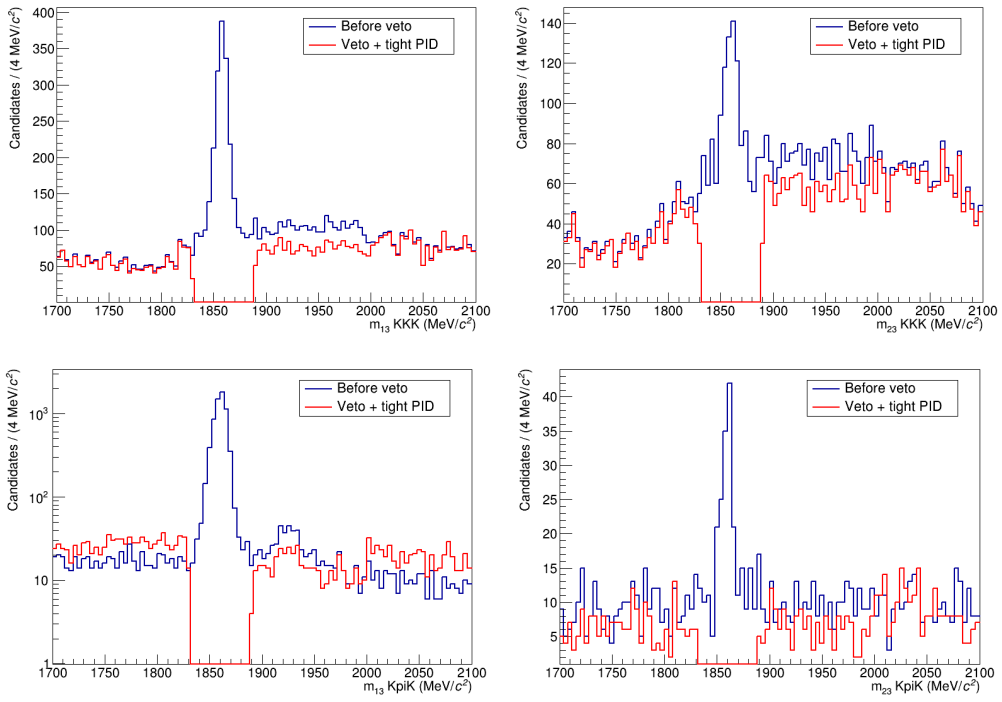

Figure 5.8: Effect of the mass vetoes and the tighter PID cuts in the region near the D mass in the m_{13} (left) and m_{23} (right) mass spectra. The top row corresponds to the KKK final state and the bottom row to the $K\pi K$ final state. Only candidates in the signal region defined for the three-body invariant mass $[5240, 5320]\,\mathrm{MeV}/c^2$ are shown.

Charmed candidates with one of the D decay products misidentified are mostly shifted outside the defined veto region. Depending on the specific decay and the misidentified particle, the shift can be towards higher or lower mass. Candidates with misID tracks populate a larger mass region than correctly-identified ones, due to the smearing caused by the misID. Vetoing a large region of phase space is not an optimal choice as it risks the danger of sculpting the shape of the combinatorial background and would remove some amount of the charmless hhh signal.

An alternative approach is taken, in which the PID requirements are tightened in the regions where misID charmed background is expected. In the $\pi \to K$ misID case, a tighter PID cut is applied to the two-body mass region $[1850, 2000]\,\mathrm{MeV}/c^2$. In the $K \to \pi$ case, the tighter cuts are applied in the region $[1700, 1850]\,\mathrm{MeV}/c^2$. The values of the tighter PID requirements are listed in Table 5.5, together with the 'global' cuts used outside these specific regions. The effect of these cuts on the two-body mass spectra are shown in Figures 5.8 and 5.9.

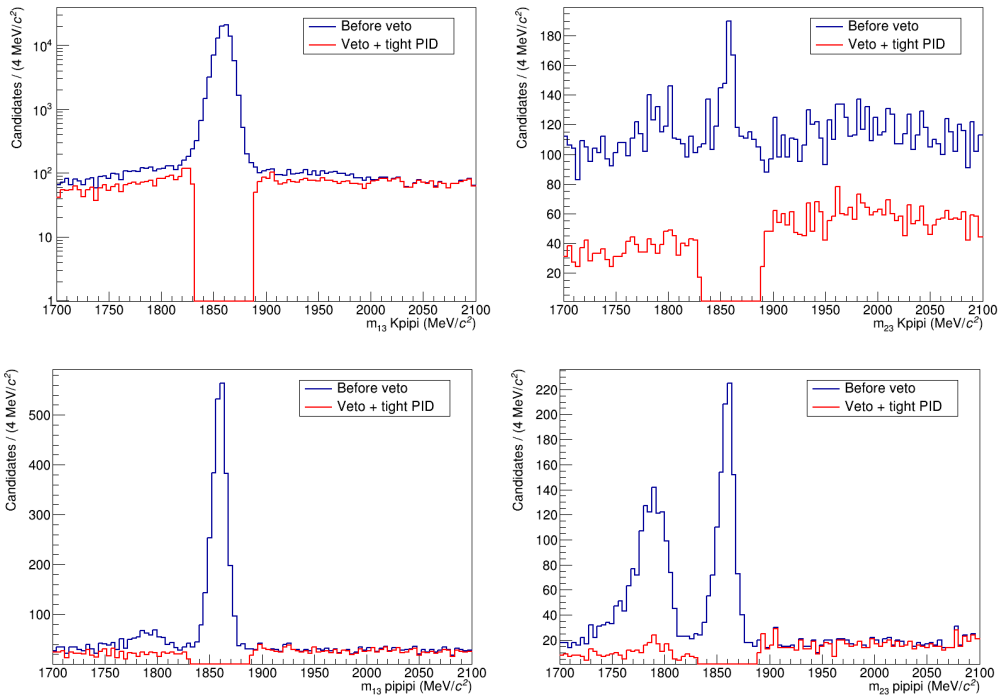

Figure 5.9: Effect of the mass vetoes and the tighter PID cuts in the region near the D mass in the m_{13} (left) and m_{23} (right) mass spectra. The top row corresponds to the $K\pi\pi$ final state and the bottom row to the $\pi\pi\pi$ final state. Only candidates in the signal region defined for the three-body invariant mass $[5240, 5320]\,\mathrm{MeV}/c^2$ are shown.

Table 5.5: Tight PID cuts applied to misID charmed candidates

Signal mode	Global cut (r_π, r_K)	Tight cut (r_π, r_K)
$B^+ \to K^+K^+K^-$	(–, 0.5)	(–, 0.9)
$B^+ \to K^+\pi^+K^-$	(0.2, 0.9)	(0.1, 1.0)
$B^+ \to K^+\pi^+\pi^-$	(0.5, 0.7)	(0.1, 1.0)
$B^+ \to \pi^+\pi^+\pi^-$	(0.5, –)	(0.1, –)

5.6 Partially Combinatorial Vetoes

Partially combinatorial candidates create a small excess in the high-mass sideband distribution. The effect becomes more pronounced once the MVA selection is applied, since the neural network training is less effective against this type of background. In fact, during the optimisation of the

MVA figure of merit, the excess of partially combinatorial candidates was seen to affect the extrapolation of the background yields, as illustrated in Figure 5.10.

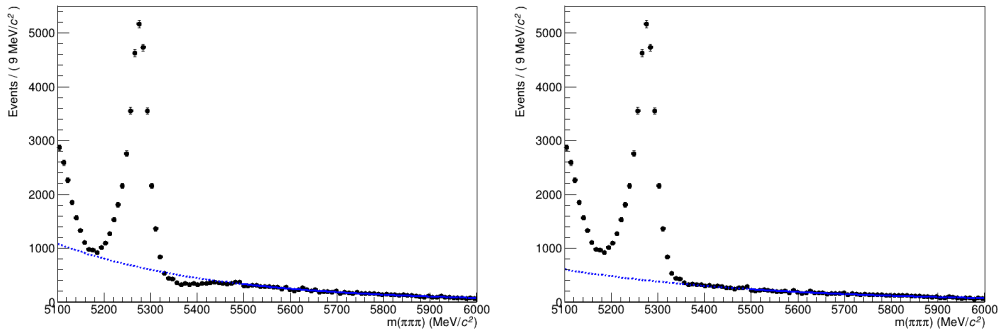

Figure 5.10: Exponential fit to the high-mass sideband in the $B^+ \to \pi^+\pi^+\pi^-$ mass spectrum, before (left) and after (right) partially combinatorial candidates are removed. The small excess in the region around $5500\,\mathrm{MeV}/c^2$ causes an overestimate of the combinatorial yield in the lower-mass regions.

The most efficient strategy for removing this contribution is a mass veto on the two-body invariant mass spectrum, around the mass of the parent particle. The distributions of m_{13} and m_{23} in the region around the B^0/B_s^0 mass are shown in Figure 5.11. The regions directly under the observed peaks are vetoed.

5.7 Multiple Candidates

A small number of events can contain more than one signal candidate. In these cases, only the candidate with the highest value of the MVA output is selected. Events that contain two or more candidates in different final states are not removed. The fraction of such events is given in Table 5.6.

Table 5.6: Fraction of events containing multiple candidates in each final state.

Channel	Fraction of events with 2 candidates (%)	Fraction of events with >2 candidates (%)
$B^+ \to K^+K^+K^-$	0.20	0.00
$B^+ \to K^+\pi^+K^-$	0.75	0.04
$B^+ \to K^+\pi^+\pi^-$	0.84	0.03
$B^+ \to \pi^+\pi^+\pi^-$	2.40	0.15

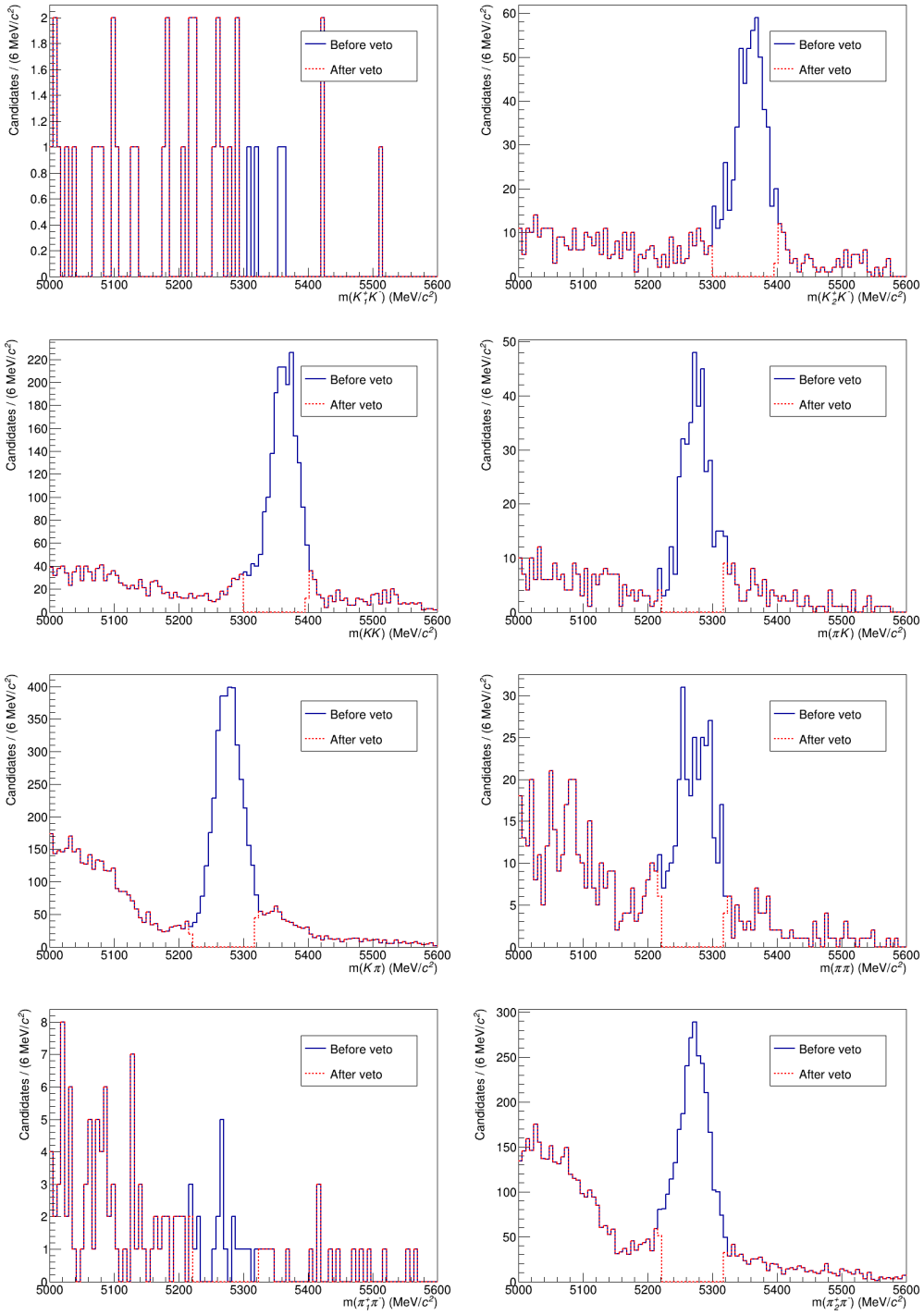

Figure 5.11: Vetoes on partially combinatorial contributions to the m_{13} (left) and m_{23} (right) two-body invariant mass distributions. From top to bottom: $B^+ \to K^+K^+K^-$, $B^+ \to K^+\pi^+K^-$, $B^+ \to K^+\pi^+\pi^-$, $B^+ \to \pi^+\pi^+\pi^-$.

CHAPTER 6

Efficiencies

> *Dividing one number by another is mere computation; knowing what to divide by what is mathematics.*
>
> Jordan Ellenberg, *How Not to Be Wrong*

Understanding the efficiency with which signal candidates are selected is crucial to the determination of the branching fraction ratios. Many parts of the selection strategy have been designed so that the same requirements are applied to all four signal modes, since this ensures that a number of potential systematic effects cancel when evaluating the branching fraction ratios. The PID efficiency is the main exception to this, but some second-order effects originating from other selection requirements are also expected to not cancel fully in the ratio.

The PID efficiency also differs in the manner it is determined; as stated in Section 5.3, a data-driven method is used to calculate the efficiency of the PID selection. In contrast, the other components of the total selection efficiency are evaluated with the MC samples described in Section 4.2. The total efficiency $\epsilon_{\rm tot}$ can therefore be factorised into $\epsilon_{\rm PID}$ and $\epsilon_{\rm sel}$ components, which can be evaluated independently. While the MC samples are expected to model $\epsilon_{\rm sel}$ with good accuracy, a number of data-driven corrections are applied to bring certain areas of the simulation closer to agreement with data; these corrections are discussed in Sections 6.3 to 6.6.

The distribution of the total efficiency is not expected to be uniform across the phase space. In fact, different steps of the selection can separately favour some regions of the Dalitz plot over others. To take this into account, the efficiencies that enter Equation 4.1 are evaluated as

a function of the Dalitz plot. Phase-space-integrated efficiencies are also evaluated but are only used for constraining the yields of cross-feed backgrounds in the maximum likelihood fit to data.

As discussed in Section 2.3, B decays mostly populate the edges of the Dalitz plot in its conventional representation, making the 'square DP' (sqDP) representation more useful in studying the variation of the efficiency across the phase space. Decays with identical particles, such as $B^+ \to K^+K^+K^-$ and $B^+ \to \pi^+\pi^+\pi^-$, are symmetric with respect to the exchange of the two like-sign particles, with this symmetry manifesting itself as a reflection axis in the Dalitz plot. The DP can therefore be 'folded' along this reflection axis without any loss of information. As a consequence of the definition of (m', θ') chosen in this analysis, the reflection axis in the sqDP is the line $\theta' = 0.5$.

