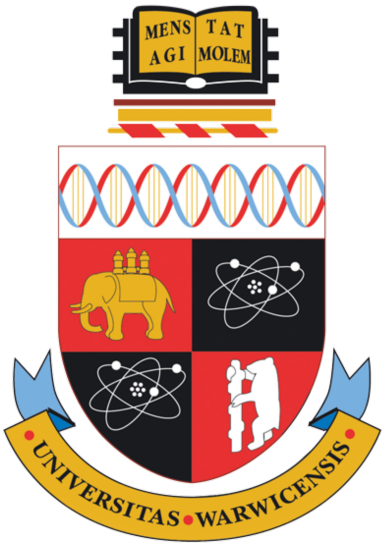

Search for *CP* violation in charmless three-body decays of strange-beauty baryons

by

Abhijit Mathad

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

November 2018

Contents

“ *Sophie . . . a philosopher is he who knows that in reality he knows very little. He knows nothing about life and about the world. And now comes the important part: it troubles him that he knows so little.* ”

Jostein Gaarder, *Sophie's World*

Acknowledgments

I would like to thank my supervisor Prof Tim Gershon, for giving me an opportunity to work in such an exciting field of physics. I am thankful to him for his excellent guidance on the research and analyses expounded in this thesis and for providing positive criticism on my academic writing. He has always found time to entertain my questions and ideas. I am truly fortunate to have worked under his guidance in conducting my doctoral research.

I am very grateful to my second supervisor Dr Thomas Latham, who had solutions to almost every problem, ranging from physics to programming. His infinite patience and support have helped me greatly during my time at Warwick. I would as well like to thank his wife Eugenia and daughter Angelica for welcoming me to the amazing barbecues hosted at their house on many occasions.

I thank Dr Anton Poluektov for helping me with my analysis, particularly in relation to resampling of global particle identification variables and the development of the 'Tensorflow-Analysis' package, without which the work presented in the later part of the thesis would not have been possible.

I would like to thank Dr Tom Blake for answering any questions I had, however small, and Dr Michal Kreps for his advice on writing of the decay files.

I would like to thank Andrew and David, my fellow PhD students at Warwick, for sharing the various problems we had encountered in our respective analyses; Cayo and Ed for distracting me with questions from which I always ended up learning something new; Dan O'Hanlon for introducing me to cool python packages, accompanied by discussions ranging from machine learning to brains in jars; Charlotte for being the voice of reason when the conversations were led astray by Mark; Dan Craik for teaching me useful bash commands; Rafael for introducing me to

‘V0hhfitter’ fitting package with which I have had a bittersweet relationship; Mark Whitehead for his help and advice in various topics related to physics, jobs and questioning my prejudice about lime/lemon; and finally thanks to Leigh Whitehead for interesting discussions in R1 which made my stay enjoyable at CERN.

I would like to thank the whole of the LHCb collaboration for creating such a great atmosphere to work. I feel privileged to be part of this collaboration.

I would like to thank my father Dr R D Mathad, who first aroused my interest in this fundamental field. My interaction with him has been a truly enjoyable and elevating exercise in physics. I would like to thank my mother Dr Pratima Mathad, who not only pestered me for my good health, but also helped me in making judicious decisions in life. Additionally, I would like to thank my late Ajja and Ajji, my paternal grand parents, Ajja for all the blessings and Ajji with whom I enjoyed many science and wild-life documentaries; Avva and Appaji, my maternal grand parents for spoiling me with delicious food; my aunts, uncles, cousins and friends who have made my summer vacations with them the most memorable ones. I offer my obeisances to Sri Swamiji of Duradundeshwar Matt, whereto I belong, for all the blessings.

I would like to thank April Cridland for her unwavering support and patience in all of my endeavours; I would also like to thank her parents and brothers as well for hosting me at their place during the holidays.

Last but not the least, I thank immensely the University of Warwick for offering me the prestigious ‘Chancellor’s International Scholarship’ award to undertake the present research ventures.

Abhijit Mathad

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.

Abhijit Mathad

November 12, 2018

Abstract

Studies of charmless three-body decays of the Ξ_b^- and Ω_b^- baryons to $ph^-h'^-$ final states, where $h^{(\prime)}$ denotes a kaon or pion, are presented. The first part of the analysis is based on a sample of proton-proton collision data collected by the LHCb experiment at centre-of-mass energies $\sqrt{s} = 7$ & $8\,\text{TeV}$, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. In this analysis, the decay $\Xi_b^- \to pK^-K^-$ is observed with a significance of 8.7 standard deviations, and evidence at the level of 3.4 standard deviations is found for the $\Xi_b^- \to pK^-\pi^-$ decay. The results are reported, relative to the $B^- \to K^+K^-K^-$ normalisation channel, for the products of branching fractions and b-hadron production fractions. The branching fractions of $\Xi_b^- \to pK^-\pi^-$ and $\Xi_b^- \to p\pi^-\pi^-$ relative to $\Xi_b^- \to pK^-K^-$ decays are also measured. In the second part of the analysis, an amplitude analysis of $\Xi_b^- \to pK^-K^-$ decays is presented. This analysis is based on a data sample corresponding to integrated luminosities of $1\,\text{fb}^{-1}$, $2\,\text{fb}^{-1}$, $0.3\,\text{fb}^{-1}$ and $1.7\,\text{fb}^{-1}$ collected by LHCb during 2011 ($\sqrt{s} = 7\,\text{TeV}$), 2012 ($\sqrt{s} = 8\,\text{TeV}$), 2015 ($\sqrt{s} = 13\,\text{TeV}$) and 2016 ($\sqrt{s} = 13\,\text{TeV}$) respectively. Sensitivity to the CP violation parameters in $\Xi_b^- \to pK^-K^-$ decays is also presented. An updated search for the previously unobserved charmless decay $\Omega_b^- \to pK^-K^-$ is also conducted.

CHAPTER 1

Introduction

> *"The time has come," the Walrus said, "to talk of many things: of shoes – and ships – and sealing-wax – of cabbages – and kings – And why the sea is boiling hot – And whether pigs have wings.*
>
> Lewis Carroll, *Through the Looking Glass*

The theory of Big Bang is a cosmological model that postulates the birth and evolution of the universe. This theory is supported by many experimental observations, among which are discoveries of the expansion of the universe in 1929 [1] and of the Cosmic Microwave Background (CMB) in 1964 [2]. The Big Bang theory though predicts production of an equal amount of matter and antimatter at the birth of the universe, yet the present day universe is dominated by matter. Therefore during the evolution of the universe, there must have prevailed certain conditions and interactions between the fundamental particles of matter and antimatter that led to the asymmetry that is observed today. In fact, three conditions are required to be met by processes during the evolution of the universe to produce this imbalance, namely baryon/lepton number violation, charge and Charge-Parity (CP) violation and a departure from thermal equilibrium [3]. The highly successful theory, the Standard Model (SM) of particle physics, that explains the properties of fundamental particles and their interactions allows CP violation in the quark sector but its effect is many orders of magnitude too small to account for the observed matter-antimatter asymmetry. This motivates the search for new particles and new sources of interactions between fundamental particles that could potentially give rise to the observed asymmetry. The charmless decays of b-hadrons provide a fertile group to conduct studies to search for such new interactions.

The present thesis deals with the study of charmless decays of $\Xi_b^-(\Omega_b^-) \to phh'^-$, where h and h' are a kaon or a pion, to search for sources of CP asymmetries using the data collected by the LHCb experiment at CERN. The thesis is organised as follows.

In chapter 2, I introduce the Standard Model of particle physics and highlight its successes and failures. I discuss in this chapter the matter-antimatter asymmetry problem and the source of CP violation within the SM. Since this thesis deals with CP violation in the quark sector of the SM, I discuss the different manifestations of CP violation that can occur in the decays of b-hadrons. I also highlight here the reasons as to why charmless decays of b-baryons act as a fertile ground to search for new sources of CP violation beyond the Standard Model.

The LHCb produces b-baryons in unprecedented quantities and has hence opened up a new field in flavour physics for precision measurement. In chapter 3, I give an overview of the CERN accelerator complex and discuss the different sub-detectors that form the LHCb experiment. The performance of the LHCb experiment is also discussed in this chapter.

In chapter 4, I present the analysis of my search for $\Xi_b^-(\Omega_b^-) \to phh'^-$ decays based on a sample of proton-proton collision data collected by LHCb experiment at centre-of-mass energies $\sqrt{s} = 7$ & 8 TeV, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. Here I present details on the event selection used to create the final data samples of $\Xi_b^-(\Omega_b^-) \to phh'^-$ events, the background studies that were performed to understand the origins of selected events, the fitting strategy used to extract the signal yields and the measurement of the branching fraction ratios.

In chapter 5, I present a search for CP violation in $\Xi_b^- \to pK^-K^-$ decays using the data sample collected by the LHCb experiment between the years 2011–2016. The various techniques employed to achieve a gain in the signal selection efficiency are discussed. An amplitude analysis of $\Xi_b^- \to pK^-K^-$ decays has been conducted to establish a baseline model that will be used to measure the CP violation parameters. The amplitude analysis is conducted using the 'TensorFlowAnalaysis' package. This analysis is currently under review within the LHCb collaboration and is blind to the CP violation parameters. The unblinding will occur when the review committee appointed by the LHCb collaboration is satisfied with the procedure employed in this analysis to search for CP violation. Therefore I have presented in this chapter the sensitivity that can be achieved on the CP violation parameters with the current and future LHCb data samples. I also discuss the possible sources and evaluation of the systematic uncertainties as part of future work.

In the final chapter 6, I summarise the thesis. During my PhD, I assisted with the implementation of the material description for the upgrade of the VErtex LOcator (VELO) sub-detector to be installed during the LHC long shutdown 2 (2019-20). In Appendix A, the work undertaken is highlighted. In particular, I describe here the detector geometry of the VELO module and its support system.

CHAPTER 2

Theoretical background

> *I cannot deny a feeling of unreality in writing about the first three minutes as if we really know what we're talking about.*

Steven Weinberg, *The First Three Minutes*

In this chapter, I will discuss the theoretical background and motivation for the work presented in this thesis. The Standard Model of particle physics is briefly discussed in Sec. 2.1 and the problem of matter-antimatter asymmetry is discussed in detail in Sec. 2.2.

2.1 The Standard Model of particle physics

The Standard Model (SM) of particle physics is an extremely successful relativistic quantum field theory that describes the fundamental interactions and the dynamics of the elementary particles of nature.

All observed particles can be classified as states that carry half integer spin, called fermions, or those that carry integer spin, which are called bosons. These states are differentiated according to how they transform under Lorentz transformations or more accurately how they form the different representations of the Poincaré group [4, 5]. The Poincaré group, $SO^+(1,3) \otimes \mathbb{R}(1,3)$, is a 10 dimensional non-Abelian Lie group (3 rotations, 3 boosts and 4 translations) corresponding to the symmetries of special relativity. The group $SO^+(1,3)$ represents the proper orthochronous Lorentz group which is a 6 dimensional non-Abelian group of all Lorentz transformations of flat Minkowski spacetime that preserve both orientation and direction of the

time axis. The group $\mathbb{R}(1,3)$ is the Abelian Lie group of all spacetime translations. It was shown by Wigner [6] that different representations of the Poincaré group are labelled by the eigenvalues of the two Casimir operators of the group. One such label is the four momentum (p^μ) of the state and the other is either spin (j) or helicity ($h = \vec{J}.\hat{p}$) depending on whether the particle is massive or massless. Here $\vec{J}$ is the spin vector of the particle and $\hat{p}$ is the unit vector along the momentum of the particle.

Massive particles have their representations labelled by spin (j) and the total number of degrees of freedom for a given representation is $2j + 1$, corresponding to $-j \ldots j$. For example, a massive spin-1 particle (Z, $W^\pm$) forms a vector representation of the Poincaré group, which would have three degrees of freedom (2 transverse components corresponding to $j = \pm 1$ and 1 longitudinal component corresponding to $j = 0$).

Massless particles have their representations labelled by helicity (h) and each representation of this group has only one degree of freedom corresponding to that value of helicity. Therefore, one would be justified in thinking of massless particles of each different helicity as different species of particle. However particles of opposite helicity are related to each other by the symmetry of space inversion (Parity). Therefore, if the massless particle interacts via parity-conserving interactions then it is necessary to assign the two representations h and $-h$ to represent the same particle. For example, in an ultra-relativistic limit, a spin-half particle (Ξ_b^- or Ω_b^-) and its conjugates, produced via parity-conserving interactions (*i.e.* strong interaction), have two degrees of freedom (2 transverse corresponding to $h = \pm 1$). This is essentially because massless spin-1 particles that mediate such interactions (*i.e.* gluons) also only have 2 degrees of freedom. On the other hand, a massless spin-half particle that undergoes only parity-violating interactions (*i.e.* weak interactions) will have only 1 degree of freedom corresponding to $h = -1$ (*e.g.* left-handed neutrinos) and it's antiparticle will have $h = +1$ (*e.g.* right handed antineutrinos). Therefore, when we refer to the spin states of a massive or the massless particle we refer to the different representations of the Poincaré group. Table 2.1 shows the various fermionic and bosonic fields that form different representations of the Poincaré group.

In the SM, the fundamental fermions, the quarks and leptons, make up the matter around us, *i.e.* atoms. The vector bosons however are the carriers of three of the four fundamental forces. The SM describes the strong, weak and electromagnetic interactions in terms of a non-Abelian gauge field theory which is renormalisable and remains invariant under local transformation of the gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_{Y_W}$ [7]. It does not, however, incorporate the geometric theory of general relativity that successfully describes the gravitational force.

The group $SU(3)_C$ represents the symmetry group of the theory of quantum chromodynamics (QCD) in which strong interactions are mediated by 8 spin-1 massless gauge bosons called gluons that carry the conserved colour quantum number C. These 8 gluon vector fields

Table 2.1: Various fermionic and bosonic fields that form different representations of the Poincaré group.

Representation	Label	Total spin	Associated quantum field
Singlet	$(0,0)$	0	Scalar field *e.g.* Higgs
Fundamental spinor	$(\frac{1}{2},0)$	$\frac{1}{2}$	Left-handed Weyl fermions
Conjugate spinor	$(0,\frac{1}{2})$	$\frac{1}{2}$	Right-handed Weyl fermions
Dirac or bispinor	$(\frac{1}{2},0)\oplus(0,\frac{1}{2})$	$\frac{1}{2}$	Dirac fermions
Vector	$(\frac{1}{2},\frac{1}{2})$	1	Vector bosons *e.g.* $\gamma, W^{\pm}$, *etc.*
Rarita-Schwinger	$(1,\frac{1}{2})$	$\frac{3}{2}$	Rarita-Schwinger fermions
Spin-2	$(1,1)$	2	Spin-2 boson *e.g.* graviton

transform in the adjoint representation of the $SU(3)_C$ group. The fundamental fermions, quarks and antiquarks, interact via strong interactions transforming as triplets corresponding to the three colour charges 'red', 'green' and 'blue', in the fundamental and complex conjugate representation of the $SU(3)_C$ group. These three colour charges are labelled by strong isospin (T^c) and its 3rd component (T^c_3) [8].

The group $SU(2)_L \otimes U(1)_{Y_W}$ represents the symmetry group of the 'chiral' electroweak (EW) theory, which predicts 4 massless gauge bosons (W^1, W^2, W^3, B) that are eigenstates of weak isospin (T), its 3rd component (T_3), and weak hypercharge Y_W [9–11]. The EW theory is a 'chiral' theory because the left-handed (L) fermions and their corresponding charge conjugates transform as doublets in the fundamental and conjugate representation respectively. The right-handed (R) fermion fields, on the other hand, are singlets and hence the charged gauge bosons of EW theory do not couple to them. Again, as for the $SU(3)_C$ group, the massless vector gauge bosons for the EW theory transform in the adjoint representation of the $SU(2)_L \otimes U(1)_{Y_W}$ symmetry group.

In nature, however, the observed weak interactions such as β decay, $n \to pe\overline{\nu}_e$, are short range interactions and therefore force mediating gauge bosons must be massive [12, 13]. Moreover, it is only the linear combination of Y_W and T_3, *i.e.* $Q = T_3 + Y_W$,[1] that is observed to be the conserved quantum number. Therefore it is necessary to find a way to break $SU(2)_L \otimes U(1)_{Y_W} \to U(1)_{EM}$, where Q is the quantum number associated to the $U(1)_{EM}$ gauge group, and preserve the renormalisability of the theory. This can be achieved through the 'Higgs mechanism' where a complex Higgs scalar field, an $SU(2)_L$ doublet $\begin{pmatrix} H^{\pm} \\ H^0 \end{pmatrix}$, is introduced into the Lagrangian. After acquiring a non-zero Vacuum Expectation Value (VEV), the Higgs

[1] The factor $\frac{1}{2}$ that is usually found in front of Y_W in the literature has been absorbed into Y_W.

field spontaneously breaks $SU(2)_L \otimes U(1)_Y$ down to $U(1)_{EM}$ [14, 15]. As a consequence all fermions of the SM acquire mass terms and the EW bosons mix to give rise to the massive spin-1 gauge bosons of the weak interaction ($W^{\pm}$, Z) and the massless spin-1 gauge boson of electromagnetism (photon). This is depicted in Figure 2.1, which also shows the mass and EW quantum numbers associated with the SM particles. In this figure, two scenarios are shown, the top 'Unbroken symmetry' which corresponds to the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_{Y_W}$ and bottom 'Broken symmetry' which corresponds to the group $SU(3)_C \otimes U(1)_{EM}$ where the Higgs boson has acquired a VEV. As a result of the broken symmetry, most of the fermions have acquired mass. This breaking of the symmetry also leads to mixing of EW gauge bosons (W^1, W^2, W^3, B), as depicted in Figure 2.1, to give rise to the massive $W^{\pm}$, Z, and a massless photon (γ). This mixing is characterised by the mixing angle (θ_{W}).

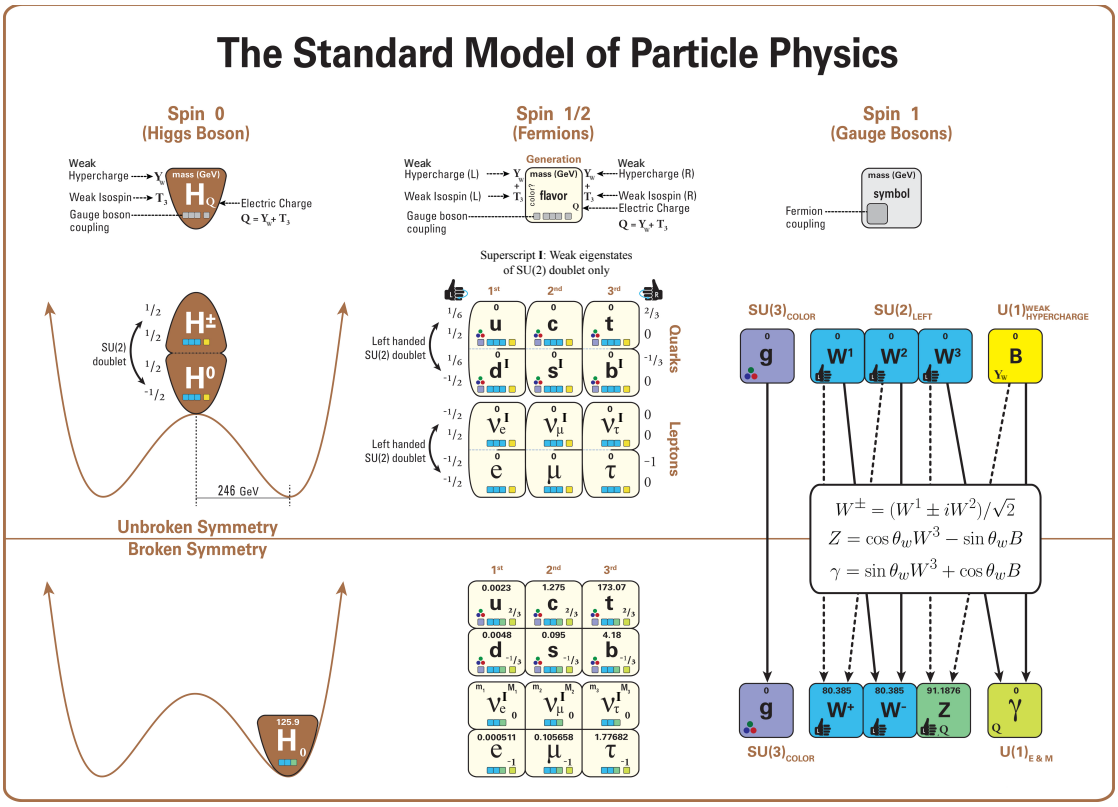

Figure 2.1: A schematic showing two scenarios, the top 'Unbroken symmetry' corresponds to $SU(3)_C \otimes SU(2)_L \otimes U(1)_{Y_W}$ and bottom 'Broken symmetry' corresponds to $SU(3)_C \otimes U(1)_{EM}$ [16].

The SM Lagrangian is composed of the following terms

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{gauge}} + \mathcal{L}_{\mathrm{fermion}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}} + \mathcal{L}_{\mathrm{GF}} + \mathcal{L}_{\mathrm{Ghost}} \,. \tag{2.1}$$

I discuss here, very superficially, the properties of the terms whose expansion can be found in Ref. [17]. The Lagrangian $\mathcal{L}_{\mathrm{gauge}}$ and $\mathcal{L}_{\mathrm{fermion}}$ contain kinetic terms for the gauge fields and fermions respectively. The $\mathcal{L}_{\mathrm{Higgs}}$ Lagrangian contains the kinetic term for the Higgs field. This term is also responsible for generating the mass terms and mixing of EW gauge bosons. The Yukawa term $\mathcal{L}_{\mathrm{Yukawa}}$ gives rise to the mass terms for fermions (except for neutrinos) and is the source of CP violating effects that occur in the SM. I will investigate the Yukawa term in more detail later. The gauge fixing term $\mathcal{L}_{\mathrm{GF}}$ is necessary to remove any extra degrees of freedom that might arise in the theory. The term $\mathcal{L}_{\mathrm{Ghost}}$ arises due to the non-Abelian nature of the SM and gives rise to unphysical Fadeev-Popov ghosts that appear in loop calculations. These unphysical ghost terms are needed in quantum field theories to yield unambiguous and non-singular solutions.