6.1 Selection Efficiency

The efficiencies of the trigger, stripping, and offline selection steps detailed in the previous chapter all contribute to ϵ_{sel}. An additional contribution originates from the acceptance efficiency of the detector. Clearly, the trajectory of all three final-state particles must remain within the LHCb angular acceptance for the signal candidate to be detected. In fact, during MC generation, decays with particles outside the detector acceptance do not go through the full detector simulation, in order to be able to produce large samples more efficiently. For each signal channel, large (10^8 entries) toy MC samples were generated with the same model used for generating the fully-reconstructed MC, but without any acceptance requirements imposed. These toy samples are scaled to the total number of generated events for each fully-reconstructed signal MC sample and are used to obtain a description of the phase space prior to any selection steps.

The variation of the selection efficiency as a function of the Dalitz plot is shown in Figure 6.1. These efficiency maps are the weighted average of the four independent MC samples $((2011, 2012), (\text{MagDown}, \text{MagUp}))$ of each signal decay, taking into account the differences in luminosity and B^+ production cross-section between them. The sharp features in which the efficiency drops significantly are due to the vetoes on the charmed backgrounds. The phase-space-integrated efficiency values are given in Table 6.1.

The statistical uncertainty in ϵ_{sel} due to finite MC samples is taken as a systematic uncertainty on the final results. Since a weighted average is used to obtain the efficiency of the combined Run 1 dataset, some care must be taken when determining this uncertainty. A Bayesian approach is taken, as detailed in Ref. [106] and summarised in Chapter 8.

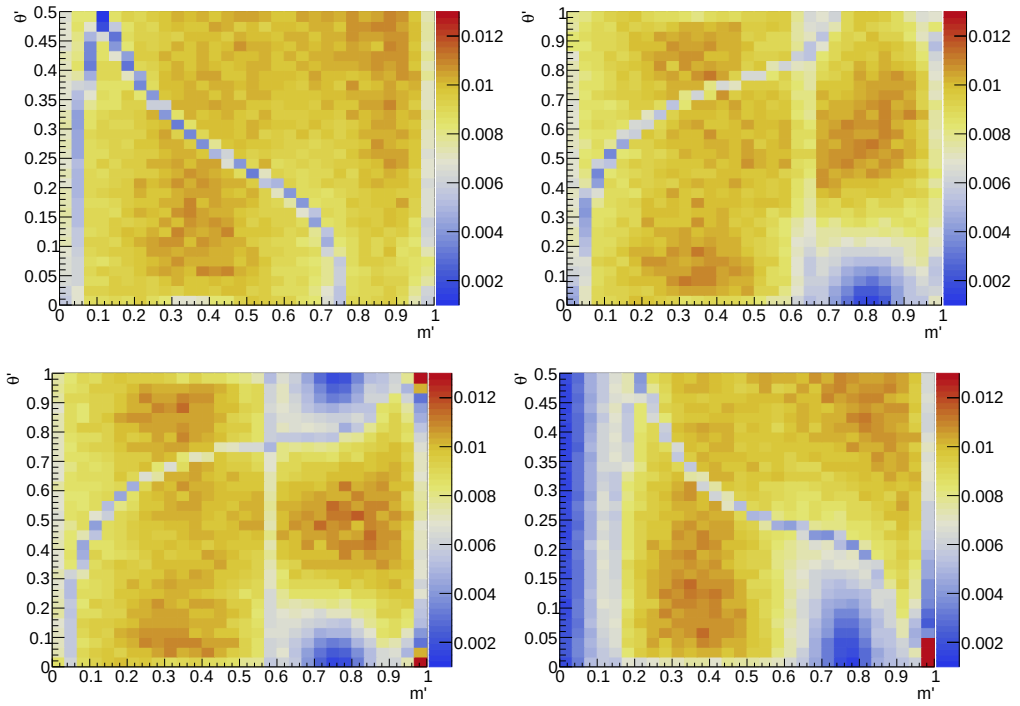

Figure 6.1: Distribution of the selection efficiency as a function of Dalitz plot position, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

Table 6.1: Phase-space-integrated selection efficiencies and their statistical uncertainties.

Signal channel	ϵ_{sel} (%)				
	2011		2012		Weighted average
	MagDown	MagUp	MagDown	MagUp	
$B^+ \to K^+K^+K^-$	1.0377 ± 0.0015	1.0296 ± 0.0015	0.8859 ± 0.0013	0.8882 ± 0.0013	0.9315 ± 0.0006
$B^+ \to K^+\pi^+K^-$	1.0006 ± 0.0014	0.9929 ± 0.0014	0.8566 ± 0.0013	0.8588 ± 0.0013	0.8998 ± 0.0006
$B^+ \to K^+\pi^+\pi^-$	0.9822 ± 0.0015	0.9788 ± 0.0014	0.8447 ± 0.0013	0.8433 ± 0.0013	0.8853 ± 0.0006
$B^+ \to \pi^+\pi^+\pi^-$	0.9141 ± 0.0013	0.9092 ± 0.0013	0.7878 ± 0.0012	0.7879 ± 0.0012	0.8253 ± 0.0006

6.2 PID Efficiency

The efficiency of the PID selection is determined from the calibration samples provided by the `PIDCalib` package [104]. It incorporates the efficiency of the cuts applied across the whole phase space and the tighter cuts applied to remove misidentified charmed backgrounds. Since the latter cuts are only applied to a very small percentage of signal candidates, it can be informative to compare the efficiency of the full PID selection with that of the ‘global‘ cuts only. For

reference, the distribution of the ‘global‘ PID efficiency across the DP is shown in Figure 6.2. The corresponding phase-space-integrated values are given in Table 6.2.

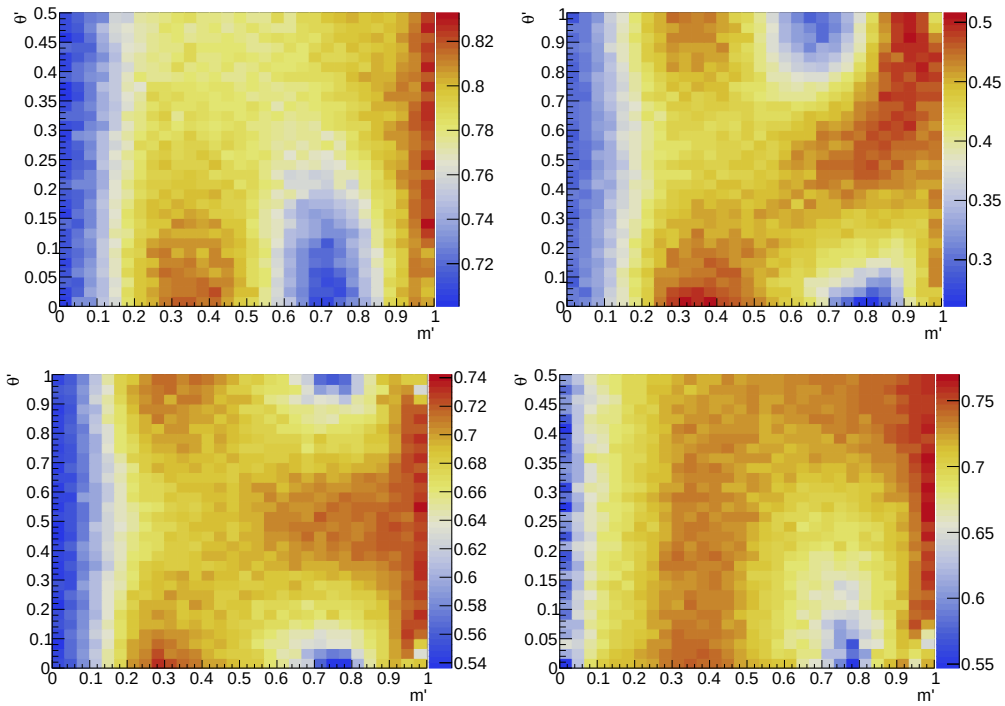

Figure 6.2: The ‘global’ PID efficiency across the Dalitz plot, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right). The ‘global’ efficiency maps are given for illustration purposes and are not used in determining the total efficiency.

Table 6.2: Phase-space integrated values of the ‘global’ PID efficiency. These values are given only for comparison with the efficiencies incorporating the ‘tight’ PID cuts.

Signal channel	‘global’ ϵ^{PID} (%)				
	2011		2012		Weighted average
	MagDown	MagUp	MagDown	MagUp	
$B^+ \to K^+K^+K^-$	76.97	76.70	78.70	78.74	77.87
$B^+ \to K^+\pi^+K^-$	43.08	42.63	42.55	41.79	42.41
$B^+ \to K^+\pi^+\pi^-$	67.74	67.27	67.90	66.87	67.43
$B^+ \to \pi^+\pi^+\pi^-$	72.26	71.79	71.67	70.30	71.34

The tighter PID cuts applied to the regions where misidentified charm backgrounds are present have a similar effect on the phase-space distribution of ϵ_{PID} as the mass vetoes have on ϵ_{sel}. The resulting sharp drops in the PID efficiency can be seen in Figure 6.3. The phase-space-integrated values given in Table 6.3 roughly show a 1% drop in efficiency due to the tighter requirements.

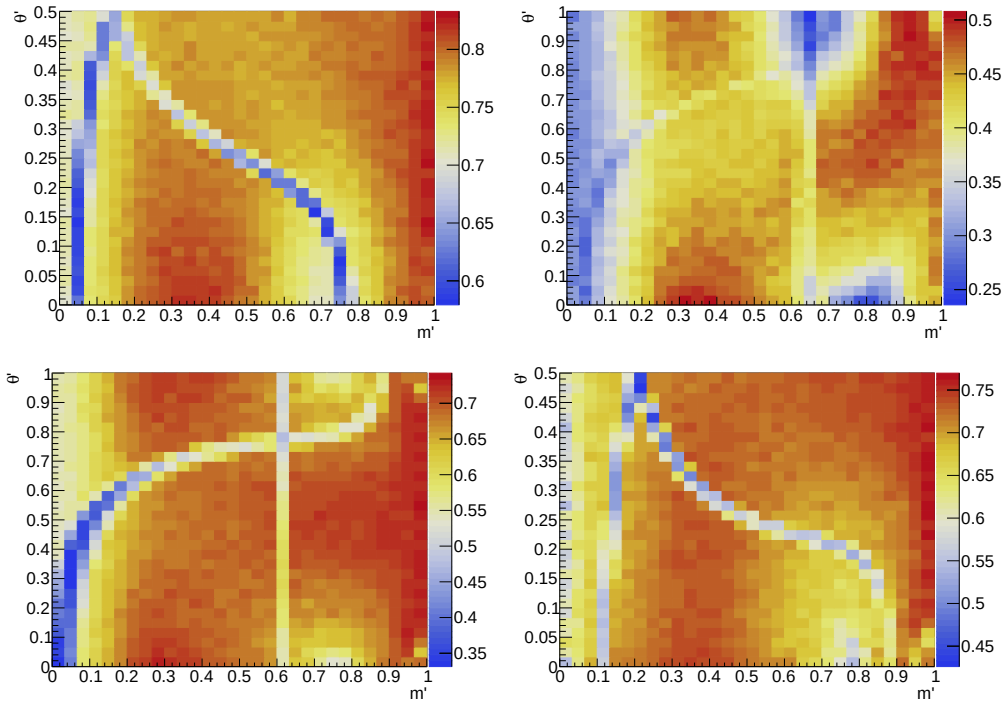

Figure 6.3: PID efficiency distribution taking into account the effect of the 'tight' PID cuts, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

A number of uncertainties are associated with the `PIDCalib` procedure. The statistical uncertainty on the PID efficiency has components from the finite size of both the signal MC and the calibration samples used. A systematic uncertainty originates from the assumption that the binning used to parameterise the efficiency ensures that the efficiency variation is smooth across the phase-space and that, within each bin, the efficiency does not vary greatly. Another systematic uncertainty is assigned due to the *sWeights* used for background subtraction of the calibration samples. These uncertainties on ϵ_{PID} are taken as systematic uncertainties on the branching fraction measurement.

Table 6.3: Phase-space integrated PID efficiency, taking into account the 'tight' PID cuts. The quoted uncertainty is the sum in quadrature of the uncertainty from the size of `PIDCalib` sample and the reference sample.

Signal channel	ϵ^{PID} (%)				
	2011		2012		Weighted average
	MagDown	MagUp	MagDown	MagUp	
$B^+ \to K^+K^+K^-$	76.18 ± 0.05	75.92 ± 0.05	77.89 ± 0.03	77.27 ± 0.03	77.07 ± 0.02
$B^+ \to K^+\pi^+K^-$	42.75 ± 0.04	42.30 ± 0.04	42.20 ± 0.04	41.44 ± 0.04	42.06 ± 0.02
$B^+ \to K^+\pi^+\pi^-$	66.24 ± 0.04	65.77 ± 0.04	66.32 ± 0.03	65.30 ± 0.04	65.88 ± 0.02
$B^+ \to \pi^+\pi^+\pi^-$	71.17 ± 0.06	70.70 ± 0.06	70.56 ± 0.04	70.56 ± 0.05	70.24 ± 0.03

6.3 B Kinematics Correction

The kinematic distribution of the parent B meson is generally well reproduced in the LHCb-specific PYTHIA tune used during MC production. Nonetheless, the large, very clean data sample of $B^+ \to K^+K^+K^-$ candidates available can be used to test this assumption.

The distributions of the kinematic variables of the B candidate in signal MC and background-subtracted data are shown in the left column of Figure 6.4. The background subtraction is performed with *sWeights* obtained from the nominal data fit described in Chapter 7. All three variables are well described, though a small amount of mismodelling can be seen in the distributions of the B pseudorapidity and transverse momentum.

A correction is performed by considering the distribution of background-subtracted data and MC in bins of (p_{T}, η) and taking the ratio of the two distributions. The binning chosen to describe the (p_{T}, η) parameter space is shown in Figure 6.5. The ratio in each bin is assigned as the weight for each MC candidate falling in that bin, which is then used to weigh the distribution of the total efficiency. A χ^2 test is performed to check the efficacy of the correction; the reweighted MC distributions shown in the right column of Figure 6.4 follow the *sWeighted* data distributions more closely.

The weights obtained with $B^+ \to K^+K^+K^-$ data and MC are used to reweigh all four signal channels, since the kinematics of the parent B should be independent of which final state it subsequently decays into. The procedure is later performed using the $B^+ \to K^+\pi^+\pi^-$ samples as a test of this assumption, with the difference being taken as a systematic uncertainty.

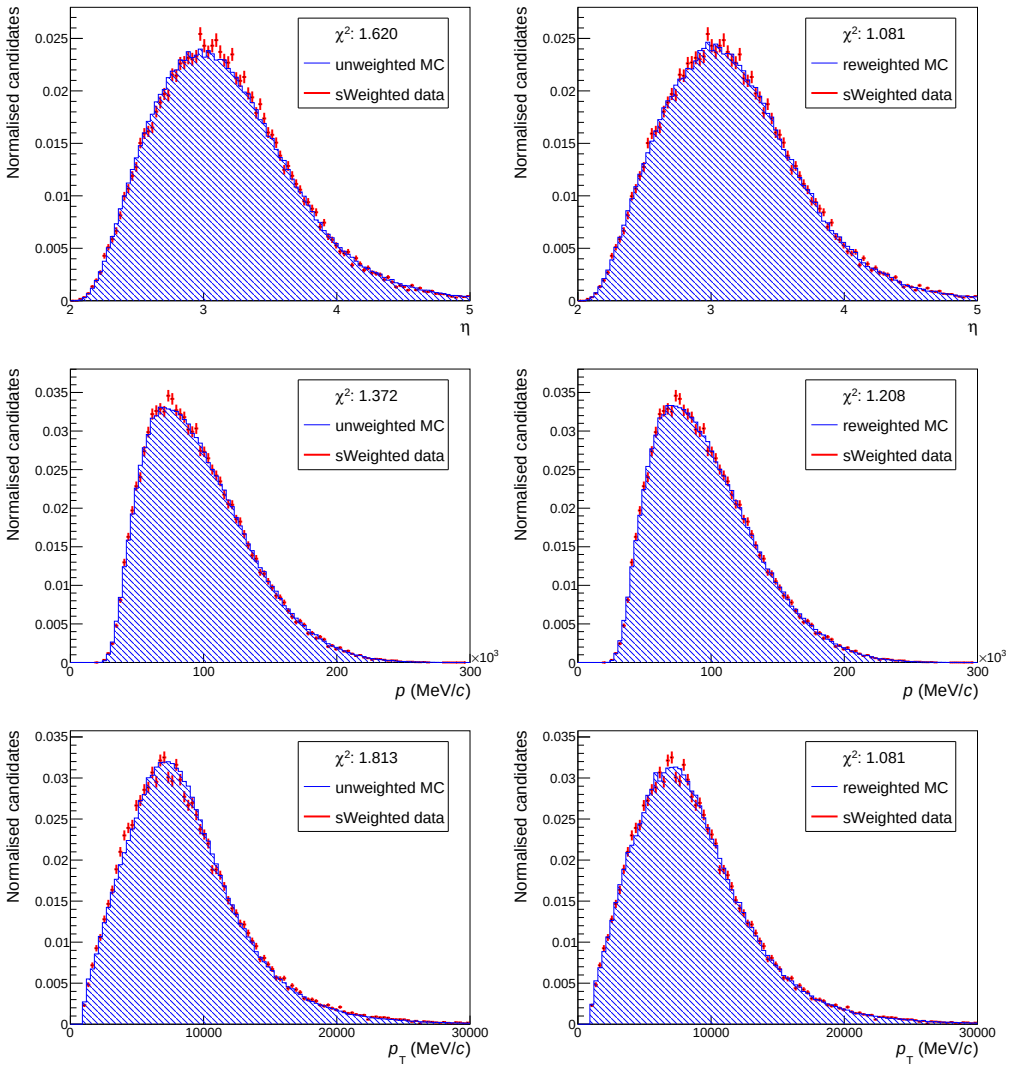

Figure 6.4: Comparison between *sWeighted* data and signal MC before (left) and after (right) the kinematic reweighting for the KKK mode.

6.4 Tracking Correction

Standard correction tables for the track reconstruction efficiency are available within the LHCb collaboration, determined using the 'tag-and-probe' method described in Section 3.3.6 [107]. The corrections are evaluated as a function of track momentum and pseudorapidity and are given for each data-taking period. The tables corresponding to the years 2011 and 2012 are shown in Figure 6.6. The weight assigned to each MC candidate is given by the product of the weights of

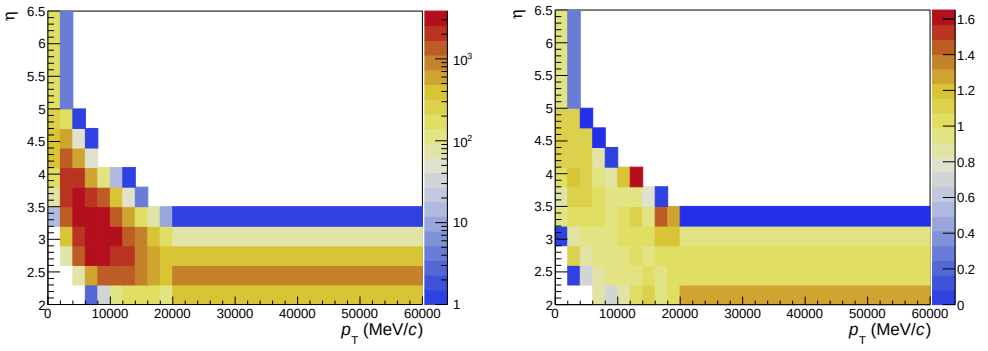

Figure 6.5: Distribution of *sWeighted* KKK data candidates (left) and the corresponding weights obtained (right). The z axis has been truncated in the weights histogram, since the value in the highest bin is ~ 16. However, the background-subtracted yield in this bin is ~ 1, making its contribution negligible.

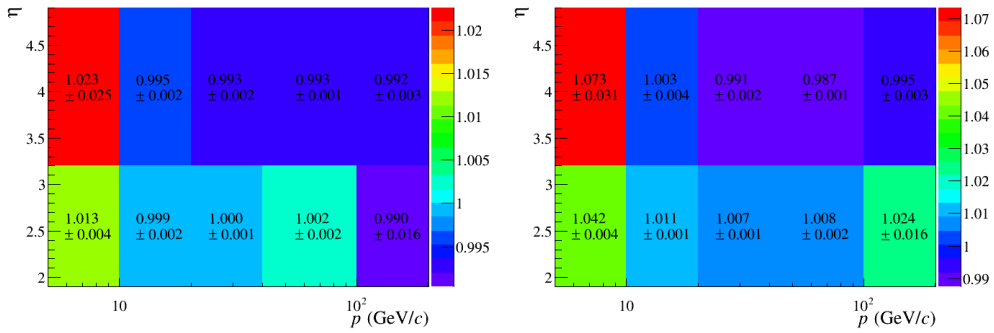

Figure 6.6: Tracking correction maps for 2011 (left) and 2012 (right).

its three associated tracks.

6.5 L0 TIS Correction

In an event categorised as TIS, the particle that has fired the trigger will most likely originate from the other b-hadron in a $pp \to b\bar{b}X$ process. Whereas the signal B candidate is forced to decay to the particular final state being considered, the other b quark will hadronise/decay according to the fragmentation/branching fractions. This can potentially introduce a mismatch with data with respect to the relative number of cases in which the TIS decision arises from, *e.g.* the muon or hadron L0 trigger. Since the kinematics of the $b\bar{b}$ pair are correlated, the frequency with which different TIS triggers fire could have some dependence on the signal channel considered.

The correction strategy taken is to reweigh the MC samples so that the proportion of

candidates in a given TIS category matches that seen in background-subtracted data. A small issue arises from the fact that in any given event, multiple TIS triggers can potentially fire (Hadron, Muon, DiMuon, Electron, Photon). A number of exclusive TIS categories are constructed out of combinations of the existing trigger decisions, to allow for a single weight to be assigned to any particular candidate. The weights obtained for this correction are shown in Figure 6.7, with the bin labels corresponding the concatenations of the trigger decisions that comprised each exclusive category. Candidates for which no TIS trigger fires are given a weight of one.

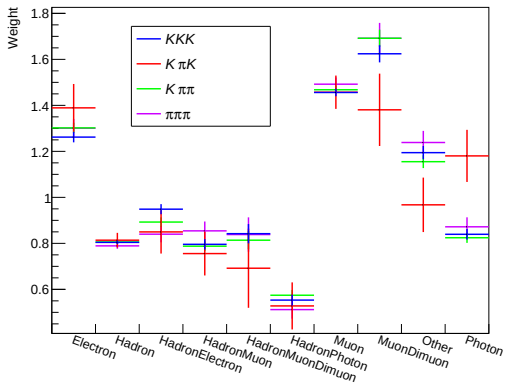

Figure 6.7: L0 TIS corrections as a function of the constructed categories. The 'other' category includes all other possible combinations of L0 TIS decisions.

6.6 L0Hadron TOS Correction

The L0Hadron response is not well modelled in simulation and so a set of standard calibration tables exists within the collaboration. These tables give the absolute L0Hadron trigger efficiency as a function of the transverse energy deposited, determined from samples of kaons and pions from $D^{*-} \to \overline{D}^0 \left(\to K^+\pi^-\right)\pi^-$ decays. The trigger efficiency is obtained through the 'TISTOS' method [108]:

$$\epsilon_{\text{had}} = \frac{N_{\text{TIS+TOS}}}{N_{\text{TIS}}} \tag{6.1}$$

where $N_{\text{TIS+TOS}}$ is the number of candidates associated with both L0Hadron TOS and any L0 TIS trigger, and N_{TIS} is the total number of candidates associated with any L0 TIS decision.

Multi-body decays must also consider an additional correction due to overlapping energy deposits in the calorimeter. Overlaps can occur when two closely-flying tracks deposit energy in adjacent calorimeter cells; the combined cluster may be enough to fire the trigger even when the individual particles would not.

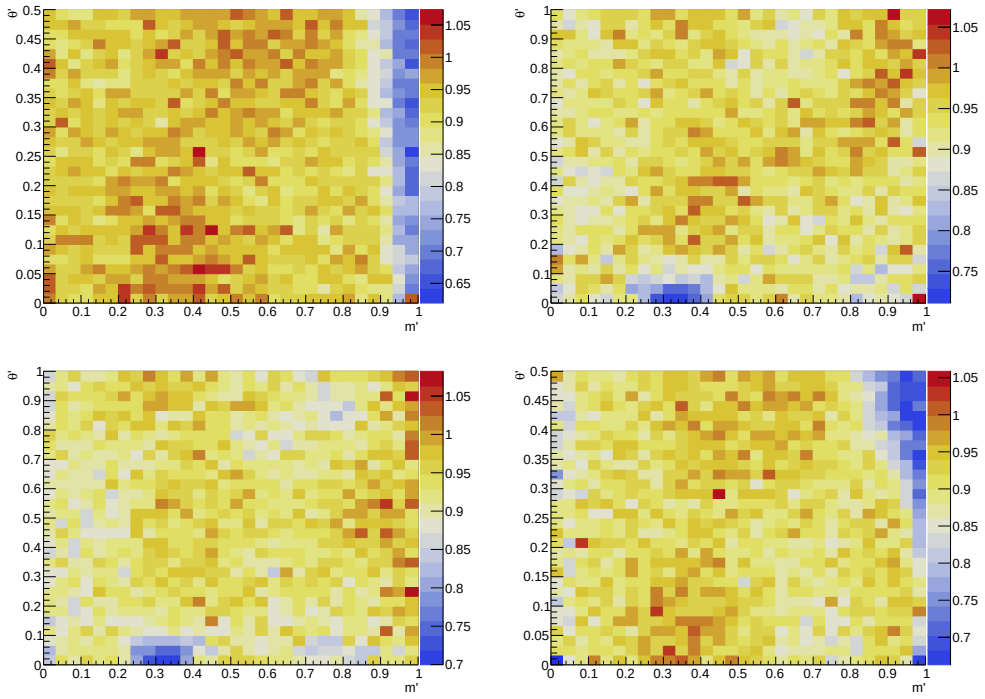

Figure 6.8: Run-1-averaged L0Hadron correction maps for candidates firing L0Hadron TOS, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

The correction is determined by taking the ratio of the absolute data efficiencies from the tables and the MC efficiencies, with the L0 TIS correction applied; the absolute MC efficiencies are also determined with the TISTOS method. Correction maps are also obtained for MC candidates not firing L0Hadron TOS, by constructing data tables with the complement $(1 - \epsilon)$ of the efficiency. The corrections are evaluated separately for each year and magnet polarisation. The correction maps, averaged over the four year+polarisation samples, are given in Figure 6.8 for candidates firing L0Hadron TOS, and in Figure 6.9 for candidates not firing it.

6.7 Total Efficiency

The phase-space distributions of the total efficiency, including the corrections described above, are shown in Figure 6.10. For reference, the phase-space-integrated values are shown in Table 6.4, before and after corrections. The uncertainties due to the correction procedures applied are taken as systematic uncertainties, given in Chapter 8.

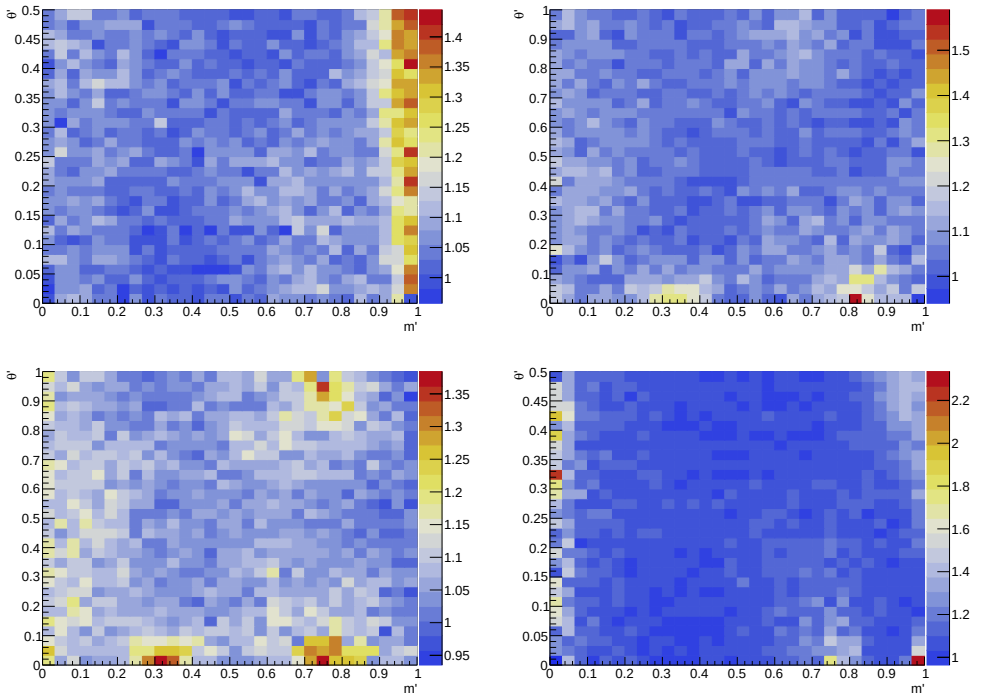

Figure 6.9: Run-1-averaged L0Hadron correction maps for candidates that do not fire L0Hadron TOS, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

Table 6.4: Phase-space integrated total efficiencies and their statistical uncertainties.

Signal channels	ϵ^{tot} (%)	$\epsilon^{\text{tot}}_{\text{corr}}$ (%)
$B^+ \to K^+K^+K^-$	0.7179 ± 0.0005	0.7053
$B^+ \to K^+\pi^+K^-$	0.3785 ± 0.0003	0.3744
$B^+ \to K^+\pi^+\pi^-$	0.5833 ± 0.0004	0.5763
$B^+ \to \pi^+\pi^+\pi^-$	0.5797 ± 0.0004	0.5674

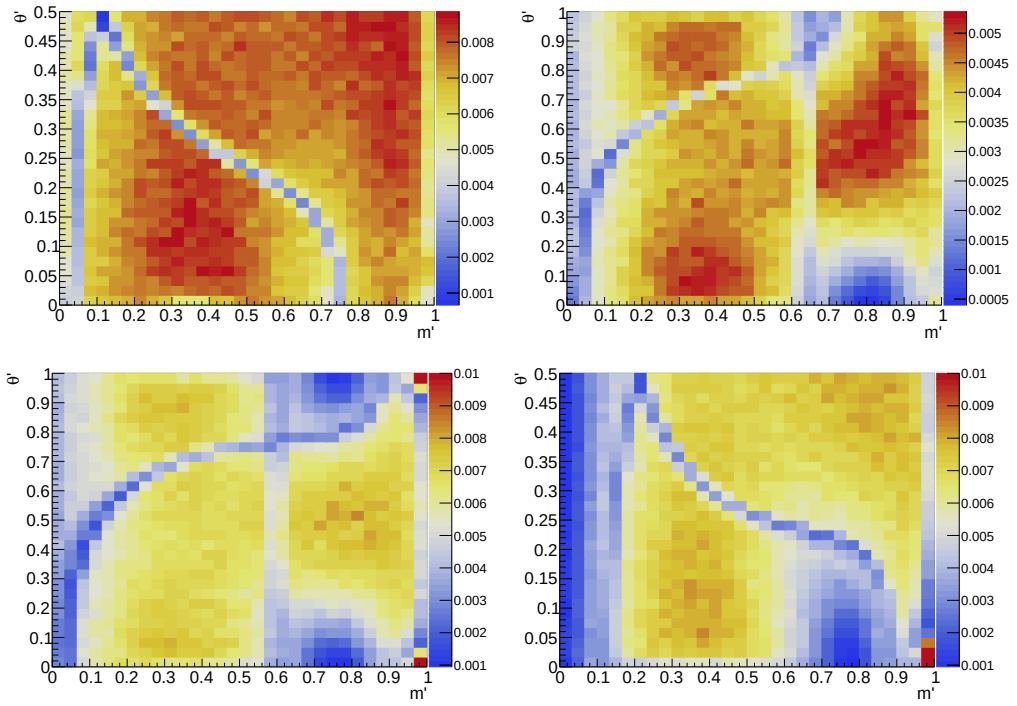

Figure 6.10: Total efficiency distribution as a function of Dalitz plot position, with all efficiency corrections applied, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

CHAPTER 7

Invariant Mass Fit

> *All models are wrong, but some are useful.*
>
> George Box, *Robustness in the Strategy of Scientific Model Building*

An accurate determination of the signal yields present in the data requires the construction of a fit model that comprehensively describes the expected background contributions to each final state. The various components in each final state model, the parameterisations chosen to describe them, and any constraints applied to the full fit model are detailed in Section 7.1. The results of applying the fit model to the dataset, and the method used to extract the signal yields and use them to determine the ratios of branching fractions are presented in Section 7.2.