Within the SM, although the Higgs mechanism generates mass for the fundamental particles of nature, it is not the only mechanism through which masses for the observed particle states are generated. At low energies ($\Lambda_{\mathrm{QCD}} \sim 200\,\mathrm{MeV}$), non-perturbative effects dominate QCD processes [18]. As a result, a phenomenon called colour confinement [19] essentially leads to hadronisation of quarks[2] and formation of colourless hadron bound states. It is these hadron states that freely exist and can either be observed through their decay products or exist as stable particles (such as protons). No free particle with fractional charge has ever been observed. The hadrons can be classified as mesons (a quark and an antiquark) and baryons (three quarks or three antiquarks). There have also been observations of tetraquark, $Z(4430)^+$ [20], and pentaquark states $P_c(4380)^+$ and $P_c(4450)^+$ [21]. To see the effect of confinement on the mass of hadrons, compare the mass of the proton ($m_p \sim 938\,\mathrm{MeV}$), which is made of uud, to the rest mass of u ($m_u \sim 2.3\,\mathrm{MeV}$) and of d ($m_d \sim 4.8\,\mathrm{MeV}$) which only sum to $9.4\,\mathrm{MeV}$. The source of the bulk of the proton's mass is the QCD binding energy, which arises out of QCD chiral symmetry breaking [22, 23]. Therefore hadrons can be viewed as being made of valence quarks and a "sea" of gluons and quark-antiquark pairs. Throughout this thesis when the quark content is spoken of, it can be assumed that it refers to the valence quark content.

The SM Lagrangian also exhibits global continuous symmetries. As a consequence, there are a set of associated conservation laws as stated by Ward-Takahashi identities [24, 25], the quantum equivalent of classical Noether's theorem [26]. The symmetries exhibited by the SM and the associated conservation laws are shown in Table 2.2.

[2]Except for the top quark which decays before hadronising.

Table 2.2: Symmetries of the SM of particle physics.

Symmetry	Group	Symmetry Type	Conserved Quantities
Poincaré	$SO^+(1,3) \otimes \mathbb{R}(1,3)$	Continuous and Global	Four momentum, Angular momentum
Gauge	$SU(3)_C \otimes SU(2)_L \otimes U(1)_{Y_W}$	Local, Continuous and exact	Colour charge, Electric charge
Global phases	$U(1)_B \otimes U(1)_e \otimes U(1)_\mu \otimes U(1)_\tau$	Accidental, exact, continuous and global	Baryon, electron, muon and tau number
Flavour	$SU(3)_{\text{flavour}}$	Approximate, continuous and global	Isospin (I, I_3), charmness, strangeness, topness, bottomness
Custodial	$SU(2)_L \otimes SU(2)_R$	Approximate, continuous and global	Symmetry of Higgs potential
CPT	Pin group	Discrete	Equal masses and lifetimes for particle and its conjugate

One can see here that the SM exhibits an approximate flavour symmetry (it will be explained later as to why this is an approximate symmetry). Due to this approximate symmetry, one can design a scheme for classifying hadrons in terms of their valence quarks and antiquarks. Such a scheme was put forth by Murray Gell-Mann and George Zweig and is called the quark model [27, 28]. In the quark model, one can use the quantum numbers of the valence quarks to determine the quantum numbers of the hadrons. The conserved quantum numbers of the quarks such as baryon number (B) and flavour quantum numbers like charm (C'), strangeness (S), top (T), bottomness (B') can be used to build a hypercharge quantum number $Y = \frac{1}{2}(B + C' + S + T + B')$. This number together with the 3rd component of the flavour isospin, I_3,[3] label the hadron states. One can classify *e.g.* the beauty baryons according to this scheme using bottom and strangeness quantum numbers as depicted in Figure 2.2. However this flavour symmetry is an approximate symmetry of the SM as it assumes that $m_u = m_d = m_s$, where subscript denotes the quark flavour. The assumption is broken in the electroweak sector by the Higgs mechanism which couples to different flavours of quarks with different Yukawa coupling. As a consequence, unlike the strong and electromagnetic interactions, the weak interactions do not conserve the flavour quantum numbers.

[3] Since u and d form isospin doublets. I_3 by definition encodes the up and down quantum numbers.

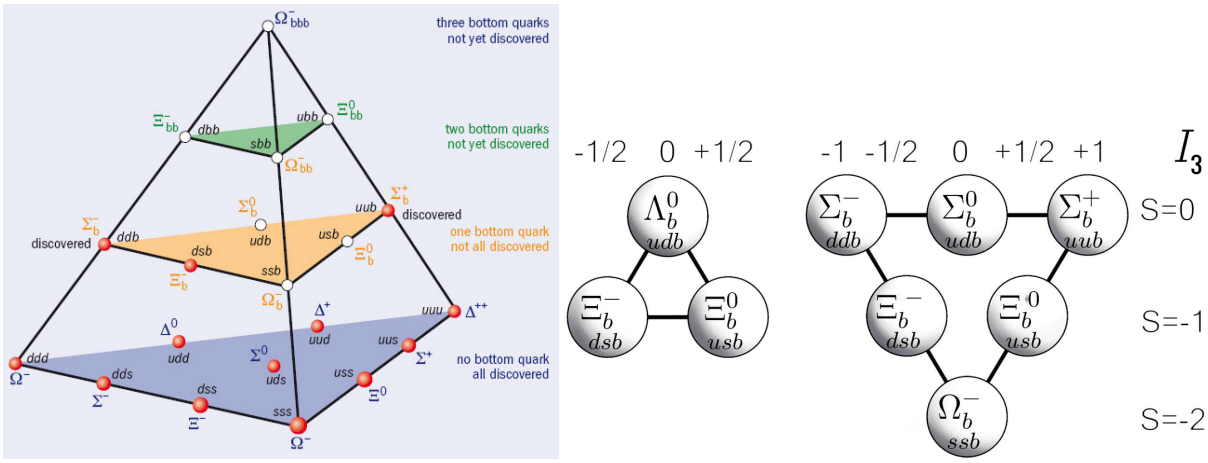

Figure 2.2: For $J^P = \frac{1}{2}^+$ baryons, I show (left) the quark model classifying them according to the bottom (B') quantum number and (right) the classification of $B' = 1$ baryons according to strangeness and isospin I_3 quantum numbers. Note here that the Σ_b^0 particle predicted by the quark model has not yet been discovered [29, 30].

As the SM is a quantum field theory it is also invariant under the combined discrete CPT transformation due to the CPT theorem [31]. Here, C is the charge conjugate unitary operator that transforms a particle state to its charge conjugate without changing momentum or spin and therefore negates all internal quantum numbers of the field, and P is also a unitary operator called parity which reverses the chirality or helicity of a given state, therefore it transforms a left-handed Weyl spinor to a right-handed Weyl spinor and vice versa. Only the strong and the electromagnetic interactions conserve C and P symmetry. T is an antiunitary operator that exchanges the initial and final states. This T operator reverses momentum and spin and also flips the sign of the time component of a state. There has been a direct observation of T violation in weak interactions of the K^0 [32] and B^0 systems [33]. One can also form a combination CP which is the subject of the thesis and will be discussed in the next section. Experimental probes to study the violation of CPT symmetry have also been conducted by many experiments such as BaBar [34, 35], LHCb [36], ALPHA [37, 38], *etc.* However the results so far have been consistent with CPT symmetry.

Despite the huge success of SM there are a few theoretical problems and experimental observations that it fails to address [39]. Below are some of the problems that the SM faces.

- Neutrino oscillations: Neutrinos have been observed to undergo oscillations from one flavour to another, hence violating individual lepton flavour numbers. Such oscillations can only occur if neutrinos are not massless, which they are in the SM. It is therefore necessary

to extend the SM to accommodate non-zero neutrino masses. Searches for neutrinoless double β decays are also being carried out aiming to establish if massive neutrinos are Dirac (with particles and antiparticles being different objects: $\nu \neq \overline{\nu}$) or Majorana (with particles and antiparticles being the same: $\nu = \overline{\nu}$) fermions.

- Matter-antimatter asymmetry: This is the primary motivation for the work in this thesis and will be discussed in detail in the following section.
- Strong CP problem: This will be discussed when covering CP violation in the SM.
- Why are the quark flavor parameters in the SM small and hierarchical? and Why are they so different from the neutrino flavor parameters?
- Why are there 3 generations of fermions?
- Hierarchy Problem: Why is the mass of the SM Higgs so small, since it has a quadratic divergence when quantum corrections are included ?
- Quantum gravity: As mentioned previously, the SM does not account for gravity. When Poincaré symmetry is broken, particles lose definite values for mass and spin. Gravity bends spacetime, changing the Minkowski metric. Thus we expect breaking of the Poincaré symmetry when gravity becomes significant.
- Number of parameters: The SM is not expressed in terms of the fundamental constants of the Universe such as speed of light (c) or Planck's constant (h), however it consists of 19 free parameters which need to be determined from experiments.
- The SM does not explain the constituents of dark matter or dark energy which constitute 23% and 73% of the Universe respectively [40].
- The SM does not achieve a true unification of the couplings of the weak, EM and strong interactions.

2.2 Matter-antimatter asymmetry

The matter-antimatter asymmetry problem refers to the fact that there is an observed imbalance between matter (stable baryons and leptons) and antimatter (stable antibaryons and antileptons) in the observable universe. The stable baryons are protons and neutrons[4] and stable leptons are

[4]Neutrons can be stable or unstable depending on the nuclide they are in. A free neutron is unstable and decays via β-decay

electrons and the lightest neutrino. The complex conjugate partners of these particles make-up the stable antibaryons and antileptons. The matter and antimatter in the early universe almost entirely annihilated, leaving only a small amount of matter that forms the currently observable universe. The possibility that there are regions of space dominated by primordial antimatter can be excluded by the lack of signal in astronomical searches for photons from annihilation processes that occur at the interface between matter and antimatter dominated regions of the universe. This predominance of matter is believed to have arisen in the early evolution of the universe.

A quantitative measure of the departure from baryon antibaryon equality is provided by a parameter called the baryon asymmetry of the universe ($\eta_{\rm BAU}$) [41]. This quantity is defined as the difference in the total number density of baryons (n_B) and antibaryons ($n_{\overline{B}}$) divided by the number density of cosmic background radiation photons (n_γ) or more accurately the entropy density of the universe (s).[5] Since there is no evidence of antibaryons in the present day universe, we can write

$$\eta_{\rm BAU} = \frac{n_B - n_{\overline{B}}}{s} = \frac{n_B}{s} \tag{2.2}$$

The $\eta_{\rm BAU}$ parameter can be inferred from the abundances of light elements in the intergalactic medium formed in the process of Big Bang nucleosynthesis and from the power spectrum of temperature fluctuations in the Cosmic Microwave Background [43]. It is measured to be

$$\eta_{\rm BAU} = (6.07 \pm 0.33) \times 10^{-10}\,. \tag{2.3}$$

To see how significant this value is, one needs to calculate $\frac{n_B}{n_\gamma}$ assuming a baryosymmetric universe *i.e.* $\eta_{\rm BAU} = 0$. In this case, as the universe cools from temperatures $T > 1\,{\rm GeV}$, the baryons and antibaryons are in thermal equilibrium with the relatively high-energy photons through processes such as

$$p + \overline{p} \rightleftharpoons \gamma + \gamma \tag{2.4}$$

Once the rate for this process becomes slower than the expansion rate (*i.e.* Hubble's constant H), the probability of a subsequent annihilation becomes negligible. Using a typical hadronic cross-section, this equality of rates occurs at a temperature of about $20\ {\rm MeV}$. Calculations of this

[5] As the universe expands, n_γ is not a conserved quantity whereas s is conserved. Both s and n_γ are related by the expression $s = 1.8 g_s(T) n_\gamma$, where $g_s(T)$ is the number of degrees of freedom contributing to the entropy at temperature T of the universe [42].

thermal freeze out of the baryons predict [41, 44]

$$\frac{n_B}{n_\gamma} \approx \frac{n_{\overline{B}}}{n_\gamma} \approx 10^{-18}\,. \tag{2.5}$$

It can be seen here that the observed value in (2.3) is significantly larger than this value.

However there is a simple path to obtain a much larger value. To generate the observed $\eta_{\rm BAU}$, in the early universe one requires for every $\sim 10^9$ antibaryons there be $\sim 10^9 + 1$ baryons. But where would such an excess come from? One might think it could appear as an initial condition, set at the beginning of the universe. However, such an initial condition is irrelevant due to inflation and moreover would violate any naturalness principle [41, 44]. What is preferable is a mechanism by which this peculiar excess arises dynamically during the evolution of the universe, a possibility known as baryogenesis.

One can in principle apply the same arguments to leptogenesis (dealing with lepton antilepton asymmetry) by defining an equivalent parameter called lepton asymmetry of the universe (η_{LAU}) [45, 46] given by

$$\eta_{LAU} = \frac{n_L - n_{\overline{L}}}{n_\gamma}, \tag{2.6}$$

where n_L and $n_{\overline{L}}$ are the number density of leptons and antileptons respectively. This parameter is however extremely hard to measure since it requires the measurement of neutrino-antineutrino asymmetry in a Universe with overall charge neutrality [46].

To explain the observed matter-antimatter asymmetry in the Universe, in 1967 Andrei Sakharov [3, 47] proposed the three following conditions that must be satisfied by a fundamental theory of particle physics.

Baryon number violation: There exists at least one baryon or lepton number violating process. The existence of this kind of interaction is a minimum starting point for the observed excess matter in the current Universe.

In the SM, the conservation of baryon number (B) and lepton number (L) are a consequence of the accidental but exact symmetries exhibited by the SM Lagrangian and as a result it is not possible to violate B and L at tree-level or at any order of perturbation theory. However, in many cases the perturbative expansion does not describe all the dynamics of the theory. In 1980, Gerard 't Hooft [48] realised that the electroweak theory of SM can exhibit non-perturbative effects at high energies and can result in instantons called sphalerons through the quantum Adler-Bell-Jackiw anomaly [49] (or chiral anomaly), which may give rise to processes that violate the combination $B + L$, but not the combination $B - L$. Such processes are heavily-suppressed

at low energies and indeed none have yet been experimentally observed. However, when the temperature of the universe was greater or equal to the critical temperature ($T_c \sim 100\,\mathrm{GeV}$) of the electroweak phase transition, such sphaleron processes were very effective.

It is because of such interactions that one can consider leptogenesis as a possible scenario of baryogenesis where the lepton asymmetry can be transformed into a baryon asymmetry by $(B - L)$ conserving processes. Therefore if neutrinos acquire Majorana mass, they can lead to large lepton asymmetries even at low energies through lepton number violating neutrinoless double β decay. This might possibly explain the puzzle of baryogenesis through leptogenesis. However, there has not yet been an observation of CP violation in the lepton sector, or of lepton number violating processes.

C and CP violation: These baryon number violating processes must also violate C and CP symmetry, hence violating T symmetry through the CPT theorem: If C is conserved, then every baryon number (B) violating reaction $X \to Y + B$ has the same width as the C-conjugate reaction

$$\Gamma(X \to Y + b) = \Gamma(\overline{X} \to \overline{Y} + \overline{b}) \tag{2.7}$$

where X, Y have $B = 0$ and b has $B = 1$. Since both processes proceed at the same rate, B is conserved over long periods of time. So C violation is a Sakharov condition. However, this is not enough. Consider a hypothetical B violating process $X \to b_L b_L$ which creates left-handed baryons (b_L). If CP is a conserved, then this proceeds at the same rate as the CP-conjugate process $\overline{X} \to \overline{b_R b_R}$, and thus

$$\Gamma(X \to b_L b_L) + \Gamma(X \to b_R b_R) = \Gamma(\overline{X} \to \overline{b_L b_L}) + \Gamma(\overline{X} \to \overline{b_R b_R}) \tag{2.8}$$

Although the C-conjugate reactions have a different width, the sum of the two will still preserve baryon number. Thus, C and CP need to be violated, so that the rate of baryon production exceeds that of antibaryon production.

In the SM, the perturbative EW Lagrangian gives rise to the weak interactions that exhibit the famous $V - A$ (vector minus axial vector) and as a result accommodate P violation [13]. The maximal C violation in weak interactions can be attributed to the presence of left-handed neutrinos and non-existence of left-handed antineutrinos. Violation of CP symmetry was first seen in the neutral kaon system through the observation of the decay of the long-lived K_{L}^0 meson to the CP even two pion final state in 1964 [50]. Since then, CP violation has been well established in weak interactions in the decays of B mesons [51]. CP violation in the SM can be accommodated through a complex phase in the Lagrangian that cannot be reabsorbed by field redefinition. This can be achieved when the Higgs scalar field acquires a vacuum expectation

value (VEV). As the work presented in this thesis deals with CP violation, I shall elaborate on this point in Sec. 2.2.1.

Universe out of thermal equilibrium: These baryon number violating processes must have occurred when the universe had moved out of the thermal equilibrium state. This condition is necessary, since in the thermal equilibrium state, the CPT symmetry would require any baryon number violating process to be balanced by the inverse reaction, hence generating no net baryon number.

Using the SM, one can confirm that the Universe undergoes a first order electroweak phase transition at temperatures equal to the electroweak breaking scale, Higgs VEV = 246 GeV. This provides a mechanism for the universe to go out of thermal equilibrium [47].

Note here that the SM meets all the Sakhorov's conditions, however the electroweak baryogenesis is not sufficient to account for the observed η_{BAU}. I will elaborate on this point later in the chapter.

2.2.1 CP violation in the SM

There are three sources of CP odd interactions that can enter the SM Lagrangian and its extensions [52], only one of which has been experimentally observed. The first source is the CP odd term that is allowed in the QCD Lagrangian which would lead to an electric dipole moment of the neutron. The second source occurs in the EW Lagrangian through the Yukawa term for quarks and are connected to the flavour mixing processes. The third source of CP violation could occur in the lepton sector through the Yukawa term for leptons which gives rise to mass terms for neutrinos, however within SM the neutrinos as massless. Therefore, CP violation in the lepton sector can only occur in the extensions of the SM. CP violation, however, has only been observed in the EW interactions of the quarks. As the work presented in this thesis deals with the quark sector, we will study the origin and manifestation of CP violation in this sector only. I will however make passing remarks on as yet unobserved CP violation in strong interactions and in the EW lepton sector for completeness. In what follows, I have also employed Einstein's summation convention.

2.2.1.1 Flavour mixing in quarks

The Yukawa term of the SM Lagrangian for quarks is given by

$$-\mathcal{L}^{\mathrm{quarks}}_{\mathrm{Yukawa}} = Y^{d}_{ij}\overline{Q^{I}_{Li}}\phi d^{I}_{Rj} + Y^{u}_{ij}\overline{Q^{I}_{Li}}\tilde{\phi}u^{I}_{Rj} + \mathrm{h.c.}\,, \tag{2.9}$$

Here $h.c.$ implies Hermitian conjugate terms, I refers to weak interaction eigenstates, the indices $i, j = 1, 2, 3$ run over the 3 generations of quarks, the up-type and down-type quarks are represented by symbols u^I and d^I respectively, Y is the 3×3 matrix of quark Yukawa couplings, Q_L^I is the left-handed quark doublet, index R refers to the right-handed singlets, ϕ is the Higgs field and $\tilde{\phi} = i\tau_2\phi^\dagger$ where τ_2 is the 2×2 antisymmetric tensor.

The complex Higgs doublet is given by

$$\phi = \begin{pmatrix} H^\pm \\ H^0 \end{pmatrix} = \begin{pmatrix} \phi^\pm \\ \mathcal{R}e(\phi^0) + i\mathcal{I}m(\phi^0) \end{pmatrix}, \tag{2.10}$$

where we have two charged (ϕ^+ and ϕ^-) and 1 neutral ($\mathcal{I}m(\phi^0)$) unphysical massless Goldstone bosons which get reabsorbed by the massive spin-1 $W^\pm$ and Z bosons to acquire their longitudinal polarisation components. When we expand the Higgs field around the VEV, one obtains $\mathcal{R}e(\phi^0) \to \frac{v+h}{\sqrt{2}}$ where h is a massive real scalar field whose excitation is the Higgs boson particle.

Inputting this into the Lagrangian, the Yukawa interactions (2.9) become

$$-\mathcal{L}_{\text{Mass}}^{\text{quarks}} = (M_d)_{ij}\overline{d_{Li}^I}d_{Rj}^I + (M_u)_{ij}\overline{u_{Li}^I}u_{Rj}^I + \text{h.c.}\,, \tag{2.11}$$

where u_L^I and d_L^I are the left-handed singlets of the Q_L^I doublet and the mass matrix is given by

$$M_q = \frac{v}{\sqrt{2}}Y^q. \quad (q = u, d \text{ types}) \tag{2.12}$$

The mass basis corresponds, by definition, to diagonal mass matrices. One can always find unitary matrices V_{qL} and V_{qR} such that

$$V_{qL}M_qV_{qR}^\dagger = M_q^{\text{diag}} \equiv \frac{v}{\sqrt{2}}\lambda_q. \tag{2.13}$$

where M_q^{diag} is the diagonal and real mass matrix and λ_q is a 3×3 diagonal matrix of Yukawa couplings. The quark mass eigenstates are then identified as

$$q_{Li} = (V_{qL})_{ij}q_{Lj}^I \quad q_{Li} = (V_{qL})_{ij}q_{Lj}^I\,. \quad (q = u, d \text{ types}) \tag{2.14}$$

Now consider the Lagrangian term associated with charged current interactions for quarks (*i.e.* interactions of $W_\mu^\pm$) taken from the kinetic term of the SM Lagrangian. In the mass basis

this term is expressed as

$$-\mathcal{L}^q_{W^\pm} = \frac{g}{\sqrt{2}}\overline{u_{Li}}\gamma^\mu (V_{uL}V^\dagger_{dL})_{ij} d_{Lj} W^+_\mu + \text{h.c.}. \tag{2.15}$$

The Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [53, 54] for quarks is therefore given by

$$V_{\rm CKM} = V_{uL}V^\dagger_{dL}\,, \tag{2.16}$$

and is a 3×3 unitary matrix ($V_{\rm CKM}{V_{\rm CKM}}^\dagger = {V_{\rm CKM}}^\dagger V_{\rm CKM} = \mathbf{1}$). The form of $V_{\rm CKM}$ is not unique:

- There is freedom in defining $V_{\rm CKM}$ in that one can permute between the various generations. This freedom is fixed by ordering the up-type quarks and the down-type quarks by their masses, *i.e.* $(u_1, u_2, u_3) \to (u, c, t)$ and $(d_1, d_2, d_3) \to (d, s, b)$. In the Q^I_{Li} doublets, if one adapts the notation $(u^I, c^I, t^I) = (u, c, t)$, then for the down-type quarks, we can write in matrix notation:

$$\begin{pmatrix} d^I \\ s^I \\ b^I \end{pmatrix}_{\text{Weak}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{Mass \& Flavour}} . \tag{2.17}$$

  Note here that the mass eigenstates are also the flavour eigenstates. It can be seen later in Sec. 2.2.3 that the same is not true in the lepton sector.

- There is further freedom in the phase structure of $V_{\rm CKM}$. This means that the number of physical parameters in $V_{\rm CKM}$ is smaller than the number of parameters in a general unitary 3×3 matrix, which is nine (three real angles and six phases). Let us define P_q ($q = u, d$) to be diagonal unitary (phase) matrices. Then, if instead of using V_{qL} and V_{qR} for the rotation (2.13) to the mass basis we use $\tilde{V}_{qL}$ and $\tilde{V}_{qR}$, defined by $\tilde{V}_{qL} = P_q V_{qL}$ and $\tilde{V}_{qR} = P_q V_{qR}$, we still maintain a legitimate mass basis since $M^{\rm diag}_q$ remains unchanged by such transformations. However, $V_{\rm CKM}$ does change:

$$V_{\rm CKM} \to P_u V_{\rm CKM} P^*_d. \tag{2.18}$$

  This freedom is fixed by demanding that $V_{\rm CKM}$ has the minimal number of phases. In the three generation case $V_{\rm CKM}$ has a single phase. (There are five phase differences between the elements of P_u and P_d and, therefore, five of the six phases in the CKM matrix can be removed.) This is the Kobayashi-Maskawa phase, $\delta_{\rm KM}$, which is the single source of *CP* violation in the quark sector of the Standard Model.

As a result of the fact that $V_{\rm CKM}$ is not diagonal, the $W^{\pm}$ gauge bosons couple to quark mass eigenstates of different generations. Within the SM, this is the only source of *flavour changing* quark interactions.

Note that we can also obtain flavour changing neutral currents (FCNC) such as $b \to s\gamma$. In the SM, these are forbidden at the tree level and occur at loop level. They are however suppressed through the Glashow-Iliopoulos-Maiani (GIM) mechanism [10]. In this mechanism the neutral current processes, along with CKM suppression factors, obtain additional loop and GIM suppression factors which are given by

$$\frac{1}{16\pi^2}\frac{m_q^2}{m_W^2}, \tag{2.19}$$

where m_W is the $W^{\pm}$ mass and m_q is the mass of the virtual quark in the loop. Therefore, the processes that involve FCNCs are most sensitive to the heavy quarks running in the loops. As a consequence of this, FCNCs are more suppressed for transitions between up-type quarks than between down-type quarks. As these FCNC processes are suppressed in the SM, they are highly sensitive to New Physics (NP) contributions which could modify their rate of transition.

2.2.1.2 CKM parameterisation

The fact that $V_{\rm CKM}$ is unitary and depends on only four independent physical parameters can be made manifest by choosing a specific parameterisation. One standard choice is [55]

$$V_{\rm CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{\rm KM}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{\rm KM}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{\rm KM}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{\rm KM}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{\rm KM}} & c_{23}c_{13} \end{pmatrix}, \tag{2.20}$$

where $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. The θ_{ij} parameters are the three real mixing angles while $\delta_{\rm KM}$ is the Kobayashi-Maskawa phase.

Following the observation of a hierarchy between the mixing angles, $s_{13} \ll s_{23} \ll s_{12} \ll 1$, Wolfenstein [56] proposed an expansion of the CKM matrix in terms of the four parameters λ, A, ρ and η ($\lambda \approx 0.23$ being the expansion parameter). We use the definitions to all orders

$$\begin{aligned} s_{12} &\equiv \lambda\,, \\ s_{23} &\equiv A\lambda^2\,, \\ s_{13}e^{-i\delta_{\rm KM}} &\equiv A\lambda^3(\rho - i\eta)\,. \end{aligned} \tag{2.21}$$

Inserting the above definitions into (2.20) gives expressions for all CKM elements

$$V_{\rm CKM} \simeq \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \tag{2.22}$$

This parameterisation shows that inter-generational mixing of the quarks is much more suppressed than mixing occurring within a given generation.

2.2.1.3 The Jarlskog Invariant

A phase-convention-independent measure of CP violation in the SM is given by

$$\begin{aligned} \mathcal{I}m(\det\left(\left[Y^d Y^{d\dagger}, Y^u Y^{u\dagger}\right]\right)) &= 2\,J\,(m_t^2-m_c^2)(m_t^2-m_u^2)(m_c^2-m_u^2) \\ &\quad\times(m_b^2-m_s^2)(m_b^2-m_d^2)(m_s^2-m_d^2)\,. \end{aligned} \tag{2.23}$$

The Jarlskog invariant [57], J, contains the dependence on the CKM elements,

$$\mathcal{I}m(V_{ij}V_{kl}V_{il}^*V_{kj}^*) = J\sum_{m,n=1}^{3}\varepsilon_{ikm}\,\varepsilon_{jln}\,, \tag{2.24}$$

where V_{ij} are the CKM matrix elements and ε_{ikm} is the total antisymmetric tensor. Owing to the unitarity of $V_{\rm CKM}$, the non-vanishing imaginary parts of all quadri-products of CKM elements are equal up to their sign. One representation of Eq. 2.24 reads, for instance, $J = \mathcal{I}m(V_{ud}V_{cs}V_{us}^*V_{cd}^*)$. A non-vanishing KM phase and hence CP violation requires $J \neq 0$. The Jarlskog parameter expressed in the parameterisation of Eq. 2.20 reads

$$J = c_{12}\,c_{23}\,c_{13}^2\,s_{12}\,s_{23}\,s_{13}\,\sin\delta_{\rm KM}\,, \tag{2.25}$$

and, using the Wolfenstein parameterisation,

$$J = A^2\lambda^6\eta\left(1-\lambda^2/2\right) + \mathcal{O}(\lambda^{10})\,. \tag{2.26}$$

Therefore to account for CP violation (see Eq. 2.23) in the quark sector, one requires all three of the following conditions to be met

- Nondegeneracy in the masses of the up-type and down-type quarks.
- None of the 3 mixing angles must be zero or $\pi/2$.

- The CP-violating phase should be neither 0 or π.

2.2.1.4 Unitarity triangles

A very useful concept is that of the *unitarity triangles*. The unitarity of the CKM matrix leads to various relations among the matrix elements, *i.e.*

$$V_{ud}V_{us}^*[\mathcal{O}(\lambda)] + V_{cd}V_{cs}^*[\mathcal{O}(\lambda)] + V_{td}V_{ts}^*[\mathcal{O}(\lambda^5)] = 0\,, \tag{2.27}$$

$$V_{us}V_{ub}^*[\mathcal{O}(\lambda^4)] + V_{cs}V_{cb}^*[\mathcal{O}(\lambda^2)] + V_{ts}V_{tb}^*[\mathcal{O}(\lambda^2)] = 0\,, \tag{2.28}$$

$$V_{ud}V_{ub}^*[\mathcal{O}(\lambda^3)] + V_{cd}V_{cb}^*[\mathcal{O}(\lambda^3)] + V_{td}V_{tb}^*[\mathcal{O}(\lambda^3)] = 0\,, \tag{2.29}$$

$$V_{cd}V_{td}^*[\mathcal{O}(\lambda^4)] + V_{cs}V_{ts}^*[\mathcal{O}(\lambda^2)] + V_{cb}V_{tb}^*[\mathcal{O}(\lambda^2)] = 0\,, \tag{2.30}$$

$$V_{ud}V_{tq}^*[\mathcal{O}(\lambda^3)] + V_{us}V_{ts}^*[\mathcal{O}(\lambda^3)] + V_{ub}V_{tb}^*[\mathcal{O}(\lambda^3)] = 0\,, \tag{2.31}$$

$$V_{ud}V_{cd}^*[\mathcal{O}(\lambda)] + V_{us}V_{cs}^*[\mathcal{O}(\lambda)] + V_{ub}V_{cb}^*[\mathcal{O}(\lambda^5)] = 0\,, \tag{2.32}$$

where the size of each term to leading order in λ is shown in the square brackets. Each of these three relations requires the sum of three complex quantities to vanish and so can be geometrically represented in the complex plane as a triangle. These are "the unitarity triangles", though the term "unitarity triangle" is usually reserved for the relation (2.29). The area of all "unitarity triangles" formed from the CKM unitary relations is half of the Jarlskog parameter J *i.e.* $|J|/2$.

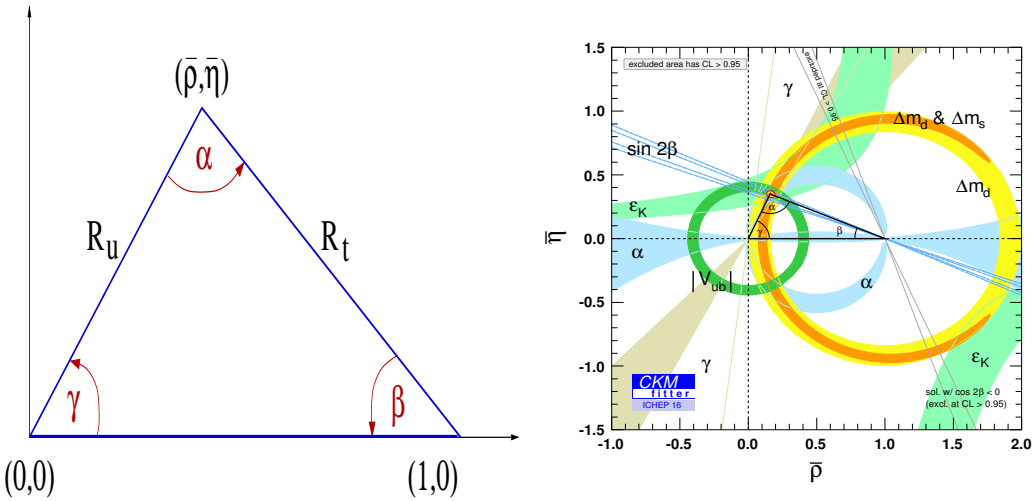

Figure 2.3: (Left) The rescaled unitarity triangle; (Right) Measurements that provide constraints on the CKM parameters where the shaded areas indicate constraints on the global parameter fit from experimental measurements of triangle side length or vertex angles [58].

One can rescale the unitarity triangle from (2.29) by dividing the lengths of all sides by

$|V_{cd}V_{cb}^*|$. The rescaled unitary triangle is sketched in the complex plan in Figure 2.3 (left), where the apex is given by the following phase-convention independent definition,

$$\overline{\rho} + i\overline{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\,. \tag{2.33}$$

When the above equation is expressed in the Wolfenstein parameterisation, one finds to all orders in λ, [6]

$$\rho + i\eta \;=\; \frac{\sqrt{1-A^2\lambda^4}(\overline{\rho}+i\overline{\eta})}{\sqrt{1-\lambda^2}\,[1-A^2\lambda^4(\overline{\rho}+i\overline{\eta})]}\,. \tag{2.34}$$

The sides R_u and R_t of the unitary triangle (the third side along the real axis is normalised to unity) read to all orders,

$$R_u = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\overline{\rho}^2+\overline{\eta}^2}\,, \qquad R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\overline{\rho})^2+\overline{\eta}^2}\,. \tag{2.35}$$

The three angles of the unitary triangle, α, β, γ, are defined by

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \qquad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \qquad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{2.36}$$

and the KM phase in the standard parameterisation (2.20) is $\delta_{\rm KM} = \gamma + A^2\lambda^4\eta + \mathcal{O}(\lambda^6)$. These angles are physical quantities and can be independently measured by CP asymmetries in B decays. It is also useful to define the two small angles of the other unitarity triangles (2.27,2.28):

$$\beta_s \equiv \arg\left[-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right], \quad \beta_K \equiv \arg\left[-\frac{V_{cs}V_{cd}^*}{V_{us}V_{ud}^*}\right]. \tag{2.37}$$

A central aim of flavor physics is to over-constrain the CKM triangles to test the unitarity conditions. The global fit to this unitarity triangle is shown in Figure 2.3 (right). The world average values for the CKM parameters and Jarlskog invariant J are shown in Table 2.3.

The experimentally measured magnitudes of the CKM matrix elements [58] are

$$|V_{\rm CKM}| = \begin{pmatrix} 0.974254^{+0.000071}_{-0.000097} & 0.22542^{+0.00042}_{-0.00031} & 0.003714^{+0.000072}_{-0.000060} \\ 0.22529^{+0.00041}_{-0.00032} & 0.973394^{+0.000074}_{-0.000096} & 0.04180^{+0.00033}_{-0.00068} \\ 0.008676^{+0.000087}_{-0.000150} & 0.04107^{+0.00031}_{-0.00067} & 0.999118^{+0.000024}_{-0.000014} \end{pmatrix}. \tag{2.38}$$

[6] Expanding Eq. 2.33 in λ leads to the well-known approximation

$$\overline{\rho} = \rho(1-\lambda^2/2) + \mathcal{O}(\lambda^4)\,, \qquad \overline{\eta} = \eta(1-\lambda^2/2) + \mathcal{O}(\lambda^4)\,.$$

Table 2.3: World average values for the CKM parameters and Jarlskog invariant J [58, 59].

Parameter	Value
A	$0.8227^{+0.0066}_{-0.0136}$
λ	$0.2254^{+0.0004}_{-0.0003}$
$\overline{\rho}$	$0.1504^{+0.0121}_{-0.0062}$
$\overline{\eta}$	$0.3540^{+0.0069}_{-0.0076}$
α	$(90.40^{+2.0}_{-1.0})^{\circ}$
β	$(22.62^{+0.44}_{-0.42})^{\circ}$
γ	$(67.01^{+0.88}_{-1.99})^{\circ}$
J	$(3.04^{+0.21}_{-0.20}) \times 10^{-5}$

The empirical value of J is small compared with its mathematical maximum of $1/(6\sqrt{3}) \approx 0.1$, showing that CP violation is suppressed as a consequence of the strong hierarchy exhibited by the CKM matrix elements.

2.2.2 Manifestations of CP violation in the quark sector

Quarks carrying colour charge are bound into colourless hadron states. As a result, CP violation in the quark flavour transitions can manifest in different ways. In this section a general formalism for, and classification of, CP violation in the decay of a hadron state B and its CP conjugate $\overline{B}$ to a multi-particle final state f and its CP conjugate $\overline{f}$ is given.

Consider a system that is initially in a mixed state of $|B\rangle$ and $|\overline{B}\rangle$ at time $t = 0$ *i.e.*

$$|\psi(t=0)\rangle = a(0)|B\rangle + b(0)|\overline{B}\rangle\,. \tag{2.39}$$

Such a mixed state can occur in SM only with a neutral meson (B^0), since the FCNC transitions give rise to "box diagrams" where $B^0 \leftrightarrow \overline{B}^0$ oscillations can occur. Conservation of charge and baryon currents in the SM forbid oscillations of charged mesons or baryons.

In the Schrodinger representation, this state evolves in time, according to the time evolution operator $\hat{U}(t)$, into a superposition of all states allowed by the conservation laws as

follows

$$\begin{aligned}|\psi(t)\rangle &= \hat{U}(t)|\psi(t=0)\rangle = e^{-i\mathcal{H}t}|\psi(0)\rangle \\ &= |B\rangle\langle B|\hat{U(t)}|B\rangle + |\overline{B}\rangle\langle\overline{B}|\hat{U(t)}|\overline{B}\rangle + \sum_i |f_i\rangle\langle f_i|\hat{U(t)}|\psi(t=0)\rangle \\ &= a(t)|B\rangle + b(t)|\overline{B}\rangle + \sum_i c_i(t)|f_i\rangle\,,\end{aligned}$$

where $\mathcal{H}$ is the Hamiltonian governing the interactions and $c_i(t)$ are the complex coefficients.

One is interested in the time-evolved states $|B(t)\rangle$ and $|\overline{B}(t)\rangle$ only. Therefore using the Wigner–Weisskopf approximation,[7] the time evolution is determined by a 2×2 effective non-Hermitian Hamiltonian matrix $\mathcal{H}_{\rm eff}$ [60]. If $\mathcal{H}_{\rm eff}$ were Hermitian then the mesons would only oscillate and not decay. This effective Hamiltonian is given by

$$\mathcal{H}_{\rm eff} = M - \frac{i}{2}\,\Gamma \tag{2.40}$$

where M and Γ are Hermitian matrices and are associated with $(B^0, \overline{B}^0) \leftrightarrow (B^0, \overline{B}^0)$ transitions via off-shell (dispersive) and on-shell (absorptive) intermediate states, respectively. Diagonal elements of M and Γ are associated with the flavor-conserving transitions $B^0 \to B^0$ and $\overline{B}^0 \to \overline{B}^0$ while off-diagonal elements are associated with flavor-changing transitions $B^0 \leftrightarrow \overline{B}^0$.

The eigenvectors of $\mathcal{H}_{\rm eff}$ have well defined masses and decay widths and are given by

$$|B_L\rangle = p|B^0\rangle - q|\overline{B}^0\rangle\,, \tag{2.41}$$

$$|B_H\rangle = p|B^0\rangle + q|\overline{B}^0\rangle\,, \tag{2.42}$$

where p and q are the complex mixing terms, B_L and B_H are light (L) and heavy (H) mass eigenstates respectively. The following relations are also satisfied

$$\begin{aligned}|p|^2 + |q|^2 &= 1 \\ \left(\frac{q}{p}\right)^2 &= \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}} \\ M_{11} = M_{22} \text{ and } \Gamma_{11} = \Gamma_{22} &\qquad (\text{if } CPT \text{ or } CP \text{ conserved}) \\ \left|\frac{q}{p}\right| &= 1 \qquad (\text{if } CP \text{ or } T \text{ conserved})\end{aligned}$$

[7] In Wigner–Weisskopf approximation one can ignore the weak interaction for those particles to which the initial hadrons decay. Therefore the decay products are considered to be stable.

The eigenvalues corresponding to the eigenvectors of $\mathcal{H}_{\rm eff}$ are

$$\mu_{L,H} = M_{L,H} - i\Gamma_{L,H}$$

where $M_{L,H}, \Gamma_{L,H}$ are masses and decay-widths of their respective $|B_{L,H}\rangle$ states.

Therefore if one were to start with the B^0 or $\overline{B}^0$ initial states, which can be expressed in the basis of eigenvectors of $\mathcal{H}_{\rm eff}$ as,

$$\begin{aligned} |\psi(t=0)\rangle_B &= |B\rangle = \frac{1}{2p}(|B_H\rangle + |B_L\rangle) \\ |\psi(t=0)\rangle_{\overline{B}} &= |\overline{B}\rangle = \frac{1}{2q}(|B_L\rangle - |B_H\rangle) \end{aligned}$$

under time evolution one would get

$$\begin{aligned} |\psi(t)\rangle_B &= \frac{1}{2p}(e^{i\mu_H t}|B_H\rangle + e^{i\mu_L t}|B_L\rangle) = g_+(t)\,|B\rangle - (q/p)\,g_-(t)|\overline{B}\rangle \\ |\psi(t)\rangle_{\overline{B}} &= \frac{1}{2q}(e^{i\mu_L t}|B_L\rangle - e^{i\mu_H t}|B_H\rangle) = g_+(t)\,|\overline{B}\rangle - (p/q)\,g_-(t)|B\rangle. \end{aligned}$$

where

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

where $\Delta m \equiv M_H - M_L$ which is positive by definition, $\Delta\Gamma \equiv \Gamma_H - \Gamma_L$ whose sign can be experimentally determined and theoretically predicted, $m \equiv \frac{M_H+M_L}{2}$ and $\Gamma \equiv \frac{\Gamma_H+\Gamma_L}{2}$. The following relations are also satisfied

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = (4|M_{12}|^2 - |\Gamma_{12}|^2), \quad \Delta m\Delta\Gamma = 4\mathcal{R}e(M_{12}\Gamma_{12}^*). \tag{2.43}$$

Using these relations the time-dependent decay rates can be written as

$$\begin{aligned}\Gamma(|\psi(t)\rangle_B \to f) &= \left|g_+(t)A_f - (q/p)g_-(t)\overline{A}_f\right|^2 \\ &= \mathcal{N}_f\,|A_f|^2\,e^{-\Gamma t}\left\{\frac{1+|\lambda_f|^2}{2}\,\cosh\frac{\Delta\Gamma\,t}{2} + \right. \\ &\qquad +\frac{1-|\lambda_f|^2}{2}\,\cos(\Delta m\,t) - \mathcal{R}e(\lambda)_f\,\sinh\frac{\Delta\Gamma\,t}{2} \\ &\qquad \left. -\mathcal{I}m(\lambda)_f\,\sin(\Delta m\,t)\right\}. \end{aligned} \tag{2.44}$$

where $\mathcal{N}_f$ is a time-independent normalization factor, $A_f = \langle f|\mathcal{S}|B\rangle$, $\overline{A}_f = \langle f|\mathcal{S}|\overline{B}\rangle$, $\mathcal{S}$ is the S-matrix of the transition and $\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}$. Similarly one can write,

$$\begin{aligned}\Gamma(|\psi(t)\rangle_{\overline{B}} \to f) &= \mathcal{N}_f\,|A_f|^2\,\frac{1}{1-a}\,e^{-\Gamma t}\left\{\frac{1+|\lambda_f|^2}{2}\,\cosh\frac{\Delta\Gamma\,t}{2} - \frac{1-|\lambda_f|^2}{2}\,\cos(\Delta m\,t)\right. \\ &\qquad \left. -\mathcal{R}e(\lambda)_f\,\sinh\frac{\Delta\Gamma\,t}{2} + \mathcal{I}m(\lambda)_f\,\sin(\Delta m\,t)\right\}. \end{aligned} \tag{2.45}$$

The decay rates to the CP-conjugate final state $\overline{f}$ are obtained analogously, with $\mathcal{N}_f = \mathcal{N}_{\overline{f}}$ and the substitutions $A_f \to A_{\overline{f}}$ and $\overline{A}_f \to \overline{A}_{\overline{f}}$ in Eqs. 2.44 and 2.45.