7.1 Fit Model

7.1.1 PDF Parameterisations

Peaking components, such as the signal and cross-feed components, are modelled in the nominal fit as the sum of two Crystal Ball (CB) functions. The Crystal Ball function is defined as a core Gaussian distribution, with one of its tails replaced by a power-law distribution [109]

$$\mathrm{CB}(m;\mu,\sigma,\alpha,n) = \mathcal{N}\cdot\begin{cases} \exp(-(m-\mu)^2/2\sigma^2) & \text{if } (m-\mu)/\sigma > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \left(\frac{n-\alpha^2}{|\alpha|} - \frac{m-\mu}{\sigma}\right)^{-n} \exp(-\alpha^2/2) & \text{if } (m-\mu)/\sigma \leq -\alpha, \end{cases} \tag{7.1}$$

where m is the reconstructed mass variable, μ and σ are the standard Gaussian mean and standard deviation, α defines the number of standard deviations from the mean at which the transition from the core to the power-law tail occurs, n is the power-law exponent, and $\mathcal{N}$ is a normalisation factor. This definition of the Crystal Ball determines the low-mass, left-hand tail as the power law component; it is possible to define it such that the right-hand tail is the power law component instead.

The addition of the two CB functions is performed with the requirement that the mean and standard deviation of the core Gaussian distribution is the same for both CB functions. Additionally, they are added with one tail to each side such that the resulting function encapsulates the different effects that modify the tails on both sides. The right-hand power-law tail describes the effect of tracking imperfections and other detector effects that are non-Gaussian in nature, while the left-hand tail not only encapsulates these effects but also the effect of missing final-state radiation on the mass distribution. Finally, an extra parameter $f = \mathcal{N}_1/(\mathcal{N}_1+\mathcal{N}_2)$ is included to allow for the two CB functions to contribute in different amounts relative to each other. A sum of two Crystal Ball functions with these additional requirements is often referred as a 'double Crystal Ball' (dCB) function.

An alternative parameterisation, the Hypatia distribution [110], is used to evaluate a systematic uncertainty associated with the choice of the dCB to describe peaking components. Much like the CB function, it is also defined by separate core and tail functions; however, the Hypatia core is the generalised hyperbolic distribution.

Partially-reconstructed backgrounds exhibit significantly different behaviour to the components originating from true $B^+ \to h^+h'^+h''^-$ candidates. Such backgrounds are better modelled as ARGUS functions [43], convoluted with a Gaussian distribution to account for the resolution of the detector. The generalised ARGUS function is defined by three parameters, $M_{\rm thr}$, c, and p,

$$A(m; M_{\rm thr}, c, p) = \frac{2^{-p}c^{2(p+1)}}{\Gamma(p+1) - \Gamma(p+1, c^2/2)} \times \frac{m}{M_{\rm thr}^2}\left(1 - \frac{m^2}{M_{\rm thr}^2}\right)^p \exp\left\{-\frac{1}{2}c^2\left(1 - \frac{m^2}{M_{\rm thr}^2}\right)\right\} \tag{7.2}$$

when $m < M_{\rm thr}$, and zero elsewhere. The expressions $\Gamma(x)$ and $\Gamma(s, x)$ represent the gamma and upper incomplete gamma functions, respectively, $M_{\rm thr}$ is the threshold mass value, c governs the curvature of the function, and p controls the falling of the slope. The 'standard' ARGUS distribution is obtained by fixing $p = 0.5$ in Equation 7.2.

A combinatorial background component is present in all four mass spectra and is described by an exponential distribution in the nominal data fit, and by Chebyshev polynomials up to second

order (*i.e.* the first three Chebyshev polynomials) when evaluating the associated model systematic uncertainty.

7.1.2 Corrections to $B^+ \to h^+h'^+h''^-$ MC distributions

The signal MC samples used in this analysis have been generated flat in the square Dalitz plot, *i.e.* the distribution of candidates is uniform across the (m', θ') parameter space. However, the true dynamics seen in $B^+ \to h^+h'^+h''^-$ decays deviate significantly from a uniform distribution. In practice this means that certain regions of the phase-space are overrepresented in MC. In MC samples reconstructed under a mass hypothesis in which one or more particles are misidentified, this effect is reflected in the invariant mass distribution, as shown in Figure 7.1. This occurs since the smearing of the invariant mass distribution is related to the momentum of the misidentified particle. The secondary peak seen in the low-mass region corresponds to $B^+ \to K^+\pi^+\pi^-$ candidates in which the kaon recoils off the pion pair and gains a large fraction of the momentum available. This scenario corresponds to one of the edges of the conventional DP, which is exactly one of the regions enlarged by the definition of the square DP.

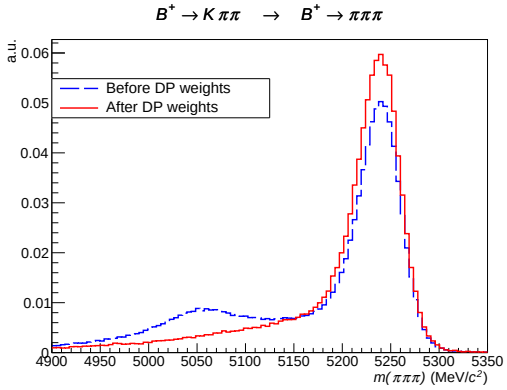

Figure 7.1: Effect of the DP-correcting weights on the invariant-mass distribution of $B^+ \to K^+\pi^+\pi^-$ MC candidates, reconstructed as $\pi\pi\pi$. The PID-correcting weights have been applied to both distributions shown.

This effect is counteracted by applying weights to the MC samples so that the generated DP distribution approximates the physical DP distribution. The result of this reweighting in the $B^+ \to K^+\pi^+\pi^-$ case is shown in Figure 7.2 as an illustration of the procedure.

It should be noted that the PID-correcting weights, evaluated from `PIDCalib`, are also applied to the MC before the mass fits, in order to model the effect of the PID cuts on the shape of the cross-feed distributions.

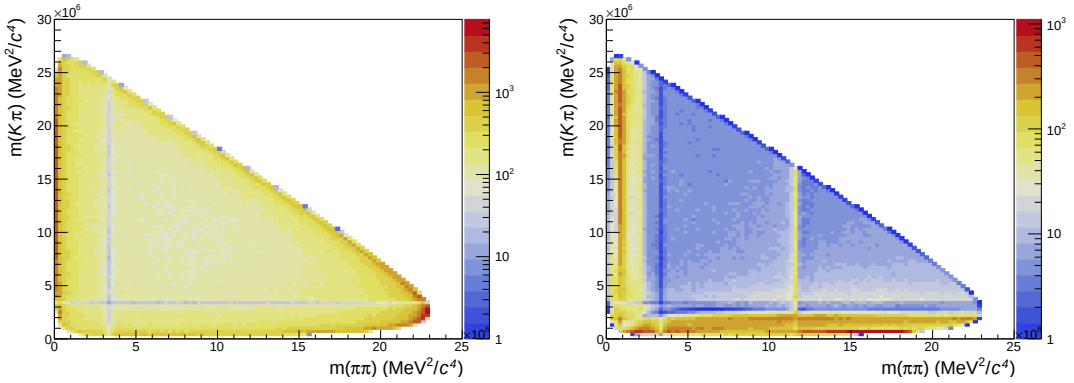

Figure 7.2: Dalitz plot distribution of $B^+ \to K^+\pi^+\pi^-$ MC candidates without (left) and with (right) applying the DP-correcting weights.

7.1.3 Individual Final-State Fit Models

Generally, the same types of background, discussed in Chapter 5, are expected in all four final states; however, their relative predominance can vary significantly between the different final states. This section summarises the individual contributions to the fit model of each final state.

All signal and cross-feed components have the tail parameters of their dCB shape fixed to the values extracted from fits to the corresponding MC sample. The parameters extracted from the signal distributions in MC are very similar for all four signal channels; the fits to these distributions are collected together in Figure 7.3 so they can be more easily compared, with the extracted PDF parameters being given in Table 7.1. Nonetheless, differences can be seen in the parameters extracted from the different modes, such as the width of the core Gaussian increasing as a function of the number of pions in the final state, which justifies the use of the individually-fitted parameters, as opposed to a common set of parameters for all final states.

Table 7.1: Fitted dCB parameters from the fits to signal MC. The parameters α_1, n_1 correspond to the left-hand tail of the function.

Signal Channel	$\mu(\,\mathrm{MeV}/c^2)$	$\sigma(\,\mathrm{MeV}/c^2)$	f	α_1	α_2	n_1	n_2
KKK	5280.8 ± 0.02	14.95 ± 0.01	0.36 ± 0.01	1.51 ± 0.03	-2.06 ± 0.02	1.81 ± 0.03	3.06 ± 0.05
$K\pi K$	5280.6 ± 0.09	15.30 ± 0.08	0.43 ± 0.08	1.50 ± 0.10	-1.95 ± 0.08	1.80 ± 0.10	3.20 ± 0.30
$K\pi\pi$	5280.4 ± 0.05	16.17 ± 0.04	0.48 ± 0.04	1.48 ± 0.05	-1.96 ± 0.04	1.61 ± 0.05	2.80 ± 0.10
$\pi\pi\pi$	5280.3 ± 0.05	16.89 ± 0.05	0.52 ± 0.04	1.44 ± 0.05	-1.86 ± 0.05	1.66 ± 0.05	3.00 ± 0.10

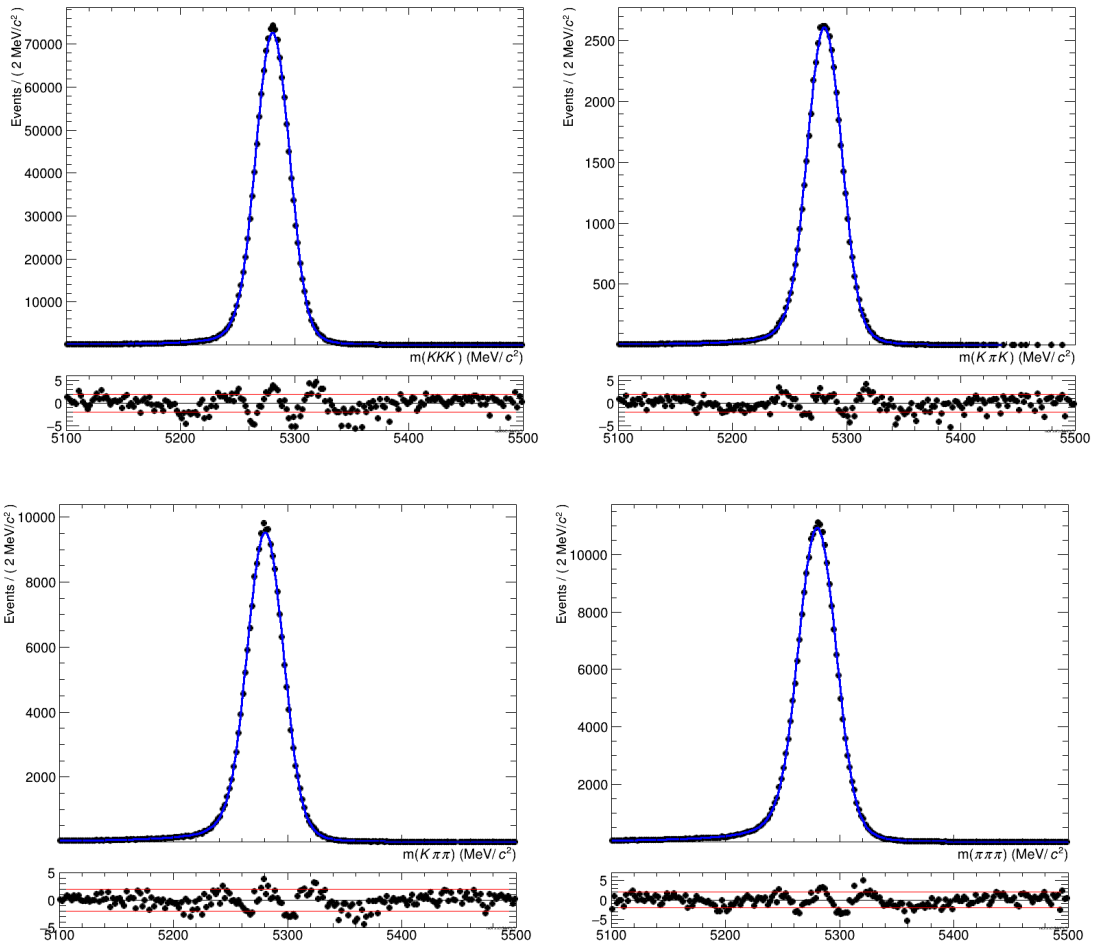

Figure 7.3: Signal MC distributions, with the fitted dCB function overlaid, for $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

7.1.3.1 $B^+ \to K^+K^+K^-$

In addition to signal and combinatorial background components, the KKK final state receives contributions from cross-feed from $B^+ \to K^+\pi^+K^-$ and $B^+ \to K^+\pi^+\pi^-$, and a partially-reconstructed contribution from four-body (B^+/B^0) decays. Fits in which an additional component, corresponding to partially-reconstructed four-body B_s^0 decays, is included were tested but the yield of the additional component was consistently found to be zero. The standard ARGUS function is used to describe the (B^+/B^0) component, with its shape parameter left floating in the data fit.

7.1.3.2 $B^+ \to K^+\pi^+K^-$

The $K\pi K$ final state receives significant cross-feed contributions from $B^+ \to K^+K^+K^-$ and $B^+ \to K^+\pi^+\pi^-$; the expected cross-feed yields of $B^+ \to \pi^+\pi^+\pi^-$ and the 'self-misID' $B^+ \to \pi^+K^+K^-$ are much smaller, but are still included in the fit model. A good description of the data is achieved only with the inclusion of three partially-reconstructed background contributions, corresponding to four-body (B^+/B^0), charmless B_s^0, and B_s^0 decays with an intermediate charm state. The charmed B_s^0 component is modelled using an MC sample of $B_s^0 \to D_s^-(\to K\pi K)\pi^+$ decays. The charmless B_s^0 component, $B_s^0 \to (K^+\pi^-)(K^-\pi^-)$, can proceed via a number of quasi-two-body intermediate states in a variety of spin combinations, *e.g.* scalar-scalar, scalar-vector, vector-vector. In the absence of an MC sample that correctly describes the ratios with which these quasi-two-body states contribute to the overall $(K^+\pi^-)(K^-\pi^-)$ final-state [111], the exact data sample used in that work is used to determine the expected shape in the $B^+ \to K^+\pi^+K^-$ mass spectrum. The shape of the (B^+/B^0) component was first left to float in the data fit, but its parameters were found to have very high correlation with other parameters in the fit; furthermore, the fit showed very little sensitivity to the shape parameters. Hence, a decision was made to fix the shape using the values seen in the preliminary data fit.

7.1.3.3 $B^+ \to K^+\pi^+\pi^-$

The $K\pi\pi$ final state receives cross-feed contributions from the other three $B^+ \to h^+h'^+h''^-$ modes and from self-misID, all significantly smaller than the very large expected signal yield. The partially-reconstructed components included are four-body (B^+/B^0) decays, B_s^0 decays, and the decay $B^+ \to \eta' K^+$, where $\eta' \to \pi^+\pi^-\gamma$ and the photon is lost. The shape of the $B^+ \to \eta' K^+$ component is taken from a fit to MC. This MC is generated with the intermediate decay chain $\eta' \to \rho^0\gamma$; however, a study of the decay $\eta' \to \pi^+\pi^-\gamma$ by the BESIII collaboration [112] has found that the $\pi\pi$ mass spectrum is not completely described by a single ρ^0 contribution. As an attempt to account for this, the $B^+ \to \eta' K^+$ MC is reweighted to match the dipion mass distribution seen by BESIII. The decay $B_s^0 \to D_s^-(\to K\pi\pi)\pi^+$ is taken to represent the four-body B_s^0 contribution in this case, and its shape is also taken from a fit to an MC sample. The shape parameters of the (B^+/B^0) component exhibited the same behaviour seen in the corresponding component in the $B^+ \to K^+\pi^+K^-$ model, and as such these parameters have also been fixed.

7.1.3.4 $B^+ \to \pi^+\pi^+\pi^-$

The cross-feed contributions in this final state come from $B^+ \to K^+\pi^+\pi^-$ and $B^+ \to K^+\pi^+K^-$ decays. A component to account for misidentified $B^+ \to \overline{D}^0(K^+\pi^-)\pi^+$ candidates is also

included. Finally, two partially-reconstructed components are included. The four-body B_s^0 contribution is included with a shape determined from a fit to $B_s^0 \to D_s^- (\to \pi\pi\pi)\pi^+$ MC, while the same approach is taken for the (B^+/B^0) component as used in the $B^+ \to K^+\pi^+K^-$ and $B^+ \to K^+\pi^+\pi^-$ models.

7.1.4 Data Fit Constraints

The number of free parameters in the fit to data is greatly reduced by fixing the large number of shape parameters described above to the values obtained from MC fits. Constraints on the relative yields of certain backgrounds are included to improve the stability of the fit and incorporate prior knowledge of these parameters, such as their estimated selection efficiency. Each constraint is implemented by multiplying the total likelihood PDF by a Gaussian PDF, with mean and width corresponding to the expected value of the parameter and its uncertainty. Furthermore, differences between simulation and data, *e.g.* in mass resolution, should affect all signal channels in the same way, and can therefore be accounted for by a single parameter, to be shared across the four fit models. The parameters in the fit to data, and whether they are floating, constrained, or fixed, are as follows:

- The yields of the four signal components and the four combinatorial components are left to float freely in the fit to data.
- A common offset parameter is added to the peak positions of all dCB PDFs, which have been fixed to the values extracted from MC. This offset is left to float in the fit.
- A scaling factor is applied to the width parameters of all dCB PDFs, after also being fixed to the values in MC. The scaling factor is also left to float.
- The yield of each cross-feed contribution is Gaussian-constrained to the product of the yield in its signal spectrum and the ratio of the corresponding phase-integrated efficiencies. The effect of the shift in mass due to misID is accounted for when determining the efficiency ratios.
- The slope parameter of the exponential describing the combinatorial background is left to float in the data fit. Furthermore it is allowed to take different values in each final-state mass spectrum.
- The ARGUS threshold parameters are fixed to the difference in the nominal masses of the parent and the missing particle. In the case of $B^+ \to \eta' K^+$, the peak position of the signal PDF is taken as the threshold.

- The Gaussian resolution PDFs that are convoluted with the ARGUS PDFs share the widths of the respective signal PDF.
- The yields of the partially-reconstructed (B^+/B^0) contributions in the KKK, $K\pi\pi$, and $\pi\pi\pi$ final states, and the B_s^0 contribution in the $K\pi\pi$ final state are left to float in the fit.
- The yield of $B^+ \to \eta' K^+$ is Gaussian-constrained to the product of the yield of $B^+ \to K^+\pi^+\pi^-$, the ratio of selection efficiencies, and the ratio of branching fractions [32].
- In the $B^+ \to K^+\pi^+K^-$ model, the ratio between charmless and charmed partially-reconstructed B_s^0 contributions is Gaussian constrained, taking into account their relative branching fractions [32] and selection efficiencies.
- In the $B^+ \to K^+\pi^+K^-$ model, the ratio between charmless (B^+/B^0) and charmless B_s^0 contributions is also Gaussian constrained. The ratio is estimated from the yields obtained in the LHCb analysis described in Ref. [111]. The yields are $N(B^0 \to K^+\pi^-K^-\pi^+) = 1013 \pm 49$ and $N(B_s^0 \to K^+\pi^-K^-\pi^+) = 6080 \pm 83$. A factor of two is included to account for both (B^+/B^0) contributions.
- The yield of the charmed partially-reconstructed B_s^0 contribution in the $B^+ \to K^+\pi^+K^-$ model is a floating parameter.
- Finally, the yields of both the $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ misID and B_s^0 partially-reconstructed components in the $B^+ \to \pi^+\pi^+\pi^-$ are Gaussian-constrained to the signal yield.

7.2 Fit Results

The results of the simultaneous extended maximum likelihood fit to four data distributions are shown in Figures 7.4 and 7.5, in linear and logarithmic scale respectively. Good overall agreement between the fit model and data is seen. A trend in the pull distribution is seen in the low-mass region of the $m(K\pi\pi)$ spectrum, identified as being due to the use of the single-parameter form of the ARGUS function to describe the (B^+/B^0) partially-reconstructed background. The potential effect of this feature in the final result is evaluated as part of the model systematic uncertainties described in Chapter 8.

The yields extracted from the fit are shown in Table 7.2. It is worth noting that many background yields do not enter the fit explicitly, but are constructed from the product of a signal yield parameter and their efficiency ratios. The background yields are still of interest, and so are explicitly listed in Table 7.2. The nuisance parameters extracted from the fit are given in Table 7.3.

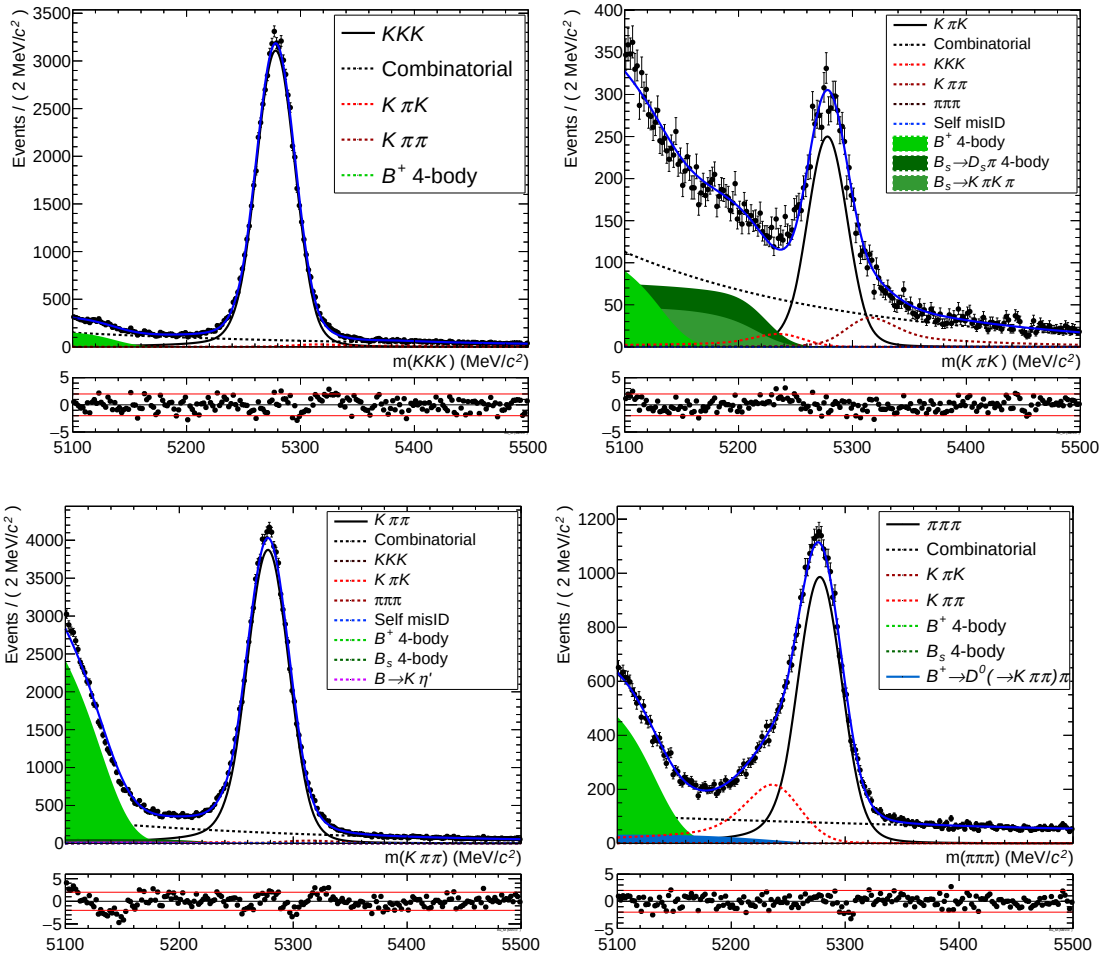

Figure 7.4: Result of the simultaneous maximum likelihood fit to the invariant mass distributions of $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right).

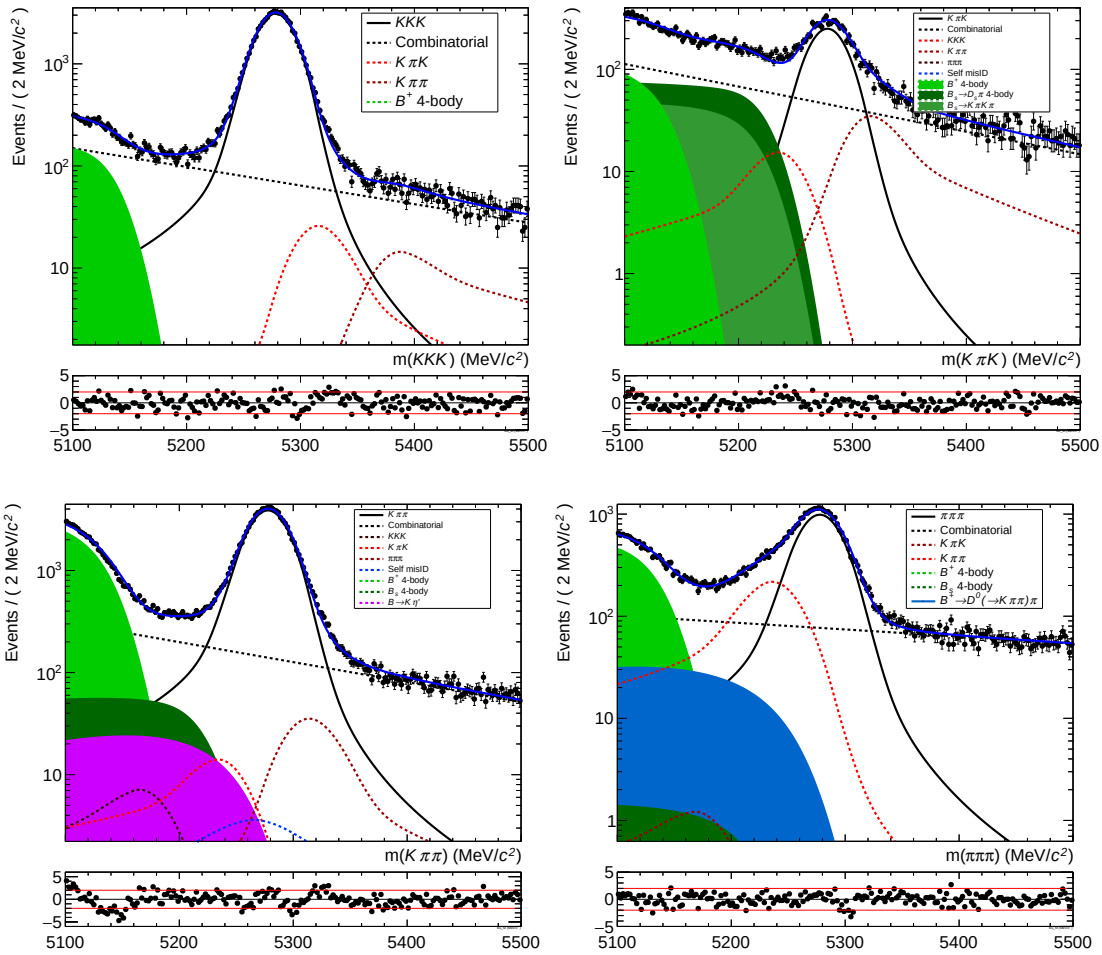

Figure 7.5: Log-scale simultaneous fit results in the $B^+ \to K^+K^+K^-$ (top-left), $B^+ \to K^+\pi^+K^-$ (top-right), $B^+ \to K^+\pi^+\pi^-$ (bottom-left), and $B^+ \to \pi^+\pi^+\pi^-$ (bottom-right) mass spectra.

Table 7.2: Fitted yields obtained from the simultaneous fit to data. The * symbol indicates the yields that are not explicit parameters in the fit. The associated errors take into account the covariances in the fit.

Mass Spectrum	Parameters	Extracted Values
	N_{signal}	69313 ± 285
	N_{comb}	14435 ± 229
KKK	$N_{K\pi K}$*	886 ± 21
	$N_{K\pi\pi}$*	700 ± 11
	$N_{B\,\text{PRB}}$	2730 ± 130
	N_{signal}	5758 ± 135
	N_{comb}	9664 ± 745
	N_{KKK}*	604 ± 3
	$N_{K\pi\pi}$*	1445 ± 10
$K\pi K$	$N_{\pi\pi\pi}$*	8.4 ± 0.3
	$N_{\pi KK}$*	2.2 ± 0.3
	$N_{B\,\text{PRB}}$*	1490 ± 122
	$N_{B_s^0\,\text{PRB}}$(charmed)	4286 ± 548
	$N_{B_s^0\,\text{PRB}}$(charmless)*	2440 ± 199
	N_{signal}	94946 ± 435
	N_{comb}	28873 ± 1450
	N_{KKK}*	274 ± 2
	$N_{K\pi K}$*	620 ± 15
$K\pi\pi$	$N_{\pi\pi\pi}$*	1245 ± 10
	$N_{\pi K\pi}$*	335 ± 7
	$N_{B\,\text{PRB}}$	39553 ± 323
	$N_{B_s^0\,\text{PRB}}$	3274 ± 978
	$N_{K\eta'}$*	1636 ± 114
	N_{signal}	25484 ± 205
	N_{comb}	15104 ± 348
	$N_{K\pi K}$*	50 ± 2
$\pi\pi\pi$	$N_{K\pi\pi}$*	8308 ± 47
	$N_{B\,\text{PRB}}$	8201 ± 149
	$N_{B_s^0\,\text{PRB}}$*	69 ± 10
	$N_{D(\to K\pi)\pi}$*	1953 ± 199

Table 7.3: Nuisance parameters obtained from the simultaneous fit to data.