Note that these equations hold for two-body decays. For multi-body decays, where there are quasi-two body contributions, these equations hold at a given point in phase space.

2.2.2.1 Classification of CP violation

We can distinguish three types of CP-violating effects occurring in hadron decays [51, 60]:

CP violation in decay: This type of CP violation can occur both in charged and neutral hadrons. This is the only possible source of CP asymmetries in charged meson or baryon decays, as mixing effects are absent. The CP asymmetries are defined as

$$\mathcal{A}_{f\overline{f}} \equiv \frac{\Gamma(|\psi\rangle_{\overline{B}} \to \overline{f}) - \Gamma(|\psi\rangle_B \to f)}{\Gamma(|\psi\rangle_{\overline{B}} \to \overline{f}) + \Gamma(|\psi\rangle_B \to f)} = \frac{|\overline{A}_{\overline{f}}/A_f|^2 - 1}{|\overline{A}_{\overline{f}}/A_f|^2 + 1}\,. \tag{2.46}$$

Therefore when $\mathcal{A}_{f\overline{f}} \neq 0$, we get $|\overline{A}_{\overline{f}}/A_f| \neq 1$. For this condition to be met, it must be possible to decompose A_f into at least two amplitude components a_1 and a_2. These components correspond, for example, to tree level, see Figure 2.4 (left), and loop level FCNC, see Figure 2.4

(right) transitions.

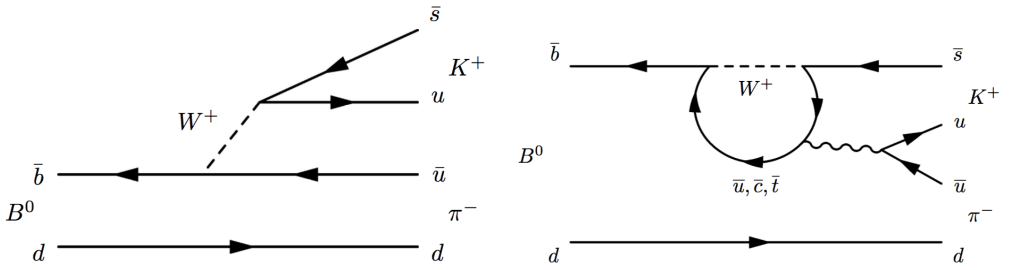

Figure 2.4: Example (left) tree diagram and (right) gluonic penguin diagram for $B^0 \to K^+\pi^-$ [61].

The amplitudes a_i are expressed in terms of the magnitude $|a_i|$, 'weak phase' ϕ_i and 'strong phase' δ_i. Therefore, A_f and $\overline{A}_{\overline{f}}$ can be expressed as

$$\begin{aligned} A_f &= |a_1|e^{i(\delta_1+\phi_1)} + |a_2|e^{i(\delta_2+\phi_2)}, \\ \overline{A}_{\overline{f}} &= |a_1|e^{i(\delta_1-\phi_1)} + |a_2|e^{i(\delta_2-\phi_2)}\,. \end{aligned} \tag{2.47}$$

The CP asymmetry can then be expressed as follows

$$\mathcal{A}_{f\overline{f}} = -\frac{2|a_1a_2|\sin(\delta_2-\delta_1)\sin(\phi_2-\phi_1)}{|a_1|^2+|a_2|^2+2|a_1a_2|\cos(\delta_2-\delta_1)\cos(\phi_2-\phi_1)}. \tag{2.48}$$

Therefore for a CP asymmetry to exist in the decay, one requires both $\phi_2 - \phi_1$ and $\delta_2 - \delta_1 \neq 0$. A diagrammatic representation of this is shown in Figure 2.5. Examples of decays that exhibit this class of CP violation are $\overline{B}^0 \to K^-\pi^+$ [62], $\overline{B}^0_s \to K^+\pi^-$ [62], $B^+ \to K^+K^-\pi^+$ [63], *etc.*

CP violation in mixing: The neutral mesons, as we know, undergo oscillations through "box diagrams" (see Figure 2.6 for $B^0 - \overline{B}^0$ oscillations). The CP violation in mixing arises when the two neutral mass eigenstate admixtures cannot be chosen to be CP eigenstates and occurs when $|q/p| \neq 1$.

The evidence of this type of CP violation can be obtained by measuring the CP asymmetry in flavour specific meson decays, *i.e.* decays where the final state tags the flavour of the decaying meson state. For example, the semileptonic decay of a neutral meson of type $\overline{(B)} \to \ell^{(+)}X$ is a flavour specific decay where the charge of the final state lepton tags the flavour

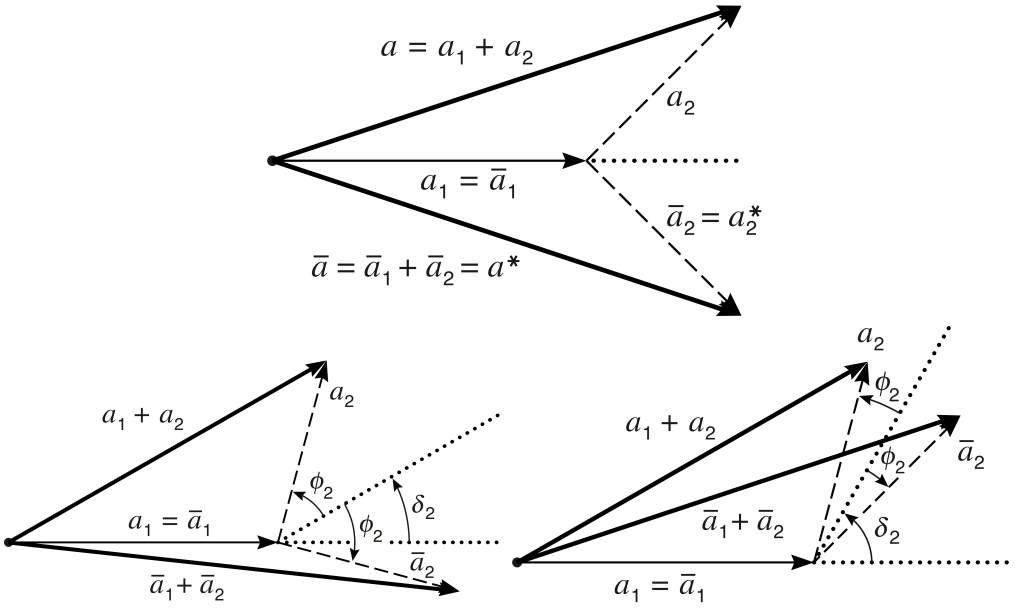

Figure 2.5: Illustrative Argand diagrams for amplitudes a_1, a_2 and their conjugates $\overline{a_1}$, $\overline{a_2}$ that can interfere to cause CP violation in decay. In the top figure there is a relative weak phase between a_2 and a_1, but no relative strong phase which leads to zero CP asymmetry since $|a| = |\overline{a}|$. In the figures at the bottom, there are both relative weak and strong phase between a_2 & a_1 which lead to $|a| \neq |\overline{a}|$, giving a CP asymmetry [64].

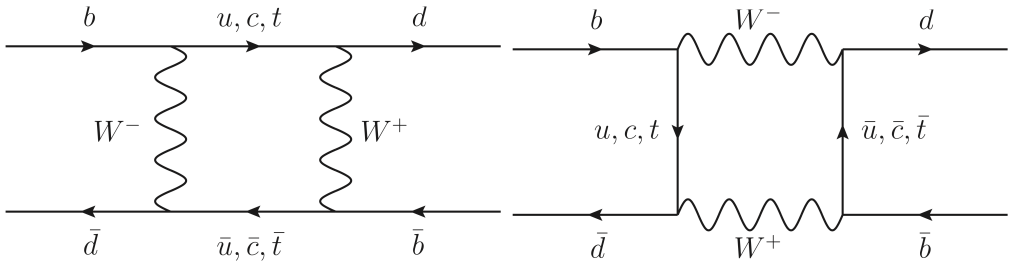

Figure 2.6: FCNC transitions that lead $B^0 - \overline{B}^0$ oscillations through box "diagrams" [65].

of the b-meson. The asymmetry in such decays is given as

$$\begin{aligned} \mathcal{A}_{\text{flavspec}}(t) &\equiv \frac{d\Gamma/dt[|\psi(t)\rangle_{\overline{B}} \to \ell^+ X] - d\Gamma/dt[|\psi(t)\rangle_B \to \ell^- X]}{d\Gamma/dt[|\psi(t)\rangle_{\overline{B}} \to \ell^+ X] + d\Gamma/dt[|\psi(t)\rangle_B \to \ell^- X]} \\ &= \frac{1 - |q/p|^4}{1 + |q/p|^4} \,. \end{aligned} \tag{2.49}$$

In the approximation that $|\Gamma_{12}/M_{12}| \ll 1$ (valid for B^0 and B_s^0 mesons), the CP asymmetry in

flavour-specific neutral-meson decays becomes

$$\mathcal{A}_{\rm flavspec} = -\left|\frac{\Gamma_{12}}{M_{12}}\right| \sin(\phi_M - \phi_\Gamma). \tag{2.50}$$

where weak phases $\phi_M = \arg(M_{12})$ and $\phi_\Gamma = \arg(\Gamma_{12})$. The extraction of $\phi_M - \phi_\Gamma$ from $A_{\rm flavspec}$ requires the prior knowledge of $|\Gamma_{12}/M_{12}|$. To date, CP violation in mixing has only been observed in the K^0 mesons.

CP violation in interference between a decay without mixing and a decay with mixing: This type of CP violation occurs when B^0 and $\overline{B}^0$ can decay to the same final state (*e.g.* a CP eigenstate, f_{CP}, with eigenvalues $\eta_{CP} = \pm 1$), so that the amplitudes for decays with and without mixing can interfere. The schematic diagram of the amplitudes involved is shown in Figure 2.7.

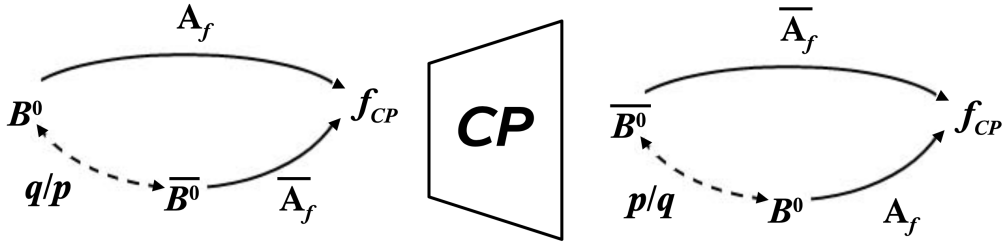

Figure 2.7: Schematic diagram of the amplitudes involved in the B^0 decay to CP eigenstate f_{CP} and its CP conjugate process. The CP asymmetry is due to the interference between mixing, described by parameters p and q, and the decay amplitudes A_f and $\overline{A}_f$.

This form of CP violation can be observed, by measuring a non-zero value for the CP asymmetry parameter which is given by

$$\begin{aligned} \mathcal{A}_{f_{CP}}(t) &\equiv \frac{d\Gamma/dt[|\psi(t)\rangle_{\overline{B}} \to f_{CP}] - d\Gamma/dt[|\psi(t)\rangle_B \to f_{CP}]}{d\Gamma/dt[|\psi(t)\rangle_{\overline{B}} \to f_{CP}] + d\Gamma/dt[|\psi(t)\rangle_B \to f_{CP}]} \\ &= \frac{S_f \sin(\Delta m t) - C_f \cos(\Delta m t)}{\cosh(\Delta\Gamma t/2) + A_{\Delta\Gamma} \sinh(\Delta\Gamma t/2)} \end{aligned}$$

where

$$S_f \equiv \frac{2\,\mathcal{I}m(\lambda_f)}{1+|\lambda_f|^2}, \quad C_f \equiv \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \quad A_{\Delta\Gamma} \equiv \frac{2\,\mathcal{I}m(\lambda_f)}{1+|\lambda_f|^2}, \quad \lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f} \ . \tag{2.51}$$

If $|\lambda_f| = 1$, the interference between decays with and without mixing is the only source of the

asymmetry and Eq. 2.51, in this case reduces to

$$\mathcal{A}_{f_{CP}}(t) = \mathcal{I}m(\lambda_f)\ \sin(\Delta m t) \tag{2.52}$$

$$= \eta_{CP} \sin(\phi_M + 2\phi_f)\ \sin(\Delta m t)\ . \tag{2.53}$$

Therefore this asymmetry arises when $\mathcal{I}m(\lambda_f) \neq 0$. Note that in this case the measured phase is purely a weak phase, and no hadronic parameters are involved in the extraction of its value from $\mathcal{I}m(\lambda_f)$. An example of a decay that exhibits this class of CP violation is $B^0 \to J/\psi\, K^0_{\rm S}$ [66], *etc.*

2.2.3 Other sources of CP violation

I mention here very briefly other sources of CP violation that might occur in QCD interactions or in the lepton sector. None of these sources have been experimentally confirmed yet.

Lepton Sector: As in the quark sector we can also write the Yukawa interactions of the lepton sector as

$$-\mathcal{L}^{\rm leptons}_{\rm Yukawa} = Y^e_{ij}\overline{L^I_{Li}}\phi E^I_{Rj} + \text{h.c.}\,. \tag{2.54}$$

Note that such mass terms in the Lagrangian require the existence of right-handed neutrinos or left-handed antineutrinos and therefore this scenario can only occur in the extensions of the SM where the neutrinos are either Dirac or Majorana particles. In Eq. 2.54, I refers to weak interaction eigenstates, the indices $i, j = 1, 2, 3$ run over the 3 generations of the leptons, Y is the 3×3 Yukawa matrix of leptons, L^I_L is the left-handed lepton doublet, E^I_R refers to the right-handed lepton singlet.

After the Higgs acquires a VEV, we can rotate from weak interaction eigenstates into mass eigenstates using the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix given by

$$\begin{pmatrix} {\nu_e}^I \\ {\nu_\mu}^I \\ {\nu_\tau}^I \end{pmatrix}_{\text{Weak and Flavour}} = \begin{pmatrix} U_{e1} & U_{e1} & U_{e1} \\ U_{\mu 2} & U_{\mu 2} & U_{\mu 2} \\ U_{\tau 3} & U_{\tau 3} & U_{\tau 3} \end{pmatrix}_{\rm PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}_{\text{Mass}} . \tag{2.55}$$

The complex 3×3 PMNS matrix is characterised by 3 mixing angles and one CP violating phase. In this scenario, the PMNS phase is responsible for all CP violating effects in the lepton sector yet none have been observed. Lepton flavour violating processes have been observed in neutrino oscillation but charged lepton flavour violation is highly suppressed due to the GIM mechanism and has not yet been observed. In the SM, the neutrinos are massless, however the observation of neutrino oscillations indicates non-zero Yukawa couplings for neutrinos suggesting an extension

of the SM.

Strong interaction: In the QCD Lagrangian, we have freedom to add the following gauge invariant term:

$$\mathcal{L}_\theta = \theta \frac{\alpha_s}{8\pi} \epsilon^{\mu\nu\rho\sigma} F^a_{\mu\nu} F^a_{\rho\sigma} = \theta \frac{\alpha_s}{8\pi} F^a_{\mu\nu} \tilde{F}^{a\mu\nu} \tag{2.56}$$

where $\epsilon^{\mu\nu\rho\sigma}$ is the antisymmetric tensor, where $\tilde{F}^{\mu\nu}$ is the dual field strength tensor and α_s is the strong coupling constant. The above term could be rewritten in terms of chromoelectric (E^a) and chromomagnetic (B^a) operators as follows

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = \mathbf{E} \cdot \mathbf{B} \,. \tag{2.57}$$

This is a CP odd term which will give rise to a non-zero electric dipole moment of the neutron (nEDM). However nEDM experiments set stringent constraints, *i.e.* $\theta < 10^{-10}$ [67]. The question as to why this θ-parameter in QCD has such a low value is referred to as the strong CP problem.[8]

It was noted before that the SM meets all Sakhorov's conditions to generate an asymmetric matter-antimatter Universe. However, accounting for the CP violation within the SM and calculating the dimensionless constant introduced in Eq. 2.23, yields $\mathcal{I}m(\det\left(\left[Y^d Y^{d\dagger}, Y^u Y^{u\dagger}\right]\right)) \sim 10^{-20}$. This number is much too small to account for the observed $\eta_{\rm BAU}$ shown in (2.3) and, thus far, attempts to utilise this source of CP violation for electroweak baryogenesis have been unsuccessful. It seems that there must be physics beyond the Standard Model (BSM) that can introduce new CP violating terms, which do not interfere with other measurements but are large enough to cause the baryon asymmetry. Therefore, we need to investigate as many potential CP violating modes as possible to search for new sources of CP violation introduced by BSM theories.

In this thesis I have chosen to investigate CP violation in charmless hadronic decays of $\Xi_b^-\,(\Omega_b^-)$ baryons to the $ph^-h'^-$ final state, where h, h' is either a K or π. These decays proceed via $b \to u$ transitions at tree level or via $b \to s$ or $b \to d$ FCNC transitions. As mentioned previously, FCNC transitions count the number of heavy degrees of freedom that occur as virtual particles in the loops and serve as a window into New Physics (NP) at scales much above the energy being probed. Therefore, many beyond the Standard Model (BSM) theories which introduce heavy new particles and new interactions to address the aforementioned matter-antimatter asymmetry problem could be tested and constrained.

[8] The QED Lagrangian also has the freedom of acquiring a similar term. However in QED this term is divergent and can be eliminated from the theory using renormalisation and regularisation techniques. In QCD this divergent term cannot be eliminated since non-perturbative effects dominate.

CHAPTER 3

Experimental set-up

“ We are the read out of the Big Bang. ”

Neil Turok, *Endless Universe*

The analyses described in the thesis use data collected by the Large Hadron Collider beauty (LHCb) experiment, which is located at one of the collision points of the Large Hadron Collider at the European Centre for Nuclear Research (CERN). In this chapter, in Sec. 3.1, I will detail the different facilities in the CERN accelerator complex used to accelerate and collide two proton beams. In Sec. 3.2, I will discuss the details of the LHCb experiment.

3.1 CERN accelerator complex

CERN is located near Geneva, on the Franco-Swiss border and hosts an accelerator complex that is a succession of machines that accelerate particles to increasingly high energies. Beams of particles are boosted in energy at each machine before being injected into the next machine in the sequence. The Large Hadron Collider (LHC) forms the final stage of the accelerator complex where two beams, composed of either protons or fully-stripped ions most often lead $^{208}\mathrm{Pb}^{82+}$ ions), are brought to collision. The other accelerators in the chain have their own experimental halls where beams are used for experiments at lower energies. This section draws on Ref. [68] to present the material that is relevant to the thesis.

Fig. 3.1 shows a schematic representation of CERN accelerator complex. Hydrogen atoms with stripped electrons are injected into the the first accelerator in the chain, Linac 2, which

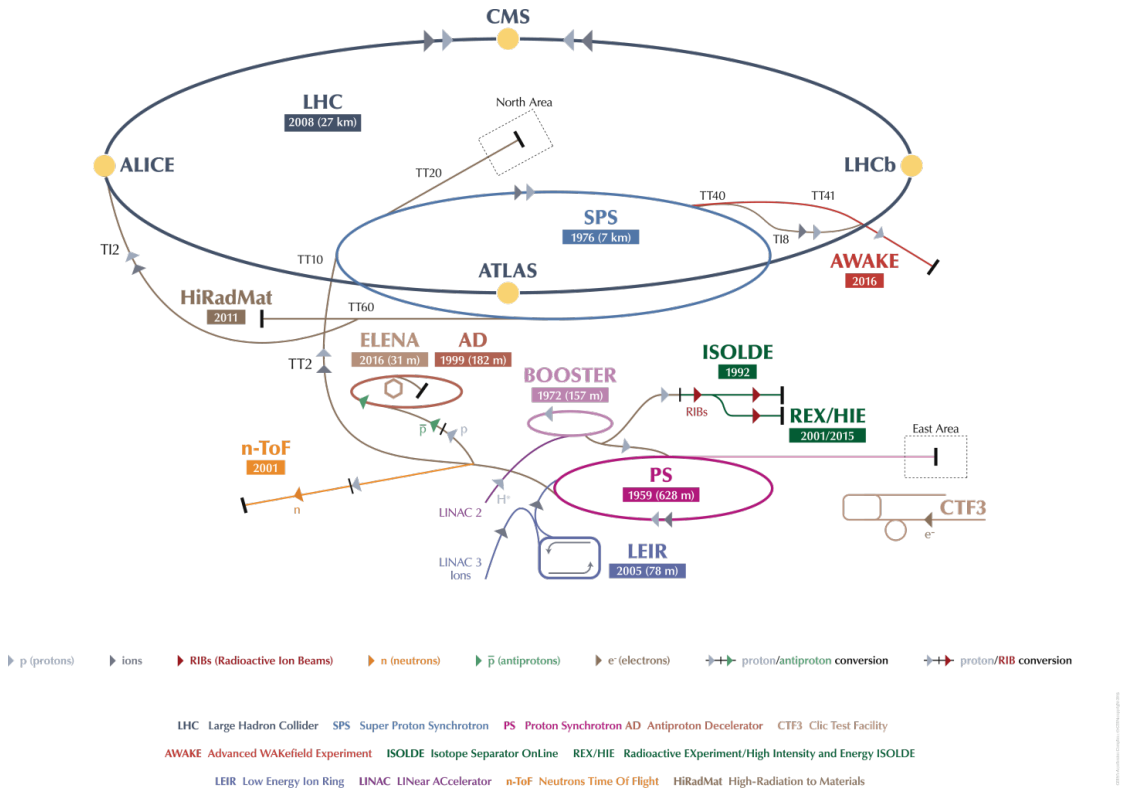

Figure 3.1: A schematic representation of the CERN accelerator complex [69].

accelerates the protons to the energy of 50×10^6 eV (50 MeV). The beam is then injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4×10^9 eV (1.4 GeV), followed by the Proton Synchrotron (PS), which further accelerates the beam to 25 GeV. Protons are then sent to the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV. The protons are finally transferred to the two beam pipes of the LHC, which circulates the beams in opposite directions to each other. The LHC is a 26.6 km circular collider, which has been designed to accelerate proton beams to a record energy of 7×10^{12} eV (7 TeV).