Mass Spectrum	Parameters	Extracted Values
	σ_{scale}	$1.144 \pm 0.003\,\text{MeV}/c^2$
	δ_{offset}	$-2.58 \pm 0.05\,\text{MeV}/c^2$
KKK	$c_{\text{combinatorial}}$	$(-42.1 \pm 1.2)\times 10^{-4}$
	$c_{B\,\text{PRB}}$	-20.5 ± 9.3
	$\epsilon(K\pi K \to KKK)$	$(15.385 \pm 0.040)\%$
	$\epsilon(K\pi\pi \to KKK)$	$(0.737 \pm 0.011)\%$
$K\pi K$	$c_{\text{combinatorial}}$	$(-50.9 \pm 4.6)\times 10^{-4}$
	$\epsilon(KKK \to K\pi K)$	$(0.871 \pm 0.003)\%$
	$\epsilon(K\pi\pi \to K\pi K)$	$(1.522 \pm 0.008)\%$
	$\epsilon(\pi\pi\pi \to K\pi K)$	$(0.0329 \pm 0.0013)\%$
	$\epsilon(\pi KK \to K\pi K)$	$(0.0382 \pm 0.0054)\%$
	$(B_s^0 \to K^+\pi^-K^-\pi^+)/(B_s^0 \to D_s^-\pi^+)$	$(56.9 \pm 8.7)\%$
	$(B \to K\pi K\pi)/(B_s^0 \to K^+\pi^-K^-\pi^+)$	$(61.0 \pm 4.1)\%$
$K\pi\pi$	$c_{\text{combinatorial}}$	$(-45.1 \pm 2.9)\times 10^{-4}$
	$\epsilon(KKK \to K\pi\pi)$	$(0.3946 \pm 0.0025)\%$
	$\epsilon(K\pi K \to K\pi\pi)$	$(10.771 \pm 0.069)\%$
	$\epsilon(\pi\pi\pi \to K\pi\pi)$	$(4.886 \pm 0.010)\%$
	$\epsilon(\pi K\pi \to K\pi\pi)$	$(0.3524 \pm 0.0068)\%$
	$(B^+ \to \eta' K^+)/(B^+ \to K^+\pi^+\pi^-)$	$(1.72 \pm 0.12)\%$
$\pi\pi\pi$	$c_{\text{combinatorial}}$	$(-15.8 \pm 1.7)\times 10^{-4}$
	$\epsilon(K\pi K \to \pi\pi\pi)$	$(0.872 \pm 0.027)\%$
	$\epsilon(K\pi\pi \to \pi\pi\pi)$	$(8.751 \pm 0.030)\%$
	$(B_s^0 \to D_s^-(\to\pi\pi\pi)\pi^+)/(B^+ \to \pi^+\pi^+\pi^-)$	$(0.269 \pm 0.040)\%$
	$(B^+ \to \bar{D}^0(\to K\pi)\pi^+)/(B^+ \to \pi^+\pi^+\pi^-)$	$(7.66 \pm 0.77)\%$

7.2.1 Signal Yield Correction

Before determining the ratios of branching fractions, the raw signal yields extracted from the fit to data must be corrected to account for the differences in the total detection efficiency across the four signal channels, per Equation 4.1. The total efficiency maps evaluated in Chapter 6 allow the efficiency correction to be performed on an event-by-event basis, by also determining the variation of the background-subtracted yields as a function of Dalitz plot position. Equation 4.1 can be rewritten as

$$\frac{\mathcal{B}(B^+ \to h^+h'^+h''^-)}{\mathcal{B}(B^+ \to K^+K^+K^-)} = \frac{\mathcal{N}^{\rm corr}_{hhh}}{\mathcal{N}^{\rm corr}_{KKK}}, \quad \mathcal{N}^{\rm corr} = \sum_{i \in N} \frac{w_i}{\epsilon_i}, \tag{7.3}$$

where the sum is performed over the number of candidates in the respective final state, w_i is the signal *sWeight* associated with the i-th candidate, and ϵ_i the total efficiency assigned to the candidate from its position in the DP.

A minor difficulty arises from imposing Gaussian constraints on the yields of cross-feed and some partially-reconstructed backgrounds. In the 'standard' *sWeight* procedure, the yields of the different contributions must be freely-floating in the fit, with all other fit parameters kept constant [113]. An extension of the method exists (described in Appendix B of Ref. [113]), that allows the inclusion of contributions with constant yields if their distributions in the variables of interest (the DP variables in this case) are known. A new invariant mass fit is performed individually for each signal channel, in which the nuisance parameters are fixed to the values extracted from the nominal simultaneous fit and the yields of the Gaussian-constrained backgrounds are explicitly fixed to the values given in Table 7.3. In the $B^+ \to K^+\pi^+K^-$ model, where the partially-reconstructed backgrounds are Gaussian-constrained relative to one another but the yield of the total partially-reconstructed contribution is freely-floating, a decision was made to keep the relative yields constant but keep the total yield floating, to avoid having to describe the Dalitz plot distribution of those backgrounds. The modified efficiency-corrected yield takes the form

$$\mathcal{N}^{\rm corr} = \sum_{\rm bins} \frac{c_n M_0(m', \theta') + \sum_{i \subset {\rm bin}} w_i}{\epsilon_{\rm bin}}, \tag{7.4}$$

where M_0 is the normalised total distribution of the fixed backgrounds, and c_n quantifies the impact of the fixed backgrounds by taking into account the signal yield and the covariances between floating yields.

The efficiency-corrected yields are given in Table 7.4. The associated statistical uncertainties also require some care in their determination. The *sWeight* fits have fewer free parameters, since all nuisance parameters are made constant, than the nominal simultaneous fit; this results

in a difference between the two fits in the statistical uncertainty associated with each signal yield. The procedure used to take this into account involves incorporating the difference in uncertainties between the two fits, $\sigma^{\text{shape}}(N) = \sqrt{\sigma^{\text{sim.fit}}(N)^2 - \sigma^{\text{sWeights}}(N)^2}$, into the total uncertainty [114]. The statistical uncertainty on the efficiency-corrected yields, as given in Table 7.4, takes the form [113]

$$\sigma^{\text{corr}}(\mathcal{N}^{\text{corr}}) = \sqrt{\sum_i \left(\frac{w_i}{\epsilon_i}\right)^2 + \left(\frac{\mathcal{N}^{\text{corr}}}{N}\sigma^{\text{shape}}(N)\right)^2}. \tag{7.5}$$

Table 7.4: Efficiency-corrected signal yields and associated statistical uncertainties.

Signal Channel	Nominal Fitted Yield	$\mathcal{N}^{\text{corr}}$
$B^+ \to K^+K^+K^-$	$69\,310 \pm 290$	$9\,323\,000 \pm 40\,100$
$B^+ \to K^+\pi^+K^-$	$5\,760 \pm 140$	$1\,417\,700 \pm 36\,500$
$B^+ \to K^+\pi^+\pi^-$	$94\,950 \pm 440$	$15\,983\,300 \pm 79\,400$
$B^+ \to \pi^+\pi^+\pi^-$	$25\,480 \pm 210$	$4\,667\,100 \pm 42\,400$

At this point, the branching fraction ratios can first be determined using Equation 7.3. The branching fraction ratios are given in Table 7.5 with their statistical uncertainties; the systematic uncertainties considered in this analysis are described in the following chapter.

Table 7.5: $B^+ \to h^+h'^+h''^-$ branching fraction ratios. The values in each cell correspond to the efficiency-corrected yield of the channel in that row divided by that of the channel in the corresponding column. The uncertainties are statistical only.

row/column	KKK	$K\pi K$	$K\pi\pi$	$\pi\pi\pi$
KKK	–	6.58 ± 0.17	0.5833 ± 0.0038	1.998 ± 0.020
$K\pi K$	0.1521 ± 0.0040	–	0.0887 ± 0.0023	0.3038 ± 0.0083
$K\pi\pi$	1.714 ± 0.011	11.27 ± 0.30	–	3.425 ± 0.035
$\pi\pi\pi$	0.5006 ± 0.0050	3.292 ± 0.090	0.2920 ± 0.0030	–

CHAPTER 8

Systematic Uncertainties

“ If you thought before science was certain, well, that’s an error on your part. ”

Richard Feynman, *The Character of Physical Law*

Sources of systematic uncertainty can introduce biases in the calculation of the results. However, by evaluating the ratios of $B^+ \to h^+h'^+h''^-$ branching fractions relative to one another, many potential sources of systematic effects are expected to cancel out. The sources of systematic uncertainty that have been considered in this analysis are detailed in this chapter. Furthermore, cross-checks are performed, by re-evaluating the branching fraction ratios with subsets of the total dataset, to determine the consistency of the final results.

8.1 Systematic Uncertainties

The systematic uncertainties considered can be divided into those related to the efficiency of the candidate selection, and those related to the fit model. Given the number of those associated with the selection efficiency, they have been grouped under those generally related to the efficiency, those specifically related to the efficiency corrections applied, and those associated with the `PIDCalib` procedure. The absolute values of the systematic uncertainties in these categories are given in Tables 8.1, 8.2, and 8.3 , respectively. The values of the systematic uncertainties related to the fit model are given in Table 8.4. Finally, a summary of all the systematic uncertainties for three of the ratios measured is given in Table 8.5, to ease the comparison between the uncertainties that have been determined.

8.1.1 Data / MC Differences

Two variables (the p_{T} asymmetry and the minimum IP χ^2) included in the MVA showed some mismatch in the distributions seen in MC and the background-subtracted data, shown in Figures 5.4 and 5.5. The effect of the data / MC difference on the efficiency variation across the Dalitz plot is estimated by reweighting the MC, so that it matches the distribution of those two variables in *sWeighted* data. The reweighted efficiency maps are then propagated through to determine branching fraction ratios; the uncertainty is then the difference between the nominal result and that determined with this reweighting.

8.1.2 Vetoes

The systematic uncertainty due to the fraction of signal candidates lost within the D^0 veto window is evaluated by increasing the window from $30\,\mathrm{MeV}/c^2$ to $45\,\mathrm{MeV}/c^2$. Both the fit to the data and the evaluation of the efficiency maps are repeated taking into account the larger veto window.

8.1.3 Dalitz Plot Binning

The choice of binning used to represent the distribution of the efficiency across the Dalitz plot is a potential source of uncertainty. The granularity of the binning is chosen with the assumption that the variation of the efficiency within each bin is negligible. The potential effect is estimated by reducing the number of bins in each dimension from 30 to 20 and re-evaluating the results.

8.1.4 Finite MC Statistics

The efficiencies determined in Chapter 6 have an associated statistical uncertainty that originates from the finite sample size of the MC samples used to evaluate them. Since the numerator / denominator in the efficiency determination of the total dataset is a weighted average of the number of selected / total MC candidates in each subsample (2011 / 2012, MagDown / MagUp), the statistical uncertainty is evaluated with the following procedure [106]. The efficiency of the total dataset and its variance are given by

$$E = \frac{K}{N} = \frac{\sum_i w_i k_i}{\sum_i w_i n_i}, \qquad V = \left(\sum_i (w_i n_i)^2 v_i\right) / \left(\sum_i w_i n_i\right)^2, \tag{8.1}$$

where k_i, n_i are the number of selected / total MC candidates in a given subsample, w_i the weight given to that subsample, and v_i is the variance of the subsample. With the efficiency and variance of the total dataset, a Beta distribution is then constructed from which the efficiency can be resampled, with the correct statistical properties. Note that the the variances of the individual

subsamples, v_i, are also determined from a Beta distribution, $\mathrm{Be}(k+\frac{1}{2}, n-k+\frac{1}{2})$. The effect on the branching fraction ratios is determined by calculating the statistical uncertainty in each bin of the four efficiency maps, creating a set of resampled efficiency maps, and using each set to re-evaluate the ratios. The systematic uncertainty on the results is taken as the standard deviation of the set of branching fraction ratios.

Table 8.1: Absolute systematic uncertainties related to the candidate selection and its efficiency.

Branching Fraction Ratio	Data / MC (10^{-2})	Veto (10^{-2})	DP Binning (10^{-2})	MC Statistics (10^{-2})
$K\pi K$ / KKK	0.033	0.43	0.052	0.032
$K\pi\pi$ / KKK	0.61	5.4	0.18	0.33
$\pi\pi\pi$ / KKK	0.16	1.3	0.39	0.10
KKK / $K\pi K$	1.4	19	2.2	1.4
$K\pi\pi$ / $K\pi K$	1.6	4.2	2.7	2.5
$\pi\pi\pi$ / $K\pi K$	0.37	0.50	1.5	0.78
KKK / $K\pi\pi$	0.21	1.9	0.062	0.11
$K\pi K$ / $K\pi\pi$	0.013	0.033	0.021	0.020
$\pi\pi\pi$ / $K\pi\pi$	0.0082	0.15	0.20	0.066
KKK / $\pi\pi\pi$	0.66	5.5	1.6	0.41
$K\pi K$ / $\pi\pi\pi$	0.035	0.046	0.14	0.073
$K\pi\pi$ / $\pi\pi\pi$	0.096	1.8	2.3	0.78

8.1.5 Trigger

A correction for the efficiency of the L0 TIS trigger was described in Section 6.5, which involved reweighting the MC samples to reflect the number of events firing the different TIS decisions seen in data. A systematic uncertainty is associated with this correction by varying the MC weights obtained, shown in Figure 6.7, within their uncertainties.

The L0Hadron TOS trigger response has also been corrected for, with the use of calibration efficiency tables determined from $D^{*-} \to \overline{D}^0 (\to K^+\pi^-)\,\pi^-$ decays. The systematic uncertainty is obtained from the use of an alternative set of calibration tables, determined with kaon and pion samples from $B^0 \to J/\psi K^{*0}$ decays, to obtain new correction maps. The absolute differences between results obtained using the nominal and alternative corrections are taken as the uncertainties.

8.1.6 Tracking

A systematic uncertainty is assigned to the tracking correction procedure, by varying the values of the track-level corrections, shown in Figure 6.6, within their uncertainties and re-evaluating

the correction.

8.1.7 B Kinematics

A correction applied to account for residual data / MC differences in the kinematics of the B meson was obtained by matching the distribution of B momentum and pseudorapidity in MC to that of $B^+ \to K^+K^+K^-$ background-subtracted data. An alternative correction is performed using the background-subtracted $B^+ \to K^+\pi^+\pi^-$ sample, with the absolute difference in the ratios determined with the two methods being taken as the uncertainty.

Table 8.2: Absolute systematic uncertainties concerning the corrections to the total efficiency.

Branching Fraction Ratio	L0 TIS (10^{-2})	L0 TOS (10^{-2})	Tracking (10^{-2})	Kinematics (10^{-2})
$K\pi K$ / KKK	0.11	0.018	0.006	0.005
$K\pi\pi$ / KKK	0.36	0.34	0.20	0.26
$\pi\pi\pi$ / KKK	0.18	0.27	0.13	0.12
KKK / $K\pi K$	4.6	0.79	0.28	0.23
$K\pi\pi$ / $K\pi K$	8.1	3.6	0.85	1.3
$\pi\pi\pi$ / $K\pi K$	2.5	2.2	0.72	0.69
KKK / $K\pi\pi$	0.12	0.11	0.067	0.089
$K\pi K$ / $K\pi\pi$	0.064	0.028	0.007	0.010
$\pi\pi\pi$ / $K\pi\pi$	0.099	0.10	0.042	0.027
KKK / $\pi\pi\pi$	0.72	1.1	0.51	0.49
$K\pi K$ / $\pi\pi\pi$	0.23	0.20	0.066	0.063
$K\pi\pi$ / $\pi\pi\pi$	1.6	1.2	0.49	0.31

8.1.8 PID Systematics

The `PIDCalib` procedure depends on two assumptions: the response of the RICH detectors can be fully parameterised in terms of the variables used in the `PIDCalib` binning (p, η, and `BestTracks`), and the efficiency variation within each bin is minimal. The first assumption is generally thought to be a reasonable one; the second is tested by evaluating the PID efficiencies with a different binning choice in `PIDCalib`. Both the granularity of the binning and the bin boundaries are varied with respect to the nominal binning defined in Table 5.2, and this alternative binning is used to recalculate the PID efficiency.