The proton beams inside the LHC ring, rather than being continuously distributed, are arranged in bunches where each bunch contains $\mathcal{O}(10^{11})$ protons. This is due to the use of radio-frequency (RF) cavities for acceleration. The space between the bunches is referred to as 'bunch spacing'. There are also gaps that are larger than the typical bunch spacing which are referred to as 'abort gaps'. These abort gaps accommodate injection or extraction kicker magnet rise times. During the period of 2015-2017, the LHC was able to achieve 2460 bunches per beam (its design value is 2808 bunches per beam), with a minimum bunch spacing of 25 ns. The

extraction kicker of the LHC beam dumping system also required an abort gap of $3\,\mu\mathrm{s}$.

To achieve record energy, the beams at the LHC are accelerated using RF cavities with a frequency of $400\,\mathrm{MHz}$. With each lap around the LHC, every proton gains on average $485\,\mathrm{keV}$ in energy. It takes approximately 4 minutes and 20 seconds to fill each LHC beam, and 20 minutes for the protons to reach their current maximum energy of $6.5\,\mathrm{TeV}$ per beam.

The LHC ring has eight bending arcs where superconducting dipole magnets generating magnetic fields up to 8.3 T are used to bend the beam. The ring also consists of eight straight sections, four of which are occupied by the radio-frequency accelerating cavities, beam dump areas, beam cleaning and collimation apparatus. The other four straight sections are where the counter-circulating beams are focused into collision by multipole magnets. At each crossing point (also referred to an interaction point) of the beams on the LHC ring, there stands one of the four large LHC experiments: ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb.

The ATLAS and CMS experiments are 'General Purpose Detectors' (GPD), which are investigating a wide range of high p_{T} phenomena. The ALICE experiment is focused on the physics of the quark-gluon plasma produced in heavy ion collisions. The LHCb experiment is designed to study decays of hadrons containing bottom and charm quarks, with particular focus on searching for non-SM sources of CP violation and evidence for new physics in rare decays. In addition, there are three smaller experiments. The TOTEM experiment (TOTal cross-section, Elastic scattering and diffractive dissociation Measurement), located either side of CMS, measures the total elastic and single or double diffractive cross-section of pp collisions. The LHCf experiment (Large Hadron Collider forward), located either side of ATLAS, measures neutral particle production for use in cosmic ray research. Finally, MoEDAL (Monopole and Exotics Detector at the LHC), located in the LHCb cavern, searches for magnetic monopoles.

The number of events per second generated in the LHC collisions at an interaction point is given by

$$\frac{dN_{\mathrm{event}}}{dt} = \mathcal{L} \times \sigma_{\mathrm{event}} \tag{3.1}$$

where σ_{event} is the cross-section of the event under study and $\mathcal{L}$ is the machine luminosity. To calculate the instantaneous machine luminosity delivered by LHC, one needs information on number of particles per bunch ($N_b = 1.15 \times 10^{11}$), number of bunches per beam ($n_b = 2808$), revolution frequency ($f_{\mathrm{rev}} = 11\,\mathrm{kHz}$), normalised transverse beam emittance ($\epsilon_n = 3.75\mathrm{mm}\ \mu\mathrm{rad}$) defined as the measure of spread of the protons in position and momentum space, beta function ($\beta^* = 0.55\,\mathrm{m}$) defined as the transverse size of the beams at the collision point and geometric luminosity reduction factor (F) that is related to beam-crossing angle ($\theta = 300\mu\mathrm{rad}$) at the interaction point. The instantaneous machine luminosity is related to these parameters by

the expression

$$\mathcal{L} = \frac{N_b^2 n_b f_{\rm rev} \gamma_r}{4\pi\epsilon_n \beta^*} F \,, \tag{3.2}$$

where γ_r is the Lorentz factor. To study possible beyond the SM contributions requires both high beam energies and high beam intensities. The LHC is designed to collide proton beams with a centre-of-mass energy of 14 TeV ($\sqrt{s}$ = 14 TeV), although the maximum achieved to date is 13 TeV.

For pp collisions, the two high luminosity experiments, ATLAS and CMS, both aimed to record a peak luminosity of $\mathcal{L} = 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. However, the maximum recorded peak instantaneous luminosity by GPDs is twice their design value $2.06 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The LHCb is a lower luminosity experiment designed to record a peak luminosity of $\mathcal{L} = 2 \times 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, but LHCb has been routinely running at twice the design value *i.e.* $\mathcal{L} = 4 \times 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

As seen in Figure 3.1, the accelerator complex also includes the Antiproton Decelerator and the Online Isotope Mass Separator (ISOLDE) facility, and feeds the Compact Linear Collider test area, as well as the neutron time-of-flight facility (nTOF).

3.2 The LHCb experiment

The LHCb detector is a forward-arm spectrometer designed primarily for the investigation of b-flavoured and c-flavoured hadron decays. At the LHC, the $b\bar{b}$ and $c\bar{c}$ quark pairs are predominantly produced via gluon initiated processes (*i.e.* at leading-order via $gg \to b\bar{b}$ and at next-to-leading-order via $gg \to b\bar{b}g$) [70, 71]. Each of these quarks, depending on their flavour, can hadronise to form either b-flavoured or c-flavoured hadrons. The production cross-section of $b\bar{b}$ quark pairs at $\sqrt{s}$ = 7 TeV measured by LHCb is $\sigma(pp \to b\bar{b}X) = 284 \pm 20(\mathrm{stat}) \pm 49(\mathrm{syst})\mu\,\mathrm{b}$ [72].[1] The production cross-section of $c\bar{c}$ quark pairs is $\sim 20 \times \sigma(pp \to b\bar{b}X)$. The b (c) and $\bar{b}$ ($\bar{c}$) quarks, after production, undergo hadronisation to give rise to various b-hadrons (c-hadrons). Interactions between the produced quarks and the beam fragments, that differ from those for produced antiquarks, can give rise to asymmetry between the production rates of b-hadrons (c-hadrons) and $\bar{b}$-hadrons ($\bar{c}$-hadrons). Such production asymmetry needs to be either known a priori or cancelled out when investigating CP violation in the decay of a hadron to a specific final state.

Heavy quark pairs are produced, at the LHC, at a small angle with respect to the beam axis. The polar angle (θ defined as angle between b-jet direction and beam axis) and pseudorapidity ($\eta = -\ln(\tan\theta/2)$) distributions of b quarks in simulated LHC collisions are shown in Figure 3.2. The polar angle coverage of LHCb extends from 15 mrad to 300 mrad in the horizontal (bending)

[1] 1 barn (b) is approximately the cross-sectional area of a uranium nucleus, *i.e.* $1\,\mathrm{b} = 10^{-24}\mathrm{cm}^2$.

plane and to 250 mrad in the vertical (non-bending) plane. This translates to a pseudorapidity range of $1.6 < \eta < 4.9$ in which roughly 25% of the $b\bar{b}$ pairs produced in LHC collisions are contained.

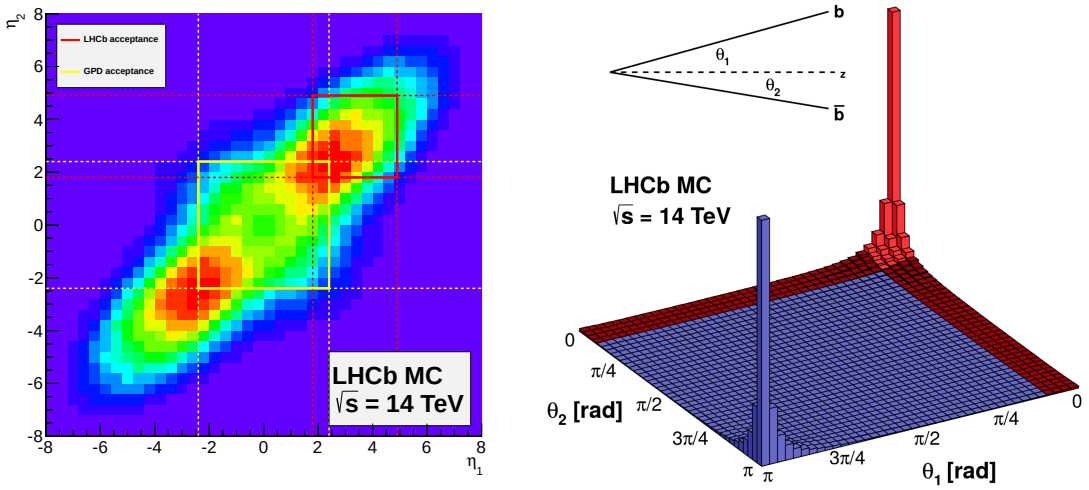

Figure 3.2: (Left) Pseudorapidity of $b\bar{b}$ quark pairs in pp collisions at $\sqrt{s} = 14\,\text{TeV}$ generated using PYTHIA 8. Here η_1 (η_2) is the pseudorapidity of the b ($\bar{b}$) quark [73]. The solid lines show the acceptance of (red) LHCb compared to (yellow) the general purpose detectors. (Right) Polar angle distribution of simulated b quarks. The red region is the acceptance of LHCb.

The LHCb experiment collected data between the years 2009-2018 with a long shutdown period between 2013-2014. In this thesis, I will be using data samples corresponding to data-taking years of 2011-2012 (this sample is referred to as 'Run I' sample) and 2015-2016 (this sample is referred to as 'Run II' sample). The data samples corresponding to years 2009 and 2010 are not used since they are too small compared to samples from years 2011 and 2012. The samples corresponding to years 2017 and 2018 are not used as they were not available in time for the analysis. The integrated luminosity recorded by LHCb for different years is shown in Figure 3.3 (left). The LHCb experiment operates at a constant instantaneous luminosity using a 'luminosity levelling' technique [74]. In this technique, the lateral separation of the beams at the crossing point is adjusted periodically to achieve a constant target luminosity. As the beam intensity reduces, the separation reduces accordingly. The instantaneous luminosity over an example fill can be seen in Figure 3.3 (right). Luminosity levelling is performed so that each event contains a constant average number of visible pp interactions per bunch-crossing (also referred to as pile-up, μ_{vis}).[2] During Run I (2011–2012), μ_{vis} of 1.5 was achieved whereas for

[2]The average number of visible pp interaction per bunch-crossing (μ_{vis}) seen by a given detector with a given

Run II (2015–2016) it was 1.1. This facilitates the measurement of b- and c-hadron decay times to a very high precision and helps in the reduction of backgrounds. Luminosity levelling also helps reduce radiation damage to the high precision silicon strip vertex tracker surrounding the pp interaction region.

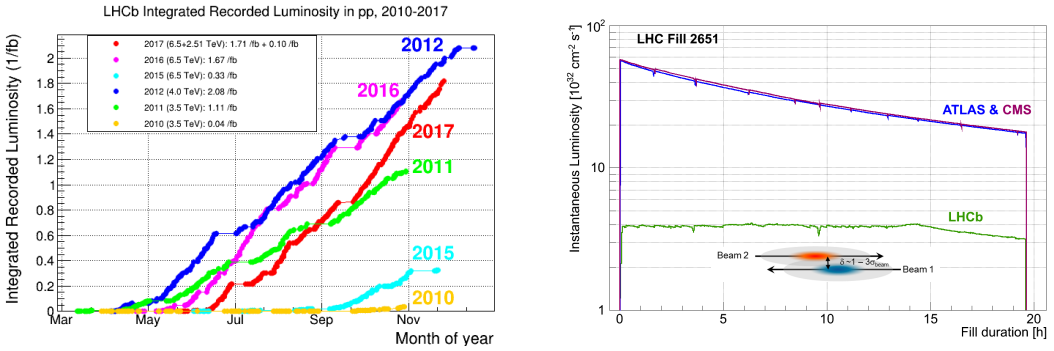

Figure 3.3: (Left) Integrated luminosity of proton-proton collisions recorded by the LHCb detector for different data-taking years. The legend shows the beam energy and integrated luminosity for each year [73]. (Right) Evolution of the instantaneous luminosity at LHCb, ATLAS and CMS during a typical 'fill' in Run I. The instantaneous luminosity at LHCb is constant until about 14 hours, when the beams have zero separation [75]. It can be also seen here that the separation between the beams (δ) is 1 minus three times the beam spread ($\sigma_{\rm beam}$).

A schematic representation of the LHCb detector is shown in Figure 3.4. A right-handed Cartesian coordinate system is used with origin located at the nominal interaction point. The z-axis is aligned with the 'downstream' direction, towards the end of the detector whereas the y-axis points vertically upwards. The different subdetectors that form the LHCb detector are also shown in Figure 3.4 and will be discussed in depth in the following sections.

Information presented in this section, to describe the various subdetectors of LHCb, has been mainly from Ref. [74]. Information on the performance of various subdetectors has been adapted from Ref. [75].

3.3 Tracking

The aim of the tracking system is to reconstruct precisely the trajectories of charged particles (tracks) using the information of the positions at which these particles have interacted with the sensors of the tracking station detectors (hits) and a dipole magnet that provides the measurement of the track momentum.

trigger conditions is defined as a fraction of the total inelastic pp interactions per bunch-crossing (ν) *i.e.* $\mu_{\rm vis} = \epsilon\nu$. Here ϵ is the fraction of inelastic events that satisfy the trigger condition.

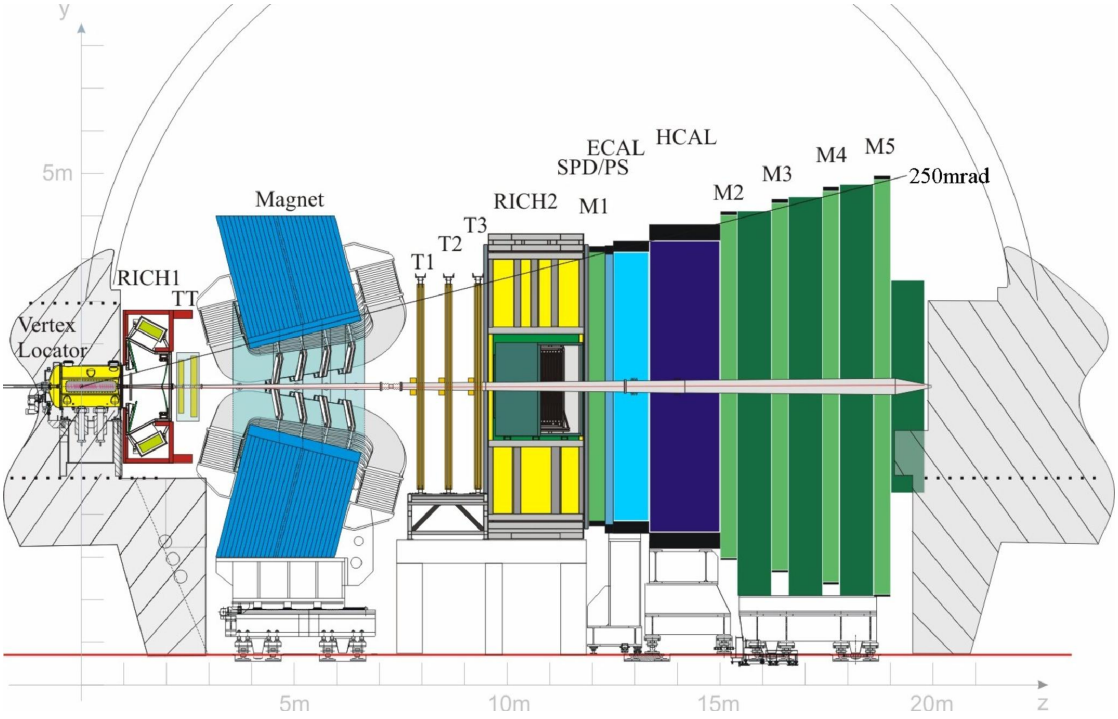

Figure 3.4: A schematic representation of the LHCb detector in the yz (non-bending) plane showing the positions of the subdetectors [74].

Information about the dipole magnet is presented in Sec. 3.3.1. The tracking system consists of the VErtex LOcator (VELO), Silicon Trackers (ST) and Outer Tracker (OT). These tracking stations are described in Secs. 3.3.2- 3.3.4. Finally in Sec. 3.3.5, I will discuss the performance of the tracking system.

The ST comprises of Tracker Turicensis (TT) and Inner Tracker (IT) since both consist of silicon microstrip detectors. As seen in Figure 3.4, there are regions labelled as IT and OT which make up the detectors T1–T3 (I will refer to them collectively as T stations). This distinction has been clarified in Figure 3.5.

Tracks are reconstructed from hits in the VELO, TT and T stations. At LHCb, the tracks are classified by which subdetectors they pass through, as illustrated in Figure 3.6. The analyses presented in this thesis use only 'long' tracks, defined as those which traverse all the tracking stations. The reconstruction of long tracks starts with a search for straight line trajectories among VELO hits.[3] Two complementary algorithms, 'forward tracking' and 'track matching' algorithms,

[3]The 'VELO tracks' must have hits in at least three R and three ϕ sensors. See Subsec. 3.3.2 for more information on the VELO sensors.

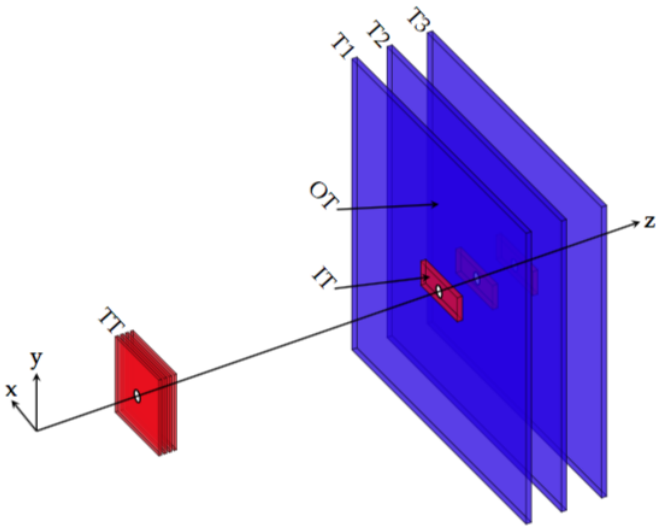

Figure 3.5: The layout of the (red) ST and (blue) OT around the beamline. The TT is to the left, while the T1 - T3 stations are on the right [76].

are then run to add information from the TT and T stations. The candidate tracks from both algorithms are combined removing any duplicates. The final step of track reconstruction is to fit the tracks with a Kalman filter. This takes into account multiple scattering and corrects for energy loss due to ionisation. Track quality is determined from the χ^2 per degree of freedom of the fit. It is possible to reconstruct 'Ghost' tracks, which do not correspond to the trajectory of a charged particle. Such tracks mostly result from incorrect matching of hits in VELO, TT and T detectors. Multivariate techniques are employed using information on track kinematics, the number of hits in the tracking stations and the result of the track fit to reduce the contribution of Ghost tracks. As a result, to control background contributions, each track is also assigned a 'Track_GHOSTPROB' which gives the probability of the track to be a Ghost track.

3.3.1 Dipole magnet

A dipolar magnetic field is used to deflect the charged tracks in the detector. Measuring the amount of deflection, *i.e.* the sagitta of the track trajectory (s) [77], provides information about the momentum of the particle. The momentum of the particle can be deduced by equating the centripetal ($\vec{F_c} = -\frac{\gamma_r m|\vec{v}|^2 \sin^2\theta_1 \hat{r}}{|\vec{r}|}$) and the Lorentz ($\vec{F_l} = -q|\vec{v}||\vec{B}|\sin\theta_2\hat{r}$) forces experienced

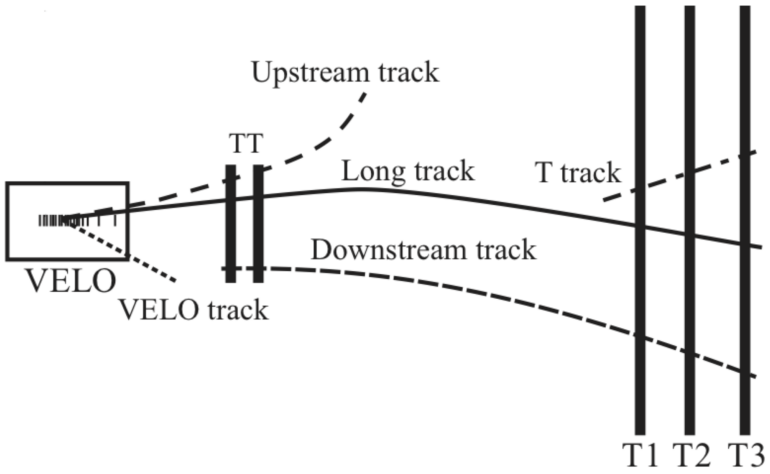

Figure 3.6: Track types at LHCb, which are defined by the subdetectors they pass through [75].

by a particle carrying a charge q moving with a linear velocity $\vec{v}$ in a magnetic field $\vec{B}$ [77]

$$|\vec{p}| = \frac{q\,|\vec{B}|\,|\vec{r}|\,\sin\theta_2}{\sin^2\theta_1}\,. \tag{3.3}$$

Here $|\vec{p}| = \gamma_r m|\vec{v}|$ is the linear momentum of the particle; $\vec{r}$ is the radius of the arc created by the charged particle in the magnetic field that is related to the sagitta of the track through the relation $|\vec{r}| = \frac{L^2}{8s}$ at leading order with L defined as the length of the straight line between the end points of the arc [77]; θ_1 is the angle between $\vec{r}$ and $\vec{v}$; θ_2 is the angle between $\vec{B}$ and $\vec{v}$; and γ_r is the Lorentz factor.

A momentum resolution of $\delta|\vec{p}|/|\vec{p}| = 4 \times 10^{-3}$ for particles with momentum of 10 GeV/c is achieved by applying an integrated magnetic field of 4 Tm over a 10 m track length between the VELO and the T stations. A warm dipole magnet is the source of this magnetic field inside the LHCb detector. This warm magnet was chosen over a super-conducting magnet due to economic constraints. The magnetic field is vertically oriented (in the y-direction), and covers $\pm 250\,\mathrm{mrad}$ vertically and $\pm 300\,\mathrm{mrad}$ horizontally. The magnetic field strength is required to be small outside the magnet so that it does not affect the operation the RICH Hybrid Photon Detectors (see Sec. 3.4.1 for more details).