An uncertainty is associated with the use of *sWeights* to subtract the background present in the calibration samples; an uncertainty of 0.1% per track in the decay channel being studied is recommended by the `PIDCalib` developers. Since ratios are being determined in this case, a

0.1% uncertainty is taken for each track that differs in type between numerator and denominator in the ratio. For example, an uncertainty of 0.2% is assigned to the ratio $K\pi\pi$ / KKK.

Finally, the statistical uncertainty due to the calibration samples used is calculated by resampling the `PIDCalib` efficiency histograms multiple times. This is achieved by resampling from Beta distributions whose parameters are the number of selected / total number of candidates in each bin. In the cases in which multiple tracks use the same performance histogram, the same set of resampled histograms is shared among the tracks, to preserve the correlations present in the original procedure.

Table 8.3: Absolute systematic uncertainties associated with `PIDCalib`.

Branching Fraction Ratio	PID binning (10^{-2})	PID *sWeights* (10^{-2})	PID statistics (10^{-2})
$K\pi K$ / KKK	0.080	0.015	0.0038
$K\pi\pi$ / KKK	0.057	0.34	0.054
$\pi\pi\pi$ / KKK	0.030	0.15	0.038
KKK / $K\pi K$	3.5	0.66	0.16
$K\pi\pi$ / $K\pi K$	5.5	1.1	0.40
$\pi\pi\pi$ / $K\pi K$	1.5	0.66	0.24
KKK / $K\pi\pi$	0.019	0.12	0.019
$K\pi K$ / $K\pi\pi$	0.043	0.0089	0.0032
$\pi\pi\pi$ / $K\pi\pi$	0.0075	0.029	0.024
KKK / $\pi\pi\pi$	0.12	0.60	0.15
$K\pi K$ / $\pi\pi\pi$	0.14	0.061	0.022
$K\pi\pi$ / $\pi\pi\pi$	0.088	0.34	0.28

8.1.9 Choice of Fit Model

The fit model uses double Crystal Ball functions to model both signal and cross-feed components, exponential functions to model the combinatorial background, and ARGUS functions to describe the partially-reconstructed contributions. Other suitable parameterisations could have been chosen instead, and so a systematic uncertainty is associated with the particular choices made.

To evaluate this uncertainty, three alternative models are built in which a different choice of parameterisation is picked for one of the categories of fit component. In 'Model I' all dCB functions are replaced with double Hypatia functions. In 'Model II' the exponential functions are replaced by second-order Chebyshev polynomials. 'Model III' implements different partially-reconstructed parameterisations: for components that have the ARGUS parameters fixed to MC, the ARGUS function is replaced by a `RooKeysPDF` kernel density estimation; the other ARGUS

components are made into generalised ARGUS functions, by allowing the extra parameter free to float.

8.1.10 Fixed Parameters

The parameters that are fixed in the fit to data, that have been taken from fits to MC, are varied within their uncertainties by using the covariance matrix of those MC fits. Multiple sets of resampled parameters are used to fit the data, the results of which are used to obtain branching fraction ratios, with the standard deviation of the set of ratios being taken as the systematic uncertainty on the results.

8.1.11 Fit Bias

Multiple pseudoexperiments are generated based on the values of the fit parameters extracted from the nominal fit to data. Subsequently, each pseudoexperiment is fitted with the nominal fit model and the distributions of the extracted signal yields are fitted with Gaussian distributions. The difference between the mean of each Gaussian and the corresponding nominal signal yield is taken as a systematic uncertainty.

Table 8.4: Absolute systematic uncertainties associated with the fit model.

Branching Fraction Ratio	Model I (10^{-2})	Model II (10^{-2})	Model III (10^{-2})	Fixed Param. (10^{-2})	Fit Bias (10^{-2})
$K\pi K$ / KKK	0.12	0.38	0.62	0.11	0.012
$K\pi\pi$ / KKK	1.5	1.1	1.8	0.71	0.038
$\pi\pi\pi$ / KKK	0.038	0.031	0.029	0.24	0.019
KKK / $K\pi K$	5.0	17	26	4.8	0.54
$K\pi\pi$ / $K\pi K$	18	21	56	9.8	0.94
$\pi\pi\pi$ / $K\pi K$	2.7	8.2	13	3.0	0.29
KKK / $K\pi\pi$	0.52	0.39	0.62	0.24	0.013
$K\pi K$ / $K\pi\pi$	0.15	0.16	0.46	0.076	0.007
$\pi\pi\pi$ / $K\pi\pi$	0.24	0.18	0.29	0.18	0.012
KKK / $\pi\pi\pi$	0.15	0.12	0.12	0.92	0.075
$K\pi K$ / $\pi\pi\pi$	0.26	0.74	1.3	0.26	0.027
$K\pi\pi$ / $\pi\pi\pi$	2.8	2.0	3.4	2.0	0.14

Table 8.5: Summary of all systematic uncertainties associated with the branching fraction ratios with respect to the $B^+ \to K^+K^+K^-$ decay.

Systematic Uncertainty (10^{-2})	$K\pi K$ / KKK	$K\pi\pi$ / KKK	$\pi\pi\pi$ / KKK
Data / MC	0.033	0.61	0.16
Veto	0.43	5.4	1.3
DP Binning	0.052	0.18	0.39
MC Statistics	0.032	0.33	0.10
L0 TIS	0.11	0.36	0.18
L0 TOS	0.018	0.34	0.27
Tracking	0.006	0.20	0.13
Kinematics	0.005	0.26	0.12
PID binning	0.080	0.057	0.030
PID *sWeights*	0.015	0.34	0.15
PID statistics	0.0038	0.054	0.038
Model I	0.12	1.5	0.038
Model II	0.38	1.1	0.031
Model III	0.62	1.8	0.029
Fixed Param.	0.11	0.71	0.24
Fit Bias	0.012	0.038	0.019

8.2 Cross-Checks

Subsets of the dataset used in this analysis have been taken under different conditions, which nonetheless should not affect the results obtained. For example, the results should be consistent when looking at data taken at one centre-of-mass energy *vs.* another (*i.e.* 2011 *vs.* 2012), and also when considering the two magnet polarisations separately. The branching fractions obtained for each cross-check category are compared against the nominal values in Figure 8.1 and the results are found to be consistent.

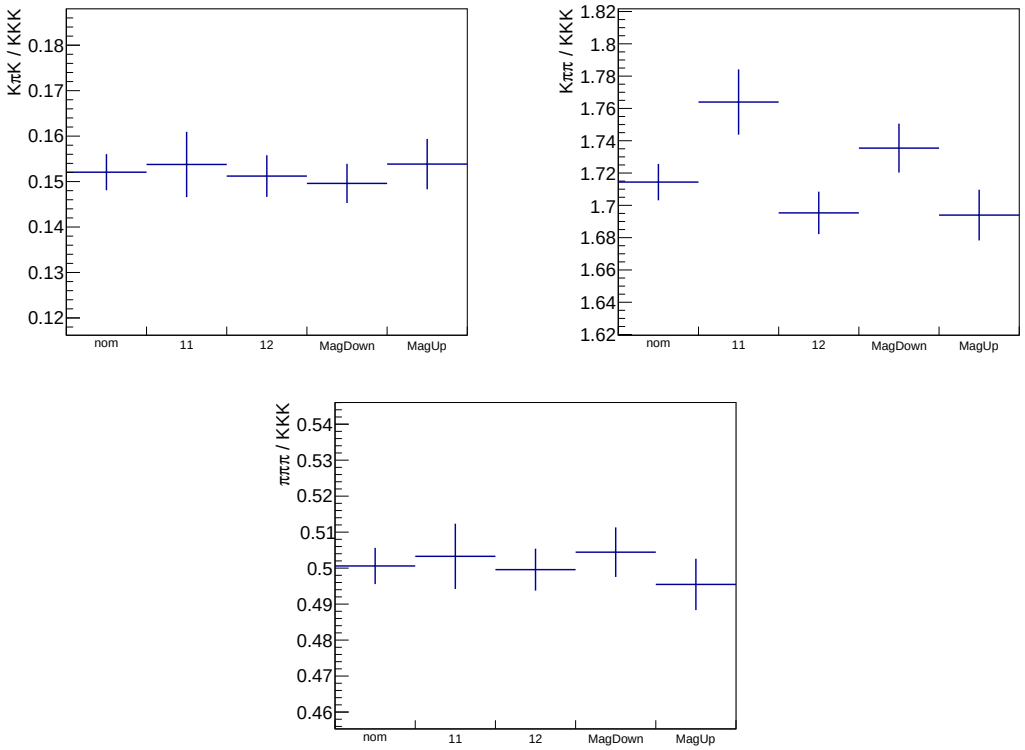

Figure 8.1: Branching fraction ratios as a function of cross-check subcategory. Values are given for the ratios $K\pi K$ / KKK(top-left), $K\pi\pi$ / KKK(top-right), and $\pi\pi\pi$ / KKK(bottom).

CHAPTER 9

Conclusions

> *Life is like topography, Hobbes. There are summits of happiness and success, flat stretches of boring routine and valleys of frustration and failure.*
>
> Bill Watterson, *Calvin and Hobbes*

This thesis documents a measurement of the relative branching fractions of the charmless decays $B^+ \to K^+K^+K^-$, $B^+ \to K^+\pi^+K^-$, $B^+ \to K^+\pi^+\pi^-$, and $B^+ \to \pi^+\pi^+\pi^-$, with the $3\,\text{fb}^{-1}$ of data collected by the LHCb experiment in the years 2011 and 2012. The complete set of possible branching fraction ratios are given in Table 9.1, along with the corresponding ratios determined using the current-best absolute branching fraction values. However, given the number of decay channels studied, the number of fully-independent ratios that can be determined is three. It can be convenient to consider the set of branching fractions relative to that of the $B^+ \to K^+K^+K^-$ decay as that independent set, since the absolute branching fraction of $B^+ \to K^+K^+K^-$ is currently the most precisely known:

$$\begin{aligned}
\mathcal{B}(B^+ \to K^+\pi^+K^-) \,/\, \mathcal{B}(B^+ \to K^+K^+K^-) &= 0.152 \pm 0.004\,(\text{stat}) \pm 0.009\,(\text{syst})\,, \\
\mathcal{B}(B^+ \to K^+\pi^+\pi^-) \,/\, \mathcal{B}(B^+ \to K^+K^+K^-) &= 1.714 \pm 0.011\,(\text{stat}) \pm 0.061\,(\text{syst})\,, \\
\mathcal{B}(B^+ \to \pi^+\pi^+\pi^-) \,/\, \mathcal{B}(B^+ \to K^+K^+K^-) &= 0.501 \pm 0.005\,(\text{stat}) \pm 0.015\,(\text{syst})\,.
\end{aligned}$$

The amplitude analyses of the $B^+ \to K^+\pi^+K^-$ and $B^+ \to \pi^+\pi^+\pi^-$ decay channels that have been recently performed by the LHCb collaboration have been previously mentioned in

Table 9.1: Final branching fraction ratios, including statistical and systematic uncertainties. The current world-average values, determined using the values in Table 2.1, are given for comparison.

Branching Fraction Ratio	Value	Current W.A Value
$K\pi K$ / KKK	$0.152 \pm 0.004\,(\text{stat}) \pm 0.009\,(\text{syst})$	0.153 ± 0.013
$K\pi\pi$ / KKK	$1.714 \pm 0.011\,(\text{stat}) \pm 0.061\,(\text{syst})$	1.50 ± 0.11
$\pi\pi\pi$ / KKK	$0.501 \pm 0.005\,(\text{stat}) \pm 0.015\,(\text{syst})$	0.447 ± 0.045
KKK / $K\pi K$	$6.58 \pm 0.17\,(\text{stat}) \pm 0.37\,(\text{syst})$	6.54 ± 0.57
$K\pi\pi$ / $K\pi K$	$11.27 \pm 0.30\,(\text{stat}) \pm 0.64\,(\text{syst})$	9.81 ± 0.94
$\pi\pi\pi$ / $K\pi K$	$3.29 \pm 0.09\,(\text{stat}) \pm 0.17\,(\text{syst})$	2.92 ± 0.35
KKK / $K\pi\pi$	$0.583 \pm 0.004\ (\text{stat}) \pm 0.022\ (\text{syst})$	0.667 ± 0.047
$K\pi K$ / $K\pi\pi$	$0.0887 \pm 0.0023\,(\text{stat}) \pm 0.0053\,(\text{syst})$	0.1020 ± 0.0098
$\pi\pi\pi$ / $K\pi\pi$	$0.2920 \pm 0.0030\,(\text{stat}) \pm 0.0054\,(\text{syst})$	0.298 ± 0.032
KKK / $\pi\pi\pi$	$1.998 \pm 0.020\,(\text{stat}) \pm 0.060\,(\text{syst})$	2.24 ± 0.23
$K\pi K$ / $\pi\pi\pi$	$0.304 \pm 0.009\,(\text{stat}) \pm 0.016\,(\text{syst})$	0.342 ± 0.041
$K\pi\pi$ / $\pi\pi\pi$	$3.425 \pm 0.035\,(\text{stat}) \pm 0.063\,(\text{syst})$	3.36 ± 0.36

this thesis. At this point, it is interesting to consider the improvement achieved in the precision of quasi-two-body branching fractions by combining the results of those two analyses and the measurements presented in this work. The improvement seen in the branching fraction of the decay $B^+ \to \rho^0(770)\pi^+$ is given as an example.

The current (not yet including the LHCb result) world average value of this branching fraction is $\mathcal{B}(B^+ \to \rho^0(770)\pi^+) = (8.3^{+1.2}_{-1.3}) \times 10^{-6}$ [33]. The relative uncertainty on this average is 16%. The amplitude analysis of $B^+ \to \pi^+\pi^+\pi^-$ has been performed via three independent methods; the values given for the $\rho^0(770)$ fit fraction are consistent across the three methods. For the purposes of this comparison, the fit fraction determined with the isobar model is used, $FF(\rho^0(770)) = (55.5 \pm 2.6)\%$ [17, 18]. The world-average absolute branching fraction value $\mathcal{B}(B^+ \to K^+K^+K^-) = (34.0 \pm 1.4) \times 10^{-6}$ [32] is used in combination with the ratio $\pi\pi\pi$ / KKK measured in this work.

The $B^+ \to \rho^0(770)\pi^+$ branching fraction can then be determined using a slightly-

modified version of Equation 2.20,

$$\begin{aligned}\mathcal{B}(B^+ \to \rho^0(770)\pi^+) &= \mathcal{B}(B^+ \to K^+K^+K^-) \cdot R\left(\frac{\pi\pi\pi}{KKK}\right) \cdot FF_{\rho^0} \\ &= (9.5 \pm 0.6) \times 10^{-6},\end{aligned} \tag{9.1}$$

with the uncertainty obtained in this calculation corresponding to a relative uncertainty of 6%. The same procedure can be used on each quasi-two-body state present in the $B^+ \to \pi^+\pi^+\pi^-$ and $B^+ \to K^+\pi^+K^-$ amplitude models.

The coming years will see further progress in the study of $B^+ \to h^+h'^+h''^-$ decays. Work is ongoing on the LHCb amplitude analyses of $B^+ \to K^+K^+K^-$ and $B^+ \to K^+\pi^+\pi^-$ decays. With the combined Run 1 and Run 2 datasets, the signal yields available to the two analyses are two orders of magnitude larger than those that were available in the B factories. Similarly to the $B^+ \to \pi^+\pi^+\pi^-$ case, the limiting factor in the precision of these two measurements is most likely to be the systematic uncertainties in the amplitude model. Model-independent CP violation analyses, in the style of Ref. [15], can potentially avoid becoming systematically limited; however, it is more difficult to associate such results to specific intermediate processes [115].

Looking further afield, prospects exist for combining the results obtained from $B^+ \to h^+h'^+h''^-$ decays with measurements of isospin-related decay channels such as $B^0 \to K^0_S h^+h'^-$ and $B^0 \to h^+h'^-\pi^0$. Constraints on CKM parameters, for example the CKM angles α and γ, can be determined by considering isospin relations between families of quasi-two-body decays such as $B \to \rho\pi$ and $B \to K^*\pi$ [116, 117]. Precise measurements of the branching fractions involved, such as those presented in this thesis, will be an essential element of this programme.

Bibliography

[1] C. D. Anderson, *The positive electron*, Phys. Rev. **43** (1933) 491.

[2] G. E. R. Lloyd, *Early Greek science: Thales to Aristotle*, W. W. Norton & Company, 1974.

[3] B. Pullman, *The atom in the history of human thought*, Oxford University Press, 1998.

[4] A. Einstein, *Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen*, Annalen der Physik **322** (1905) 549.

[5] J. Perrin, *Mouvement brownien et réalité moléculaire*, Ann. Chim. Phys. 8e série **18** (1909) 1.

[6] J. J. Thomson, *Cathode rays*, Lond. Edinb. Dubl. Phil. Mag. **44** (1897) 293.

[7] E. Rutherford, *The scattering of α and β particles by matter and the structure of the atom*, Lond. Edinb. Dubl. Phil. Mag. **21** (1911) 669.

[8] D. Hanneke, S. Fogwell, and G. Gabrielse, *New measurement of the electron magnetic moment and the fine structure constant*, Phys. Rev. Lett. **100** (2008) 120801, arXiv:0801.1134.

[9] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, *Tenth-order electron anomalous magnetic moment — Contribution of diagrams without closed lepton loops*, Phys. Rev. **D91** (2015) 033006, arXiv:1412.8284.

[10] 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.

[11] 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.

[12] Planck collaboration, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209.

[13] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[14] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and antimatter in the universe*, New J. Phys. **14** (2012) 095012, arXiv:1204.4186.

[15] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the three-body phase space of charmless $B^\pm$ decays*, Phys. Rev. **D90** (2014) 112004, arXiv:1408.5373.

[16] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^\pm \to \pi^\pm K^+ K^-$ decays*, arXiv:1905.09244, submitted to Phys. Rev. Lett.

[17] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of the $B^+ \to \pi^+\pi^+\pi^-$ decay*, arXiv:1909.05212, submitted to Phys. Rev. D.

[18] LHCb collaboration, R. Aaij *et al.*, *Observation of several sources of CP violation in $B^+ \to \pi^+\pi^+\pi^-$ decays*, arXiv:1909.05211, submitted to Phys. Rev. Lett.

[19] Wikimedia Commons, *File:standard model of elementary particles.svg — wikimedia commons, the free media repository*, https://commons.wikimedia.org/w/index.php?title=File:Standard_Model_of_Elementary_Particles.svg&oldid=341423479, 2019. [Online; accessed 6-May-2019].

[20] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014) 222002, arXiv:1404.1903.

[21] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[22] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[23] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[24] A. Salam, *Weak and electromagnetic interactions*, Conf. Proc. **C680519** (1968) 367.

[25] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[26] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[27] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[28] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[29] 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.

[30] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, arXiv:hep-ex/0107013.

[31] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.

[32] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[33] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895, arXiv:1612.07233, updated results and plots available at https://hflav.web.cern.ch.

[34] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in charm decays*, Phys. Rev. Lett. **122** (2019) 211803, arXiv:1903.08726.

[35] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[36] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[37] C. Jarlskog, *Commutator of the quark mass matrices in the Standard Electroweak Model and a measure of maximal CP violation*, Phys. Rev. Lett. **55** (1985) 1039.

[38] 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.

[39] R. H. Dalitz, *On the analysis of τ-meson data and the nature of the τ-meson*, Phil. Mag. **44** (1953) 1068.

[40] J. Back *et al.*, LAURA^{++}*: A Dalitz plot fitter*, Comput. Phys. Commun. **231** (2018) 198, arXiv:1711.09854.

[41] BaBar collaboration, B. Aubert *et al.*, *An amplitude analysis of the decay* $B^\pm \to \pi^\pm\pi^\pm\pi^\mp$, Phys. Rev. **D72** (2005) 052002, arXiv:hep-ex/0507025.

[42] CLEO collaboration, D. Bortoletto *et al.*, *A search for* $b \to u$ *transitions in exclusive hadronic* B *meson decays*, Phys. Rev. Lett. **62** (1989) 2436.

[43] ARGUS collaboration, H. Albrecht *et al.*, *Search for hadronic* $b \to u$ *decays*, Phys. Lett. **B241** (1990) 278.

[44] ARGUS collaboration, H. Albrecht *et al.*, *Search for* $b \to sX^+X^-$ *in exclusive decays of* B^+ *mesons*, Phys. Lett. **B262** (1991) 148.

[45] DELPHI collaboration, P. Abreu *et al.*, *Search for exclusive charmless B meson decays with the DELPHI detector at LEP*, Phys. Lett. **B357** (1995) 255.

[46] DELPHI collaboration, W. Adam *et al.*, *Study of rare b decays with the DELPHI detector at LEP*, Z. Phys. **C72** (1996) 207.

[47] Belle collaboration, A. Garmash *et al.*, *Dalitz analysis of the three-body charmless decays* $B^+ \to K^+\pi^+\pi^-$ *and* $B^+ \to K^+K^+K^-$, Phys. Rev. **D71** (2005) 092003, arXiv:hep-ex/0412066.

[48] BaBar collaboration, J. P. Lees *et al.*, *Study of CP violation in Dalitz-plot analyses of* $B^0 \to K^+K^-K^0_S$, $B^+ \to K^+K^-K^+$, *and* $B^+ \to K^0_S K^0_S K^+$, Phys. Rev. **D85** (2012) 112010, arXiv:1201.5897.

[49] BaBar collaboration, B. Aubert *et al.*, *Observation of the decay* $B^+ \to K^+K^-\pi^+$, Phys. Rev. Lett. **99** (2007) 221801, arXiv:0708.0376.

[50] Belle collaboration, C.-L. Hsu *et al.*, *Measurement of branching fraction and direct CP asymmetry in charmless* $B^+ \to K^+K^-\pi^+$ *decays at Belle*, Phys. Rev. **D96** (2017) 031101, arXiv:1705.02640.

[51] Belle collaboration, A. Garmash *et al.*, *Evidence for large direct CP violation in* $B^\pm \to \rho(770)^0 K^\pm$ *from analysis of the three-body charmless* $B^\pm \to K^\pm\pi^\pm\pi^\mp$ *decay*, Phys. Rev. Lett. **96** (2006) 251803, arXiv:hep-ex/0512066.

[52] BaBar collaboration, B. Aubert *et al.*, *Evidence for direct CP violation from Dalitz-plot analysis of* $B^\pm \to K^\pm\pi^\mp\pi^\pm$, Phys. Rev. **D78** (2008) 012004, arXiv:0803.4451.

[53] BaBar collaboration, B. Aubert *et al.*, *Dalitz plot analysis of $B^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}$ decays*, Phys. Rev. **D79** (2009) 072006, arXiv:0902.2051.

[54] K. G. Wilson and W. Zimmermann, *Operator product expansions and composite field operators in the general framework of quantum field theory*, Commun. Math. Phys. **24** (1972) 87.

[55] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, arXiv:hep-ph/9512380.

[56] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, Comput. Phys. Commun. **210** (2017) 103, arXiv:1601.05437.

[57] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *QCD factorization for exclusive, nonleptonic B meson decays: General arguments and the case of heavy light final states*, Nucl. Phys. **B591** (2000) 313, arXiv:hep-ph/0006124.

[58] Y.-Y. Keum, H.-n. Li, and A. I. Sanda, *Penguin enhancement and $B \to K\pi$ decays in perturbative QCD*, Phys. Rev. **D63** (2001) 054008, arXiv:hep-ph/0004173.

[59] Y.-Y. Keum and H.-n. Li, *Nonleptonic charmless B decays: Factorization versus perturbative QCD*, Phys. Rev. **D63** (2001) 074006, arXiv:hep-ph/0006001.

[60] C. W. Bauer, S. Fleming, and M. E. Luke, *Summing Sudakov logarithms in $B \to X_s\gamma$ in effective field theory*, Phys. Rev. **D63** (2000) 014006, arXiv:hep-ph/0005275.

[61] C. W. Bauer, *Hadronic B decays from SCET*, eConf **C060409** (2006) 039, arXiv:hep-ph/0606018.

[62] M. Neubert, *Aspects of QCD factorization*, AIP Conf. Proc. **602** (2001) 168, arXiv:hep-ph/0110093.

[63] C.-W. Chiang and Y.-F. Zhou, *Flavor symmetry analysis of charmless $B \to VP$ decays*, JHEP **03** (2009) 055, arXiv:0809.0841.

[64] H.-Y. Cheng, C.-W. Chiang, and A.-L. Kuo, *Updating $B \to PP, VP$ decays in the framework of flavor symmetry*, Phys. Rev. **D91** (2015) 014011, arXiv:1409.5026.

[65] H.-Y. Cheng, C.-K. Chua, and Z.-Q. Zhang, *Direct CP violation in charmless three-body decays of B mesons*, Phys. Rev. **D94** (2016) 094015, arXiv:1607.08313.

[66] W. Wang, Y.-M. Wang, D.-S. Yang, and C.-D. Lu, *Charmless two-body B (B_s^0) $\to VP$ decays in Soft-Collinear-Effective-Theory*, Phys. Rev. **D78** (2008) 034011, arXiv:0801.3123.

[67] H.-n. Li and S. Mishima, *Penguin-dominated $B \to PV$ decays in NLO perturbative QCD*, Phys. Rev. **D74** (2006) 094020, arXiv:hep-ph/0608277.

[68] Y. Li, W.-F. Wang, A.-J. Ma, and Z.-J. Xiao, *Quasi-two-body decays $B_{(s)} \to K^*(892)h \to K\pi h$ in perturbative QCD approach*, Eur. Phys. J. **C79** (2019) 37, arXiv:1809.09816.

[69] Y. Li, A.-J. Ma, W.-F. Wang, and Z.-J. Xiao, *Quasi-two-body decays $B_{(s)} \to P\rho \to P\pi\pi$ in perturbative QCD approach*, Phys. Rev. **D95** (2017) 056008, arXiv:1612.05934.

[70] H.-Y. Cheng and K.-C. Yang, *Charmless hadronic B decays into a tensor meson*, Phys. Rev. **D83** (2011) 034001, arXiv:1010.3309.

[71] Y. Li *et al.*, *Quasi-two-body decays $B_{(s)} \to Pf_2(1270) \to P\pi\pi$ in the perturbative QCD approach*, Phys. Rev. **D98** (2018) 056019, arXiv:1807.02641.

[72] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.

[73] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[74] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[75] E. Mobs, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, https://cds.cern.ch/record/2197559, 2016.

[76] LHCb collaboration, *LHCb: letter of intent*, CERN-LHCC-95-5. LHCC-I-8, 1995.

[77] M. Pepe Altarelli and F. Teubert, *B physics at LHCb*, Int. J. Mod. Phys. **A23** (2008) 5117, arXiv:0802.1901.

[78] LHCb collaboration, R. Aaij *et al.*, *Measurement of $Z \to \tau^+\tau^-$ production in proton-proton collisions at $\sqrt{s}$ =8 TeV*, JHEP **09** (2018) 159, arXiv:1806.05008.

[79] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward top pair production in the dilepton channel in pp collisions at $\sqrt{s} = 13\,\text{TeV}$*, JHEP **08** (2018) 174, arXiv:1803.05188.

[80] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s}$ =13* TeV, JHEP **10** (2015) 172, Erratum ibid. **05** (2017) 063, arXiv:1509.00771.

[81] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\psi(2S)$ production cross-sections in proton-proton collisions at $\sqrt{s}$ =7 and 13* TeV, arXiv:1908.03099, submitted to EPJC.

[82] LHCb collaboration, R. Aaij *et al.*, *Search for dark photons produced in 13* TeV *pp collisions*, Phys. Rev. Lett. **120** (2018) 061801, arXiv:1710.02867.

[83] LHCb collaboration, R. Aaij *et al.*, *Search for Higgs-like boson decaying into pair of long-lived particles*, Eur. Phys. J. **C76** (2016) 664, arXiv:1609.03124.

[84] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-flavour-violating decays of Higgs-like bosons*, Eur. Phys. J. **C78** (2018) 1008, arXiv:1808.07135.

[85] *LHCb material for presentations*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html. Accessed: 2019-01-10.

[86] C. Abellan Beteta *et al.*, *Monitoring radiation damage in the LHCb Tracker Turicensis*, arXiv:1809.05063.

[87] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[88] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, 2003.

[89] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, arXiv:1812.10790.

[90] 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.

[91] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0_s \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017) 191801, arXiv:1703.05747.

[92] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.

[93] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[94] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[95] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031, 2003.

[96] T. Head, *The LHCb trigger system*, JINST **9** (2014) C09015.

[97] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[98] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.

[99] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb-PUB-2011-016, 2011.

[100] R. Aaij *et al.*, *Tesla: an application for real-time data analysis in High Energy Physics*, Comput. Phys. Commun. **208** (2016) 35, arXiv:1604.05596.

[101] LHCb collaboration, R. Aaij *et al.*, *Search for the suppressed decays $B^+ \to K^+K^+\pi^-$ and $B^+ \to \pi^+\pi^+K^-$*, Phys. Lett. **B765** (2017) 307, arXiv:1608.01478.

[102] G. D. McGregor, *B counting at BaBar*, Master's thesis, University of Canterbury, 2008. SLAC-R-912.

[103] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[104] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.

[105] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, Nucl. Instrum. Meth. **A559** (2006) 190.

[106] D. Casadei, *Estimating the selection efficiency*, JINST **7** (2012) P08021, arXiv:0908.0130.

[107] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[108] A. Martin Sanchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, LHCb-PUB-2011-026, 2012.

[109] 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.

[110] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.

[111] LHCb collaboration, R. Aaij *et al.*, *First measurement of the CP-violating phase $\phi_s^{d\bar{d}}$ in $B_s^0 \to (K^+\pi^-)(K^-\pi^+)$ decays*, JHEP **03** (2018) 140, arXiv:1712.08683.

[112] BESIII collaboration, M. Ablikim *et al.*, *Precision study of $\eta' \to \gamma\pi^+\pi^-$ decay dynamics*, Phys. Rev. Lett. **120** (2018) 242003, arXiv:1712.01525.

[113] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[114] M. Whitehead, *Observation of the decay $B^0 \to D^0K^+K^-$ with the LHCb detector at CERN*, PhD thesis, University of Warwick, 2012, CERN-THESIS-2012-130.

[115] LHCb collaboration, *Physics case for an LHCb Upgrade II — Opportunities in flavour physics, and beyond, in the HL-LHC era*, arXiv:1808.08865.

[116] M. Gronau, D. Pirjol, A. Soni, and J. Zupan, *Improved method for CKM constraints in charmless three-body B and B_s^0 decays*, Phys. Rev. **D75** (2007) 014002, arXiv:hep-ph/0608243.

[117] A. E. Snyder and H. R. Quinn, *Measuring CP asymmetry in $B \to \rho\pi$ decays without ambiguities*, Phys. Rev. **D48** (1993) 2139.