Figure 3.7, on the left, shows the schematic diagram of the dipole magnet. The magnet is composed of two separate aluminium coils, shaped like a saddle, attached to a rectangular yoke with slanted poles, with a wedge-shaped window. The yoke is made of twenty-seven layers of laminated low-carbon steel, each $100\,\mathrm{mm}$ thick and with a total mass of 25 tonnes. Each coil is made of fifteen layers of hollow aluminium conductor with a central channel for water-cooling.

The nominal current passing through the coils is 5.85 kA.

To achieve the required momentum resolution, the magnetic field is measured to a relative precision of $\mathcal{O}(10^{-4})$ prior to data-taking. From this the position of the B field peak is known to within a few millimetres. This was achieved using arrays of Hall probes, with which the components of the field were measured in a fine grid spanning from the interaction point to the RICH2 detector. In order to eliminate artificial CP asymmetries induced by the detector, the polarity of the magnet is reversed regularly during the operational year, so that the collected data is split evenly between the two polarities. The data sample collected when the magnet field is pointing in the positive y-direction is referred to as 'MagUp' and the one collected when the field is pointing in the negative y-direction is referred to as 'MagDown'. The magnetic field as a function of z-position can be seen in Figure 3.7 (right) for both magnet polarities.

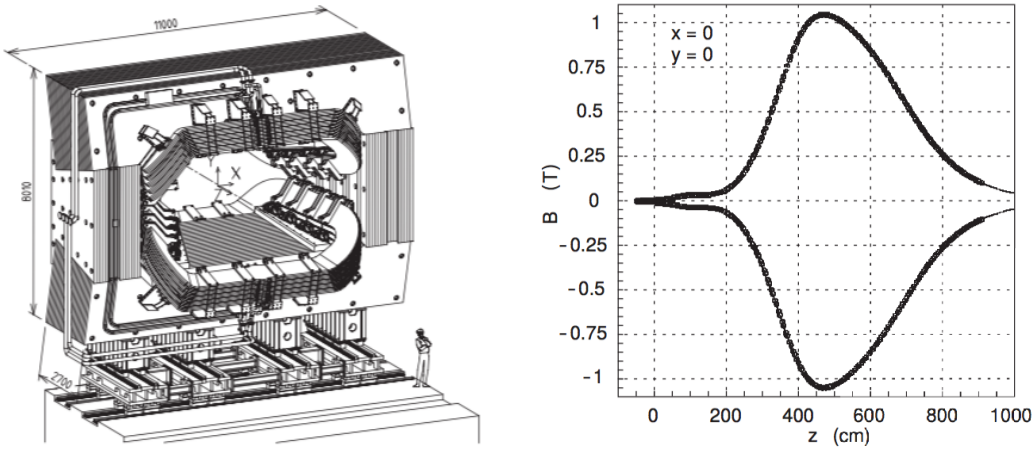

Figure 3.7: (left) A schematic representation of the magnet [74]. (right) Magnetic field as a function of z-position, at $x = 0\,\mathrm{cm}$ and $y = 0\,\mathrm{cm}$ for both magnet polarities (MagDown and MagUp) [74].

3.3.2 Vertex locator

The VErtex LOcater (VELO) immediately surrounds the interaction point and is used to obtain precise measurements of track coordinates near the interaction region. This is crucial for reconstructing the production and decay vertices of b- and c-hadrons, and for measuring the impact parameter (the distance-of-closest approach of a track to a vertex) of particles.

The VELO sub-detector is built in two halves (A and C side) staggered around the beam line between $-0.2 < z < 0.8$ m. Each half contains 21 modules composed of two semi-circular silicon strip sensors, namely R (radial) and ϕ (azimuthal) sensors. The modules are placed

perpendicular to the beam line. There is a 3 cm irregular spacing between each of the modules which has been optimised to allow each track in the detector acceptance to traverse at least four modules. The modules are densely packed around the interaction point to reduce the extrapolation distance from the first measured hit to the vertex. In addition, there are also four R sensors which have been placed upstream of the interaction point, that form the pile-up system, to detect when multiple proton–proton collisions have occurred. Moreover, all the VELO modules are placed inside an aluminium box in order to prevent RF pickup from the LHC beam; the side of the box is 300 µm thick to reduce the material budget, therefore the VELO operates in a secondary vacuum. An overview of the VELO system can be seen in Figure 3.8.

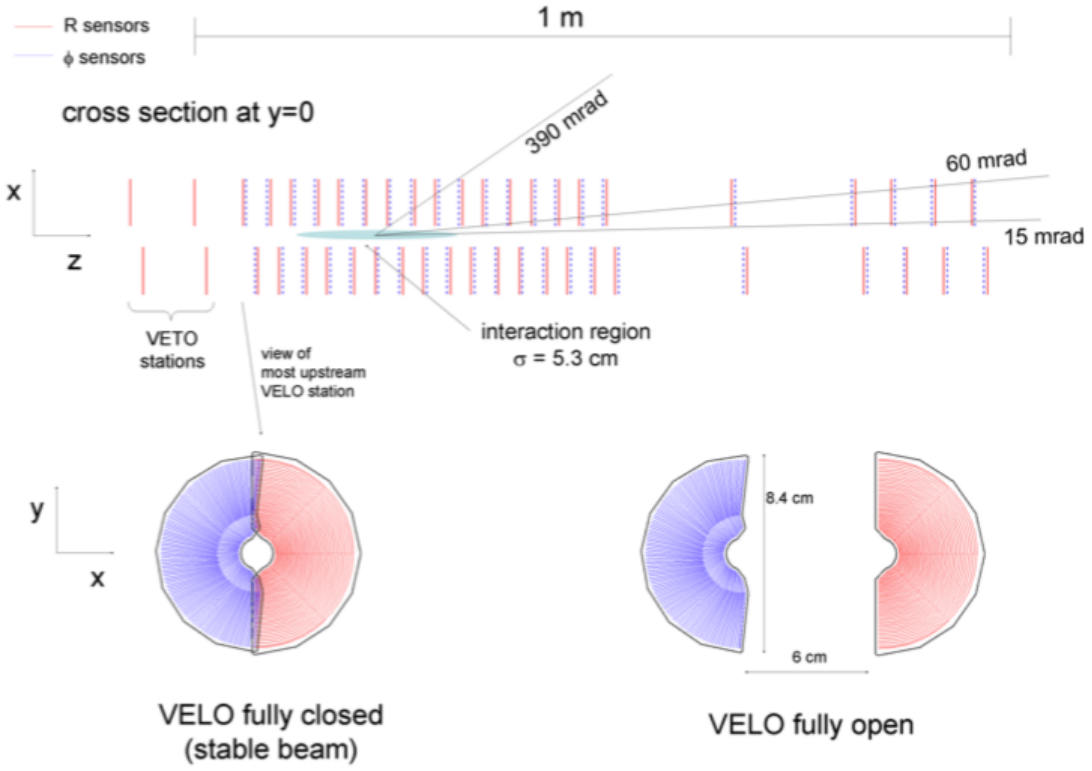

Figure 3.8: An overview of the VELO system. (top) View of the distribution of modules along the beamline within the VELO. Front views of a pair of modules in the (bottom left) closed and (bottom right) open VELO configurations [74].

During data-taking, the active area of the VELO starts at 8 mm from the nominal beam position, while during the beam injection the two halves are retracted by 29 mm. This is done to prevent damage whilst the beam profile is large during injection and ramping. When stable

beam is declared the VELO is closed around the real position of the beam which can change from fill to fill. The corresponding modules in the two halves overlap by $\sim 1.5\,\mathrm{mm}$ to cover the full azimuthal acceptance and to help with the alignment. The strip geometry of a VELO sensor module is shown in Figure 3.9.

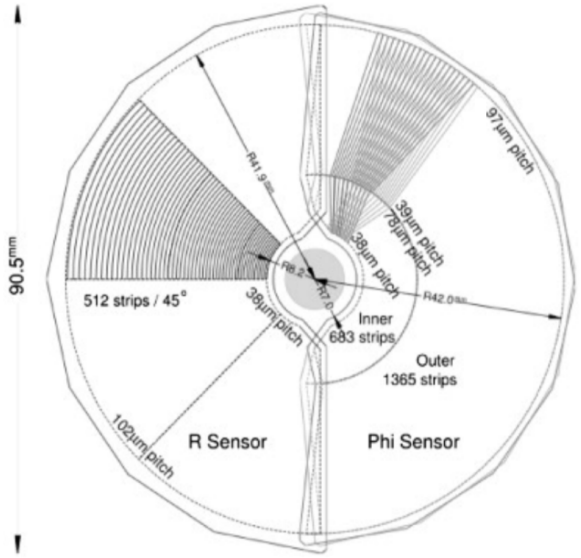

Figure 3.9: The strip geometry of R and ϕ VELO sensors [74].

To reduce the occupancy each R-sensor is divided into four sectors where each sector is 45° wide and accommodates 512 circular silicon strips. The inter-strip distance, or strip pitch, increases linearly from 38 to $102\,\mu\mathrm{m}$ to keep the strip occupancy approximately constant. On the ϕ-sensors there are 683 short strips at small radii with a pitch of 38–78 $\mu\mathrm{m}$ and covering the remaining area to the outside edge of the sensor are 1365 longer strips with a pitch of 39 to $97\,\mu\mathrm{m}$. These two radial sections are defined on the ϕ-sensors to reduce occupancy and to prevent too large a strip pitch at the outer edge. The strips on the ϕ-sensors are not perfectly radial but are inclined by a so-called stereo-angle of 10° (20°) for the inner (outer) part to improve the pattern recognition capability.

The VELO has managed to achieve an impressive resolution on the measurement of the impact parameter for charged particles of $(15 + 29/p_\mathrm{T})\,\mu\mathrm{m}$, where p_T is the component of the momentum transverse to the beam measured in GeV/c.

The LHCb experiment is planned to undergo a phase I upgrade during the period 2019 to 2020, which will allow operation at luminosities of $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, *i.e.* five times the current operational luminosity. In this period the R-ϕ sensors of the current VELO sub-detector will be

replaced by hybrid pixels to deal with the high occupancy. I have assisted with the implementation of the material description for the upgrade of the VELO sub-detector. The work undertaken is highlighted in appendix A. This work has formed part of the review process to choose one of the designs as the VELO module support system for the upgrade.

3.3.3 Silicon tracker

The Silicon Tracker (ST) consists of the Tracker Turicensis, located upstream from the magnet, and the Inner Tracker, located downstream from the magnet. The TT covers the full acceptance of the detector, while the IT covers a $120\,\mathrm{cm}$ wide and $40\,\mathrm{cm}$ high cross-shaped area in the center of each of the three downstream tracking stations.

3.3.3.1 Tracker turicensis

Long-lived particles such as K_{S}^{0} and Λ hadrons have a low reconstruction efficiency compared to short-lived particles as they tend to decay outside the VELO subdetector. Therefore tracking information provided by the TT helps immensely in the reconstruction of tracks originating from these candidates. The TT also helps in the reconstruction of low momentum tracks which will be bent out of the acceptance by the magnet.

The TT consists of four layers of silicon microstrip sensors with a pitch of $183\,\mu\mathrm{m}$ between the strips. The first and last layers are oriented vertically, and the second and third are rotated by -5° and 5° from the vertical. A single layer of the TT is composed of modules, which are columns of 14 silicon sensors. These sensors are placed in an electrically and thermally insulated light-tight box. The temperature inside the box is maintained below 5°C, and nitrogen gas is continuously flushed through the box to prevent condensation. To aid reconstruction, the four detector layers are arranged in pairs, spaced approximately $27\,\mathrm{cm}$ apart in z. The four layers are shown in Figure 3.10. Adjacent vertical modules within a layer are staggered by approximately $1\,\mathrm{cm}$ in z and a few millimetres in x to ensure coverage. The readout boards, structural supports and cooling system are located at the ends of the modules, outside the acceptance of the detector.

3.3.3.2 Inner tracker

The inner part of each of the three T stations, which is the region where the track multiplicity is highest, houses the Inner Tracker (IT). The IT therefore helps in reconstructing tracks that have passed through the magnetic field and lie near the beam axis. Each IT station, like those of the TT, consists of four layers of silicon microstrip sensors rotated by (0°, -5°, 5°,0°) about the vertical z-axis. The pitch of the strips is $196\,\mu\mathrm{m}$. These sensors are housed in light-tight boxes, four for each station, surrounding the beam pipe as shown in Figure 3.11. The boxes either side of the

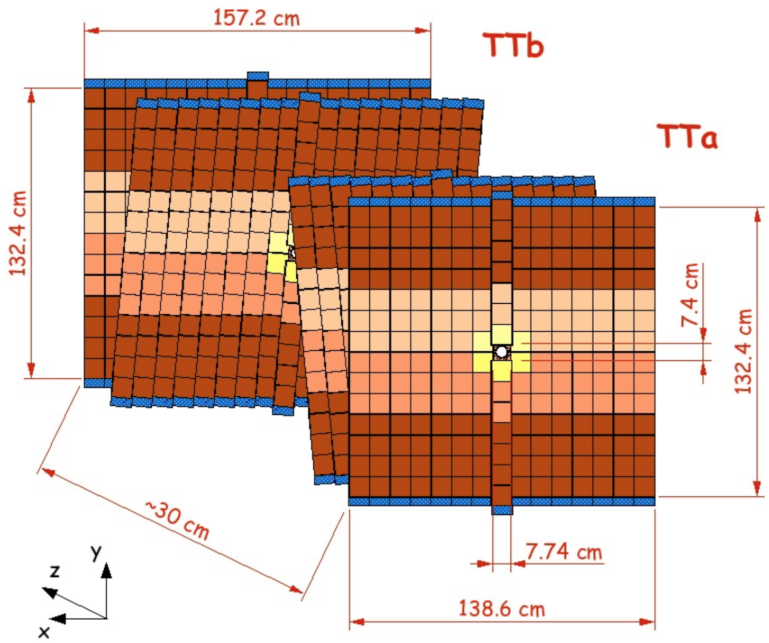

Figure 3.10: The four layers of silicon strip sensors in the TT [74].

beam pipe contain modules of two silicon sensors each. The ones above and below the beam pipe contain single-sensor modules. The boxes are staggered in z by $4\,\mathrm{mm}$ and overlap in x by $3\,\mathrm{mm}$ to ensure coverage and facilitate the relative alignment of the modules. These boxes, like those of the TT, are also maintained below $5°$C, and nitrogen gas is continuously flushed through to prevent condensation. Unlike the TT, the readout boards, cooling system and structural supports of the IT are inside the acceptance of the detector.

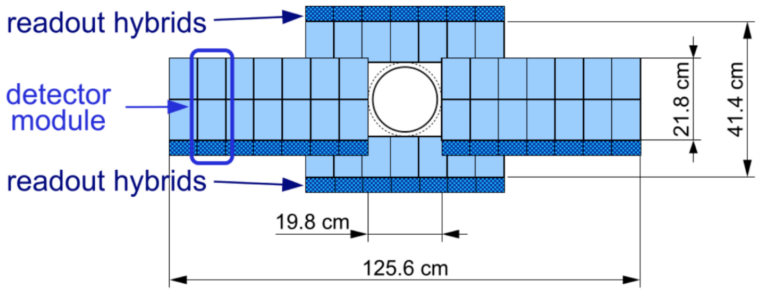

Figure 3.11: A schematic view of the IT around the beam pipe. The read out hybrids are shown in blue [74].

3.3.4 Outer tracker

The Outer Tracker (OT) is a drift-time detector, for the tracking of charged particles and the measurement of their momentum over a large acceptance area. A schematic view of the OT is presented in Figure 3.12.

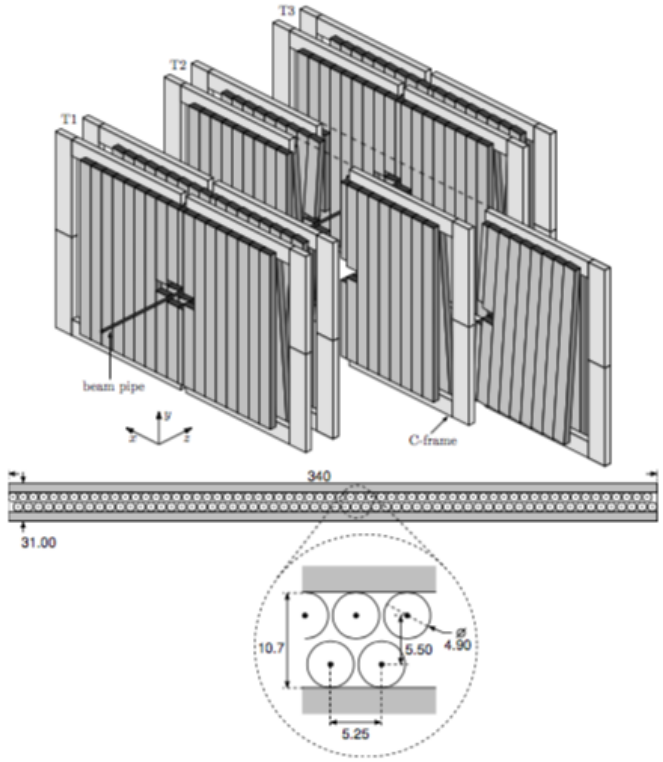

Figure 3.12: (Top) A view of the OT stations, with part of T2 retracted, showing the orientations of the straw tube modules. (Bottom) Cross-section of an OT module packed with straw tubes [74].

The OT surrounds the IT and hence covers the rest of the acceptance not taken up by the IT. There are four OT modules in each T station. These modules have the same orientation as the strips in other tracking stations (IT and TT): the first and last are along the y-axis, and the second and third are rotated by -5° and $+5^\circ$, respectively, about the z-direction. Each module consists of two staggered layers of 64 'straw' drift tubes. These tubes have an inner diameter of $4.9\,\mathrm{mm}$ and contain a mixture of 70% Ar, 28.5% CO_2, and 1.5% O_2 in order ensure a minimum drift-time across the tubes of less than 50 ns and drift-coordinate resolution of $200\,\mu\mathrm{m}$. Each tube is formed of two layers of thin foil. The outer layer is a laminate of polyamide and aluminium, which provides gas-tightness and shielding. The inner layer is carbon-doped polyamide and acts as the cathode. A gold-plated tungsten anode wire runs through the middle of each tube.

When a charged particle traverses the drift tubes, it ionises the gas contained within the tubes. The gas ions drift to the anode wires. Measurement of the time taken by the gas ions to reach the anode upon application of a current gives a measure of the relative position of the charged particle within the tube.

3.3.5 Vertexing and tracking performance

I will discuss in this section the performance achieved by the tracking system. Please note that the performance plots in this section are sometimes only shown for 2012 ($\sqrt{s}$ = 8 TeV) and 2015 ($\sqrt{s}$ = 13 TeV) data-taking periods only. Performance plots corresponding to 2011 ($\sqrt{s}$ = 7 TeV) are similar to 2012 plots and the plots corresponding to 2016 ($\sqrt{s}$ = 13 TeV) are similar to those of 2015 data-taking period.

Information from the VELO enables the reconstruction of the primary vertices (PV) and secondary vertices (SV) to allow for precise determination of particle decay-time and impact parameter (IP). I show in Figure 3.13, the vertexing performance achieved during Run 1 and Run 2 periods. It can be seen here that PV with 25 tracks has a resolution of 77 μm. At asymptotically high p_{T} the IP resolution along the x-axis is around 13 μm. A typical decay time resolution in LHCb for a four track vertex in around 45 fs.

The tracking efficiency for various types of tracks (as classified in Figure 3.6) has been studied, however, here I will discuss the performance of long tracks only. I show in Figure 3.14, for 2012 and 2015 data-taking periods, the tracking efficiency of long tracks as a function of their momentum which is studied using a 'tag-and-probe' method with $J/\psi \to \mu^+\mu^-$ decays. One of the daughter muons (the 'tag') is fully reconstructed, whereas the other (the 'probe') is only partially reconstructed (*i.e.* only the track segment from the muon system). The tracking efficiency is calculated as the proportion of probe muons which can be matched successfully to fully reconstructed long tracks. It can be seen here that the average efficiency is above 96 % in the momentum range $5 < p < 200$ GeV/c for "reconstructible" tracks in the LHCb acceptance. Note that the tracking efficiency does not only depend on track momentum but also on pseudorapidity, track multiplicity and the number of primary vertices [75].

I show in Figure 3.15 on the left, the relative momentum resolution achieved at LHCb for muon tracks in the decay of $J/\psi \to \mu^-\mu^+$ as a function of track momentum (Note also that the muon tracks here also include information from Muon detector which is described in Subsec. 3.4.3). The relative momentum resolution varies from 0.5% at low momentum to 1.0% at a momentum of 200 GeV/c. Measured tracks momenta are used to calculate invariant mass. Therefore, the mass resolution achieved is strongly correlated with the momentum resolution. I also show in Figure 3.15 on the right, the relative mass resolution achieved for six $\mu^-\mu^+$

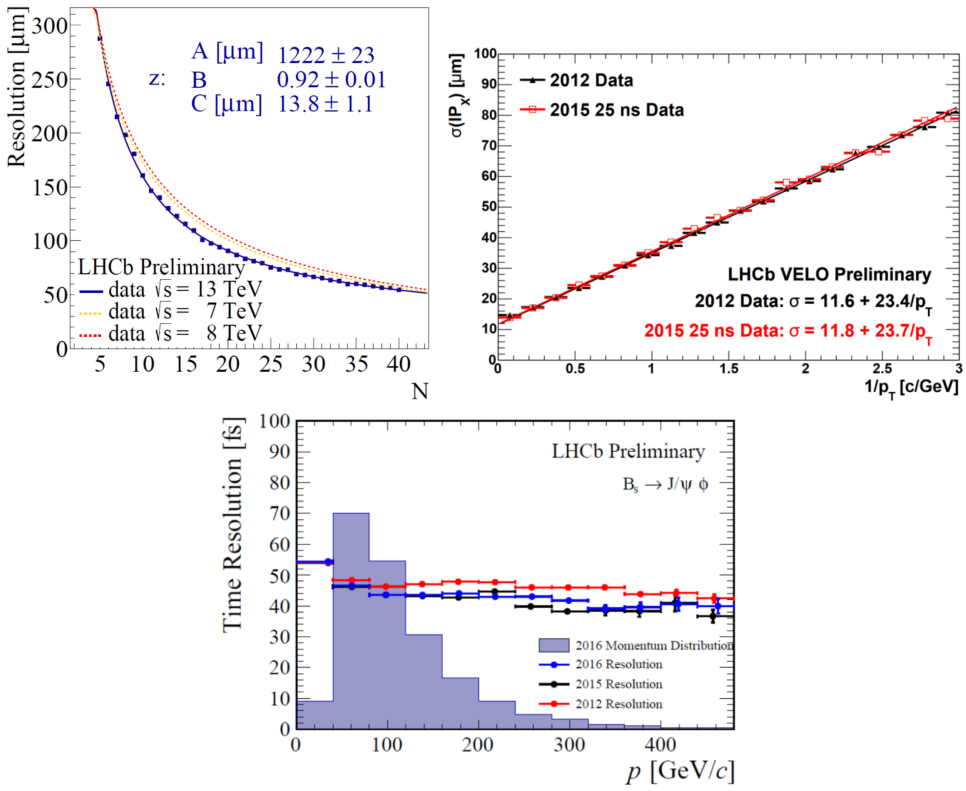

Figure 3.13: (Top left) Primary vertex resolution for events as a function of track multiplicity in the z-direction for 2011 ($\sqrt{s}$ = 7 TeV), 2012 ($\sqrt{s}$ = 8 TeV) and Run II ($\sqrt{s}$ = 13 TeV) years. (Top right) Resolution of the x-direction component of track IP, as a function of $1/p_{\mathrm{T}}$ for 2012 and 2015 data-taking years [78]. (Bottom) The resolution of decay time of B_s^0 candidate reconstructed using $J/\psi \to \mu^+\mu^-$ and $\phi \to K^-K^+$ candidates as a function of B_s^0 momentum for 2012, 2015 and 2016 data-taking years [79].

resonances (J/ψ, $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$ and Z) as a function of $m(\mu^-\mu^+)$ invariant mass.

3.4 Particle identification

The signal and background events that differ only by the species of particles in the final state can be differentiate using accurate particle identification (PID) information. At LHCb two Ring-Imaging Cherenkov (RICH) detectors (see Sec. 3.4.1), RICH1 and RICH2, are used to distinguish the species of long-lived charged hadrons: namely pions (π), kaons (K) and protons (p). The calorimeter system (see Sec. 3.4.2) is used to identify and measure the energy of photons, electrons and hadrons. The muon detector system (see Sec. 3.4.3) is used to identify and measure

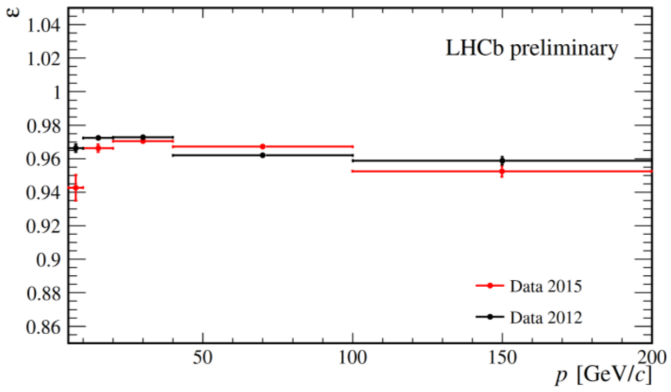

Figure 3.14: Tracking efficiency as a function of momentum for 2012 and 2015 data-taking periods [79]. The tracking efficiency as a function of pseudorapidity, track multiplicity and number of primary vertices can be seen in Ref. [75].

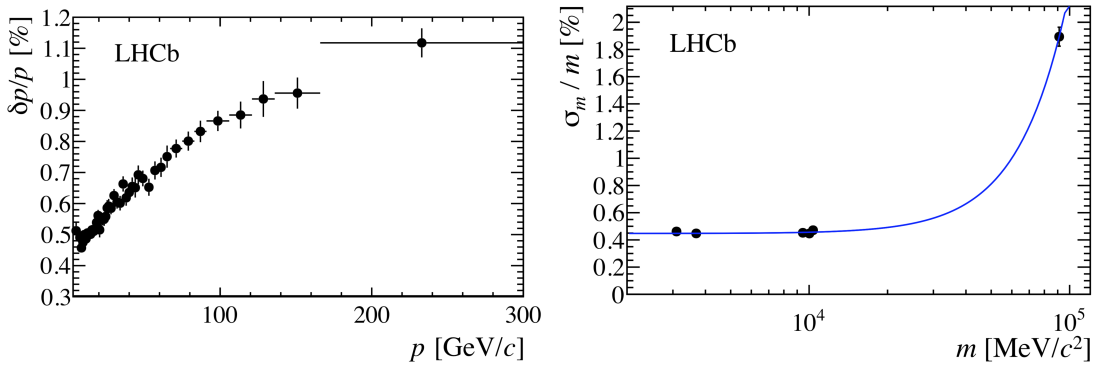

Figure 3.15: (Left) Relative momentum resolution of muon tracks from $J/\psi \to \mu^-\mu^+$ decays as a function of track momentum [75]. (Right) Relative mass resolution for six $\mu^-\mu^+$ resonances (J/ψ, $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$ and Z) as a function of $m(\mu^-\mu^+)$ invariant mass. The solid curve is an empirical power-law fit through the points. [75].

the momentum of muons. In Sec. 3.4.4, I will discuss the combined PID performance.

Using information obtained from the RICH, calorimeter and muon detector systems, a likelihood can be assigned to a particle for being a particular species. For example, the likelihood for a charged particle to be a proton is given by $\mathcal{L}_p = \mathcal{L}_p^{\text{RICH}} \times \mathcal{L}_{\text{hadron}}^{\text{CALO}} \times \mathcal{L}_{\text{not}\mu}^{\text{MUON}}$. The PID requirement would then be a requirement on the difference in the logarithm of two such likelihood values *e.g.* $\Delta \log \mathcal{L}(p - \pi) = \log \mathcal{L}_p - \log \mathcal{L}_\pi$. The analyses presented in this thesis, however, use 'ProbNN' variables as PID variables. These variables represent the Bayesian probability for a particular particle hypothesis and are an output of a trained neural network [80]. A dedicated neural network is used in the training for each particle hypothesis (K, π, p, e, μ). The information

from the tracking stations and the subdetectors involved in PID, in particular the RICH detectors, is used to train these neural networks. The difference in the log likelihood variables, defined above, also forms an important input to these neural networks.

3.4.1 Ring imaging cherenkov detectors

The two LHCb RICH detector systems (RICH1 and RICH2) provide information about the probable species of charged particles over a wide range of momenta. The two RICH systems detect Cherenkov photons that are emitted in the forward direction with an angle θ by a charged particle when it traverses a radiator of known refractive index (n) with a velocity (v) greater than the velocity of light in that medium (phase velocity). The emitted Cherenkov photons are collected and focused onto planes of hybrid photodetectors (HPDs) outside the LHCb acceptance using a combination of spherical and flat mirrors. These HPDs, which are designed to detect photons with a wavelength range of 200-600 nm, are surrounded by external iron shields and placed in MuMetal cylinders to shield them from magnetic fields up to $50\mathrm{mT}$. The Cherenkov photons detected by the photodetectors form a ring of radius r. The measured value of this radius along with the known focal length of the spherical mirror, f, are used to calculate θ using the relation: $\tan(\theta) = r/f$. Using the measured value of θ for a particle and the value of the particle's momentum, $\vec{p}$, determined using the tracking system, one can calculate the mass of the particle using the relation:

$$\cos(\theta) = \frac{1}{n\beta} = \frac{E}{n|\vec{p}|} = \frac{\sqrt{|\vec{p}|^2 + m^2}}{n|\vec{p}|} \tag{3.4}$$

A schematic view of how a charged particle creates a light cone with a cherenkov angle can be seen in Fig. 3.16.

In this way, different charged particle species can be differentiated using the RICH detectors. The sensitivity to particles of different masses and momenta is made possible by using radiators with different refractive indices. Figure 3.17 shows the dependence of θ on $|\vec{p}|$ for different charged particle species in different RICH radiators. Note that for a particle with velocity less than the phase velocity of light in the medium (*i.e.* $\beta < 1/n$), no Cherenkov light will be emitted and for a particle with high velocity (*i.e.* $\beta \to 1 \implies \cos(\theta) \to 1/n$) the discrimination power is lost.

3.4.1.1 RICH1

RICH1 is located between the VELO and TT (See Fig. 3.4) and covers the full angular acceptance of 25–$300\,\mathrm{mrad}$. It offers PID for low momentum particles in the range 2–$40\,\mathrm{GeV}/c$ using

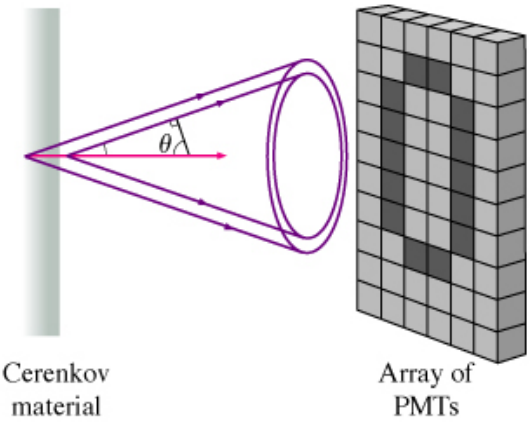

Figure 3.16: A charged particle passing through a medium faster than the local speed of light emitting Cherenkov radiation in a light cone with Cherenkov angle, θ, which is then focused on the PMTs [81].

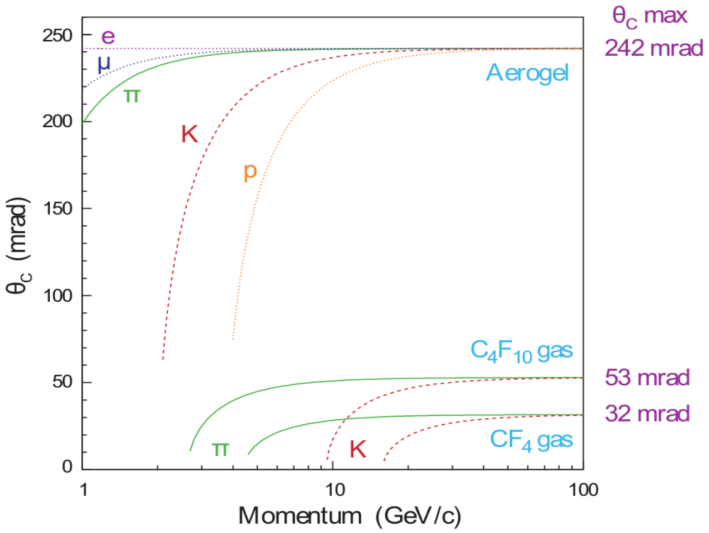

Figure 3.17: The Cherenkov angle dependence on particle momentum shown for the different RICH radiators [75]. RICH1 and RICH2 detectors provide PID information over the momentum ranges 10–50 GeV/c and 50–100 GeV/c, respectively.

aerogel and C_4F_{10} radiators that have refractive indices of 1.03 and 1.0014 for 400 nm light.

A side-view of the RICH1 subdetector is shown in Figure 3.18 (left). The C_4F_{10} gas, aerogel tiles and the mirror assembly sit in a sealed aluminium box. The detector contains $3.5\,\mathrm{m}^3$ of C_4F_{10} gas, and the silica aerogel forms a 50 mm thick wall on the upstream side of the aluminium box. There are four spherical mirrors arranged symmetrically around the beam

pipe that are located within the LHCb acceptance and hence are constructed from lightweight carbon-fibre. All other components of the optical system, including the two planes of flat mirrors, are located above and below the beam pipe, outside the particle acceptance.

During the Run I data-taking period, it was noticed that the aerogel did not improve low-momentum performance as much as predicted and therefore was removed prior to the Run II data-taking period. Removal of the aerogel results in no significant loss of performance of PID, speeds up reconstruction, and results in a larger C_4F_{10} volume.

3.4.1.2 RICH2

The RICH2 subdetector is located after the T1–T3 tracking stations to reduce the amount of material before the tracking stations. It provides PID information for particles in the high momentum region of 15–100 GeV/c and covers a smaller angular range 15–120 mrad. RICH2 uses a $95\,\mathrm{m}^3$ volume of CF_4 gas for its radiator.

A top-view of the RICH2 subdetector is shown in Figure 3.18 on the right. The CF_4 gas is contained within a sealed box with the mirror array. Within RICH2 there are 52 spherical primary mirrors which focus Cherenkov light onto two planes of flat mirrors to the left and right of the beam pipe.

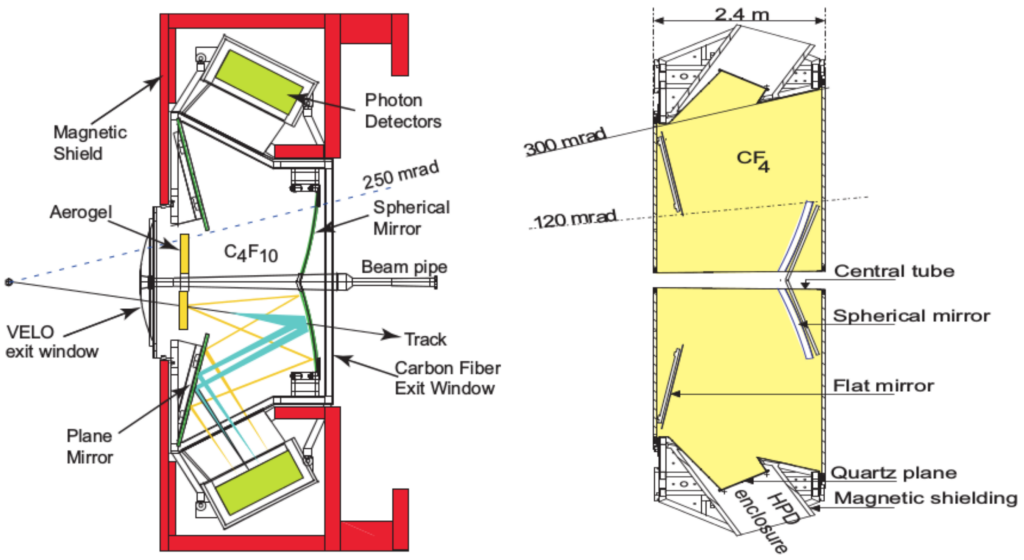

Figure 3.18: (Left) Cross-section of the RICH1 detector in the y-z plane [75]. (Right) Cross-section of the RICH2 detector in the x-z plane [75].

3.4.2 Calorimeters

The calorimeter system at LHCb consists of four subdetectors: Scintillator Pad Detector (SPD), Pre-shower detector (PS), Electromagnetic calorimeter (ECAL) and Hadronic calorimeter (HCAL). These subdetectors are located between the first and second muon stations (described in Sec. 3.4.3). These calorimeters are essential to reconstruct decays with a photon or a neutral pion in the final state. They can also help reconstruct decays that feature high momentum electrons that radiate photons via bremsstrahlung. As mentioned before, information from the calorimeter system is also used for PID and is an important input for the first level of the trigger where the decision to keep or throw away an event must be made just 4 μs after a pp interaction.

When a particle traverses a calorimeter, it interacts with the calorimeter material and loses energy. This interaction produces a cascade shower of particles which undergo further interactions to produce scintillation light. Scintillation light from the calorimeters is transferred to multianode photomultiplier tubes (PMTs) using wavelength-shifting (WLS) fibres. At the end, the photodetectors convert light into an electrical signal, which is used to infer the energy deposited by the original particle. All the four calorimeters at LHCb use absorber material with a large radiation length to stop particles in the detector and measure the energy deposited.

3.4.2.1 SPD, PS and ECAL

The SPD and PS are two scintillator pad detectors which are placed before the ECAL and are separated by 15 mm thickness of lead. The dimensions of the SPD are about 0.45% smaller than those of the PS to maintain the same angular acceptance. The SPD determines the electric charge of a particle before it showers, allowing electrons to be differentiated from high energy γ and $\pi^0 \to \gamma\gamma$ backgrounds. The PS detector is used to distinguish electrons, which are more likely to shower in the lead absorber, from charged hadrons.

The ECAL, on the other hand, is a sampling calorimeter with 66 alternating layers of lead absorber and polystyrene scintillator with thicknesses of 2 mm and 4 mm, respectively. The layers are arranged perpendicular to the beam line with total thickness of 42 cm. This thickness ensures showers from high energy electrons and photons are fully contained within the calorimeter. The granularity of the ECAL is somewhat coarser than that of the SPD and PS. The ECAL has a design energy resolution of $\sigma_E/E = 10\%/\sqrt{E/\,\mathrm{GeV}} \oplus 1\%$.

The PS, SPD, and ECAL are separated into three regions namely inner, middle and outer each with cell areas of approximately $4 \times 4\,\mathrm{cm}^2$, $6 \times 6\,\mathrm{cm}^2$ and $12 \times 12\,\mathrm{cm}^2$, respectively. This size has been designed based on the expected occupancy and shower size in each region. The layout of the scintillator pads in the PS, SPD and ECAL is shown in Figure 3.19 on the left. It can be seen here that the PS, SPD, and ECAL all have finer granularity near the beam line where

particle flux is higher, since the hit density varies by two orders of magnitude over the surface of the calorimeter.

3.4.2.2 HCAL

The HCAL subdetector uses iron and scintillating tiles as the absorber and active material, respectively. The scintillating tiles of the HCAL lie parallel to the beam axis unlike those of the ECAL. The HCAL depth is limited to about 1.6 m (5.6 nuclear interaction lengths) due to space constraints in the LHCb cavern. The HCAL is separated into two regions, namely inner and outer, unlike the other three calorimeter systems which are segmented into three regions. The read-out cells in the HCAL are larger than in the ECAL due to the differences in dimension and structure of hadronic and EM showers.[4] A schematic of the HCAL scintillating pad regions is shown in Figure 3.19 on the right. Like the ECAL, finer granularity can also be seen for the HCAL near the beam line. The HCAL has a design energy resolution of $\sigma_E/E = 69\%/\sqrt{E/\,\mathrm{GeV}} \oplus 0.9\%$.

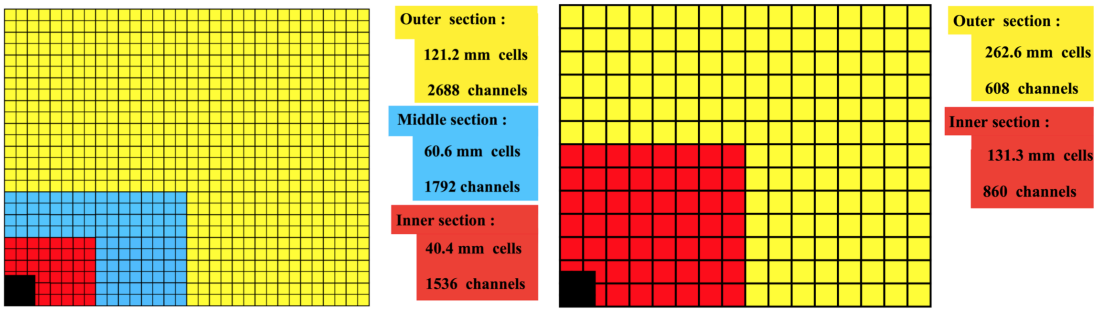

Figure 3.19: (Left) Schematic of SPD, PS and ECAL scintillating pads within a layer. The beam pipe hole is shown in black [74]. (Right) Schematic of the HCAL scintillating pad regions [74].

3.4.3 Muon chamber

The detection and identification of muons at LHCb is crucial in the studies of leptonic and semi-leptonic b- and c-hadron decays. The information from the muon stations also gets utilised in the flavour tagging and trigger decision for events.

The LHCb muon system consists of five stations (M1–M5) orientated perpendicular to the beam line. The schematic view of the muon system is shown in Figure 3.20. All the five muon stations have an acceptance of 20–306 mrad in the horizontal direction and 16–258 mrad in the

[4]Hadrons interact with the HCAL via long-range nuclear interactions creating a wider shower structure than that produced by electrons or photons in the ECAL which interact via electromagnetism.

vertical direction. To give a better p_{T} estimate for the muon candidates before they pass through the dense calorimeter material, the first muon station (M1) is located upstream of the calorimeter system. The remaining four layers are all downstream of the calorimeters. The M2–M4 stations are separated by $80\,\mathrm{cm}$ thick iron blocks to absorb any hadronic backgrounds surviving past the HCAL. Only muons with momentum greater than $6\,\mathrm{GeV}/c$ will be able to penetrate through all the muon stations. A good p_{T} resolution of $\sim 20\%$ in the horizontal plane is offered by stations M1–M3. The final two stations M4 and M5 merely confirm that the candidate has penetrated all of the iron absorber. A muon candidate is required to have hits in all five layers of the detector.

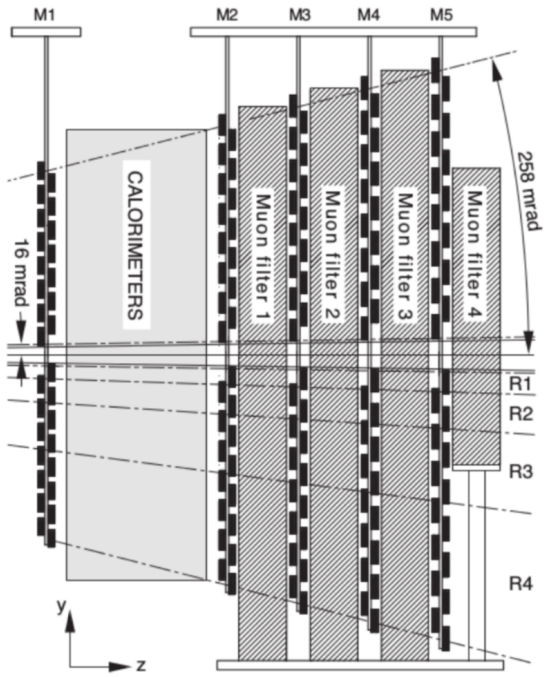

Figure 3.20: Schematic view of the LHCb muon detector [74].

Each muon station consists of chambers that contain rectangular 'logical pads' of different dimensions. The layout of chambers and logical pads in M1 is shown in Figure 3.21. Smaller logical pads with finer segmentation are placed closer to the beam line to offer greater transverse momentum resolution. The pad dimensions scale in the ratio 1:2:4:8 with distance from the beam pipe to give comparable particle flux across each pad.

With the exception of the inner region of M1, the muon system is constructed from Multi-wire Proportional Chambers (MWPCs). The MWPCs have vertical anode wires with $5\,\mathrm{mm}$ gas gaps between cathode plates. The wires are 20–$30\,\mathrm{cm}$ long with a spacing of $2\,\mathrm{mm}$ in the x-direction. With MWPCs a time resolution of around 5 ns is achieved by a gas mixture of Ar,

CO_2 and CF_4 in a ratio of 40:55:5, respectively. The inner region of M1, the region that has the highest occupancy, is constructed from triple Gaseous-Electron-Multipliers (GEMs), which have a higher radiation tolerance. The twelve chambers in the R1 region of M1 each have two triple-GEM detectors from which a logical OR result is returned. A time resolution of around $3\,\mathrm{ns}$ is achieved by a gas mixture of Ar, CO_2 and CF_4 in a ratio of 45:15:40, respectively.

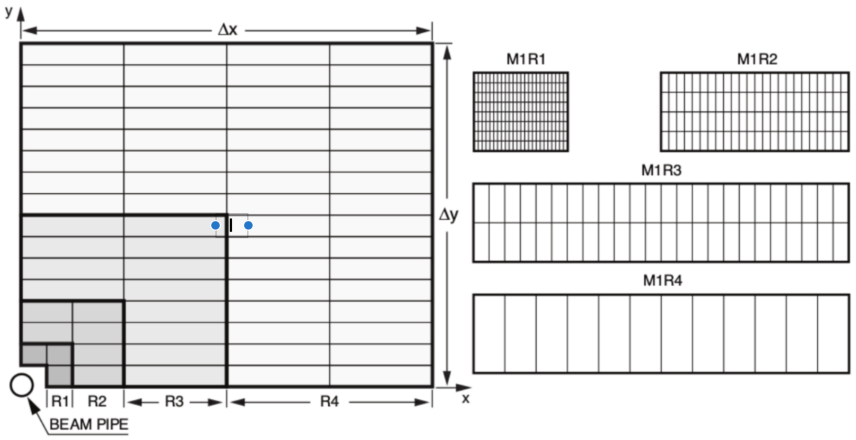

Figure 3.21: (Left) Division of a quarter of the M1 station into regions R1–R4 shown in different shades of grey. The rectangles denote individual chambers [74]. (Right) The division of each chamber into logical pads in M1. The number of pad rows per chamber is the same for all muon stations, however the number of pad columns per chamber varies between muon stations [74].

3.4.4 PID performance

The efficiency of K identification (ID) and $\pi \to K$ mis-ID as a function of track momentum for 2012 and 2015 data-taking periods and different requirements on the PID variables ($\Delta \log \mathcal{L}(K - \pi)$) can be seen in Figure 3.22. By comparing these two figures, it can be seen that the separation between pions and kaons has been improved in Run II, especially in the momentum region of 2–20 GeV/c due to the removal of the aerogel in the RICH subdetector [82]. The average efficiency of kaon ID achieved at LHCb is $\sim 95\%$ whereas the average $\pi \to K$ mis-identification (mis-ID) probability is $\sim 5\%$ [73]. The average efficiency of μ ID is $\sim 97\%$ whereas the $\pi \to \mu$ mis-ID is $\sim 1 - 3\%$ [73].

3.5 Trigger

The peak bunch crossing frequency of the LHC is $40\,\mathrm{MHz}$. However, as detailed earlier, there are gaps in the beam (*e.g.* abort gaps) that lead to an average crossing frequency of $30\,\mathrm{MHz}$

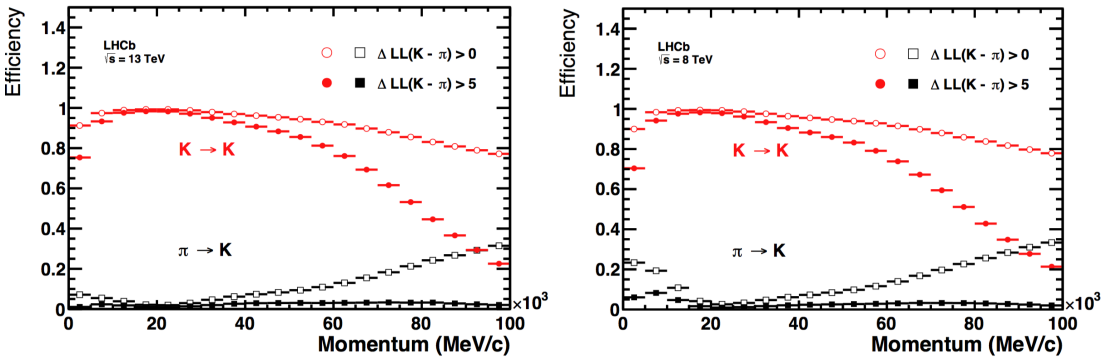

Figure 3.22: The efficiency of (red) K identification and (black) $\pi \to K$ mis-identification in samples of $D_s^+ \to K^+K^-\pi^+$ and $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ decays for the (left) 2012 and (right) 2015 data-taking periods. The distributions are shown as a function of track momentum for two different PID requirements: $\Delta \log \mathcal{L}(K-\pi) > 0$ (open shapes) and $\Delta \log \mathcal{L}(K-\pi) > 5$ (filled shapes). (Right) The same, for the 2015 data-taking period. The efficiency plots corresponding to 2011 data-taking period are similar to 2012 plots and the plots corresponding to 2016 are very similar to that of 2015 data-taking period [82].

seen by the LHCb detector. Within the LHCb acceptance, assuming that the cross-section of inelastic pp collisions is 60 mb at $\sqrt{s}$ = 14 TeV and the instantaneous luminosity is around $\mathcal{L} = 2 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$, the frequency of pp interactions[5] is 12 MHz. Due to constraints on the storage of the data, this rate needs to be reduced dramatically while still retaining events of interest. The role of the trigger is to do exactly this. It reduces the rate at which events are read out of the detector to the storage recording rate of 3.5–12.5 kHz [83]. The trigger efficiency for dimuon channels is around $\sim 90\%$ whereas for multi-body hadronic final states it is around $\sim 30\%$ [73].

The LHCb trigger consists of two levels: the hardware 'level zero' trigger (L0) and the software 'High Level Trigger' (HLT). The design of the trigger for 2011, 2012 and 2015-2017 data-taking years is compared in Figure 3.23. In this section I will discuss the design and structure of each different trigger level in detail. Note also that, in LHCb offline analysis, all the trigger decisions can be associated with the offline reconstructed signal candidate and hence can be classified into three categories with respect to that candidate: 'Trigger on Signal' (TOS), 'Trigger Independently of Signal' (TIS) and 'Trigger on Both' (TOB). TOS candidates are those which are triggered by the final state particles of the signal candidate under study. TIS candidates are those which are triggered not by the final state particles of the signal candidate but by other particles present in the same event as the signal. For b-hadron decays, this is usually from decay products

[5] A pp interaction is defined as visible if there are at least two charged particles, created from an interaction, that leave enough hits in the VELO and tracking stations to be reconstructed.

of the b-hadrons formed by the other b-quark in the $pp \to b\bar{b}X$ process. TOB candidates are those which require both the signal and the rest of the event in order to trigger.

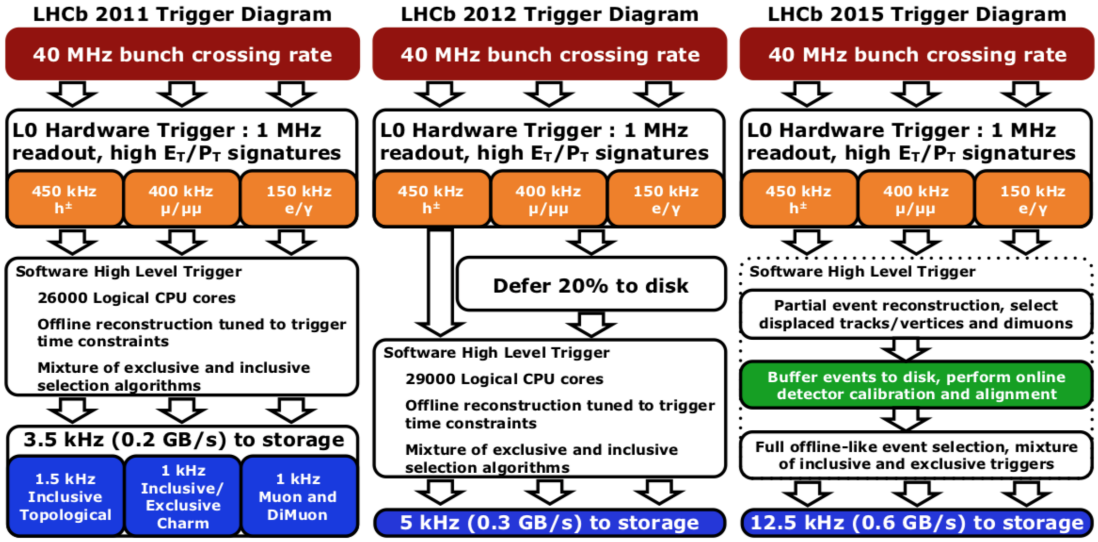

Figure 3.23: From left to right: HLT schemes in 2011 (no events deferred), 2012 (20% of L0 accepted events deferred for later processing by HLT) and 2015-2017 (all HLT1 accepted events deferred to be processed by HLT2) [83].

3.5.1 Hardware trigger

The LHCb hardware trigger (L0) uses custom electronics which is synchronised with the LHC clock. It reduces the rate which the subdetectors can be read out to 1 MHz. The L0 consists of two independent systems: the calorimeter trigger and the muon trigger. The information from these two triggers is used to make a final decision of whether or not to pass the full event to the HLT using the L0 decision unit. This decision unit allows for pre-scaling and the overlapping of several conditions. It has a budget of 2 μs to make a decision, after accounting for the latency of the electronics, particle flight time and cable lengths.

The calorimeter trigger identifies high E_{T} electrons (`L0Electron`), photons (`L0Photon`) and hadrons (`L0Hadron`). It forms clusters by adding the E_{T} of 2×2 cells in the ECAL and HCAL and selects the clusters with the highest E_{T}. Hadrons are selected from an HCAL, PS and ECAL clusters with $E_{\mathrm{T}} \gtrsim 3.5\,\mathrm{GeV}$. Electrons and photons are selected from ECAL and PS clusters with $E_{\mathrm{T}} \gtrsim 3\,\mathrm{GeV}$. Electrons are discriminated from photons by requiring that there be a matching hit in the SPD. A requirement that the total number of hits in the SPD is

less than 600 is also used to reduce contributions from high-multiplicity events that would take too long to process in the HLT.

The muon trigger looks for the two highest p_{T} muons in each quadrant of the cluster formed in the muon system. A stand-alone track reconstruction is performed by the muon system, using M1 and M2 stations to measure the muon p_{T}. Muon candidates are combination of hits that form a straight line through all five stations and are required to be consistent with originating from the interaction point. An event is triggered if either the highest p_{T} of any muon candidate in an event (`L0Muon`) is above 1.5–1.8 GeV/c with hits in SPD is less than 600 or the product of the highest and second-highest p_{T} in an event (`L0DiMuon`) is above $(1.3\text{–}1.6\,\mathrm{GeV}/c)^2$ with hits in SPD less than 900.

Note that during Run I and Run II data-taking periods, the same L0 hardware trigger system is used with different E_{T}, p_{T} and SPD hit thresholds.[6]

3.5.2 Software trigger

The HLT, unlike L0, runs asynchronously from the LHC clock. The HLT reduces the event rate from 1 MHz to an output chosen to match the available computing resources typically by making decisions based on the full event data. It is split into two stages: HLT1 and HLT2.

The HLT1 stage performs a partial event reconstruction wherein the primary vertices (PVs) are reconstructed from the VELO tracks, the tracks from VELO are also matched to tracks in the downstream tracking stations and a Kalman fit [84] is performed. There also requirements imposed on track fit χ^2, p_{T} of the track and impact parameter of the track. The HLT1 stage reduces the output rate to $\sim 50 - 150$ kHz. The analyses described in this thesis use the `Hlt1TrackAllL0` line.[7] This line selects hadron decays that are significantly displaced from a primary vertex by applying loose requirements on the number of hits in the tracking stations for each track, the magnitude of track impact parameter from the primary vertex, and the reconstructed secondary vertex quality, along with requirements on the track p and p_{T}.

At the HLT2 stage, a full reconstruction is performed on all events passing HLT1. Therefore, full event information is used to design HLT2 exclusive (required to select certain specific final states) and inclusive (generic topological requirements are imposed on the final state) trigger lines. The inclusive 'topological' trigger lines are used in the analyses presented in this thesis. These are designed to trigger on partially reconstructed b-hadron decays with at least two charged particles in the final state and a displaced decay vertex. Tracks are selected using fit quality, impact parameter and particle identification information. Vertices are constructed using com-

[6] Generally in Run II higher L0 thresholds are used.

[7] A set of conditions which, if satisfied, causes any level of the trigger to pass the event to the next stage, or write it to storage, is referred to as a 'trigger line'.

binations of two, three or four of the selected tracks. Signal candidates are selected based on several kinematic and isolation variables using a novel Bonsai Boosted Decision Tree (BBDT) algorithm [85], trained on simulated signal events and previously collected collision data.

During the 2011 data-taking period (see left of Figure 3.23), the HLT farm was idle during the inter-fill inactivity. To circumvent this waste of resources during the 2012 period (see center of Figure 3.23) $\sim 20\%$ of the L0-passed events were buffered to local disks and the HLT selection applied later to these events leading to a sizeable improvement of $\sim 25\%$ in the HLT processing rate. However, for long periods without beam, the HLT farm was still idle. It was decided for Run II that the HLT1 will be the only stage of the software trigger that will be operating with the LHC bunch crossing frequency (see right of Figure 3.23). Hence, HLT1 and HLT2 will be two asynchronous processes running on the same or different computing nodes. To achieve this all the events passing the HLT1 selection are buffered in the local disk and later processed by HLT2. Further, during Run I, HLT2 did a simplified event reconstruction, with preliminary alignment and calibration of the detector[8] and only used marginal information from the RICH subdetector. Due to this, the quality of reconstruction performed 'online' (during data-taking) by the HLT2 was of slightly lower quality when compared to 'offline' (after data being written to tape) reconstruction. It was decided in Run II that the buffered HLT1-passed samples will be used to perform real-time alignment and calibration of the detector before the HLT2 stage. The reconstruction performed by HLT2 now also uses full PID information. More offline computing resources and more efficient use of them have helped boost the HLT2 processing rate from 5 kHz in Run I to ~ 12.5 kHz in Run II.

The implementation of this separation of the software trigger in Run II also enabled a substantial change in the flow of event data. The flow of event data for Run II is shown in Figure 3.24. In the Full stream that existed during Run I, an offline reconstruction of the event is performed and the full event record *i.e.* including detector raw data (detector hits, *etc.*) is written to tape. After a few pre-selection requirements, referred to as 'Stripping' lines, that have been defined by the analyst and optimised to select events that might contain the decay mode under study, the data is written to disk in the `DST` format. The user then only needs to process the reconstructed candidates flagged as passing the pre-defined Stripping line. In Run I, the 'Full' and 'Calibration' streams existed while in Run II part of the 'Calibration' stream became the 'Turbo' stream [86]. In the Turbo stream, a compact event record (without the detector raw data) is written, in the `micro-DST` format, directly from the trigger and is prepared for physics analysis by a new application called 'Tesla' [86]. The by-product of the 'Turbo' stream is to

[8] The alignment and calibration during Run 1 was applied after data-taking. The alignment is performed for VELO, Tracking stations, RICH mirrors and Muon chamber. The calibration is performed on RICH refractive index, HPDs, Calorimeters, *etc.*

facilitate physics analysis just few days after data-taking. Examples of the analyses that have used 'Turbo' stream are Refs [87, 88]. The Calibration stream separates events for further processing to calculate data-driven efficiencies for both the Full and Turbo streams.

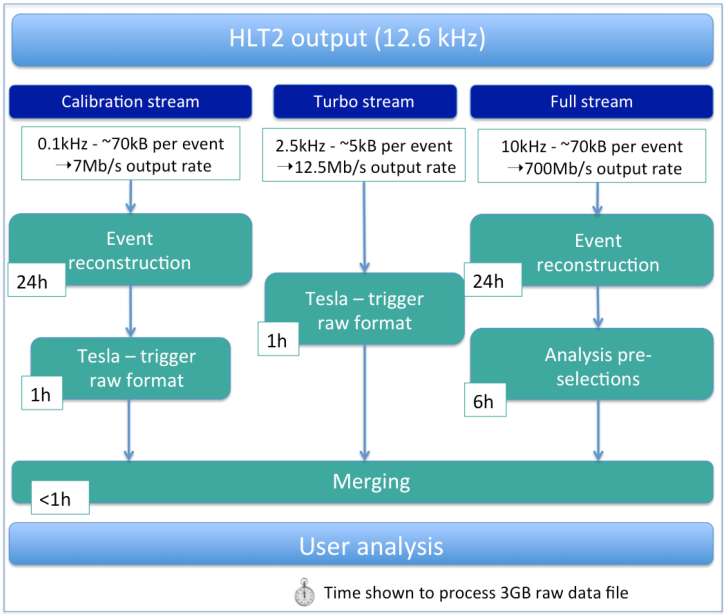

Figure 3.24: The data flow during Run II data-taking periods where two new streams, namely Turbo and Calibration, have been introduced. During Run I only the 'Full' stream existed. The time taken for each step in hours is provided for a 3 GB raw LHCb data file [86].

3.6 LHCb core software

The LHCb software [89] consists of several applications constructed in the Gaudi [90] framework. Figure 3.25 shows the data flow and how the various applications use this data. To produce simulation samples for direct comparison with the real collision data, I use PYTHIA [91] to generate pp collision events, EVTGEN [92] to describe the decay of b- or c-hadrons present in the generated events into a specific final state, GEANT4 [93] to propagate the final state particles through the LHCb detector, BOOLE [94] to digitise the data by modelling the response of the LHCb subdetectors and GAUSS [95] to mediate and propagate data from generation to propagation stage. The simulated data then undergoes the same processing as the real data. The MOORE [96] package is used to implement the trigger algorithms, BRUNEL [97] is used for reconstruction of tracks and DAVINCI [98] helps in selection of reconstructed candidates for a particular analysis.

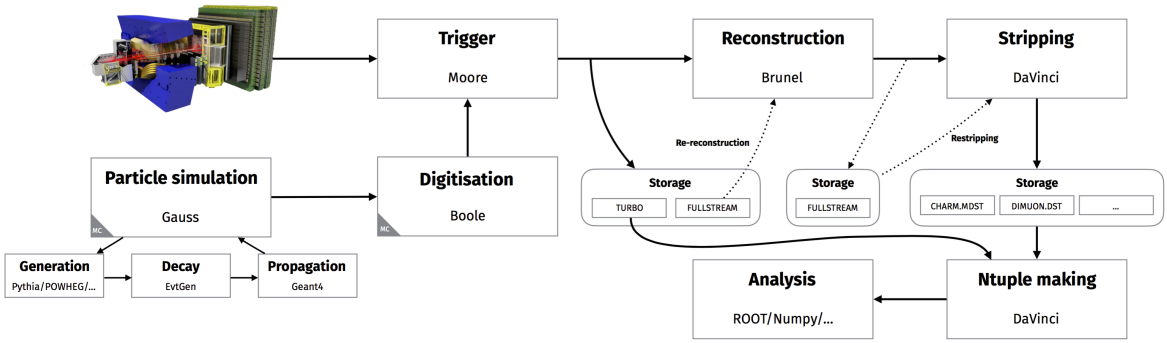

Figure 3.25: Various software applications with their function shown in bold [89].

CHAPTER 4

Search for $\Xi_b^-(\Omega_b^-) \to phh'^-$ decays

“ Recipe for dragon stew: First, find a dragon . . . ”

Matt Visser, *Lorentzian Wormholes*

4.1 Introduction

The charmless non-leptonic decays of B mesons have been well explored and provide fertile ground to study CP violation effects. Large CP-violating asymmetries have been observed in certain B meson decays [63, 99, 100]. At the LHC, b-baryons are produced at an unprecedented level, allowing a detailed study of their decays to be conducted for the first time.

Although the two-body $\Lambda_b^0 \to pK^-$ and $p\pi^-$ decays were observed some time ago [101], there had been no observation of a charmless three-body b-baryon decay until more recently, when the $\Lambda_b^0 \to K_S^0 p\pi^-$ decay mode was observed [102]. Since then, a few other charmless b-baryon decays have been observed: $\Lambda_b^0 \to \Lambda\phi$ [103]; $\Lambda_b^0 \to \Lambda K^+\pi^-$ and $\Lambda_b^0 \to \Lambda K^+K^-$ [104]; $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$, $\Lambda_b^0 \to p\pi^-K^+K^-$, $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$, $\Xi_b^0 \to pK^-\pi^+\pi^-$ and $\Xi_b^0 \to pK^-\pi^+K^-$ [105]. Among the three-body decays, the inclusive CP asymmetry parameters for $\Lambda_b^0 \to K_S^0 p\pi^-$, $\Lambda_b^0 \to \Lambda K^+\pi^-$ and $\Lambda_b^0 \to \Lambda K^+K^-$ have been measured and found to be consistent with zero, albeit with rather large uncertainties [104]. There have been searches for CP violation in two-body charmless Λ_b^0 decays [106], as-yet with no evidence for a non-zero effect. Searches for asymmetries in regions of the phase-space of the four-body charmless hadronic $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$, $\Lambda_b^0 \to p\pi^-K^+K^-$, $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$ and

$\Xi_b^0 \to pK^-\pi^+K^-$ decays have also been carried out, with $3.3\,\sigma$ evidence of a CP violation effect in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays [107, 108]. Therefore, CP violation in the baryon sector stands unobserved to date.

In this thesis a detailed study of $\Xi_b^-(\Omega_b^-) \to phh'^-$ decay channels is conducted. The topology of the $\Xi_b^-(\Omega_b^-) \to phh'^-$ decays are very similar to their mesonic counterparts $B^- \to h^-h'^-h'^+$ where one meson (h) is replaced by a proton (p) in the final state, although there is in principle an extra dependence of decay kinematics on additional angular variables in the baryonic case (due to presence of spin-half initial and final state particles). Recent analyses from LHCb of $B^- \to h^-h'^-h'^+$ have observed unexpectedly large CP asymmetries in certain regions of the phase space (Dalitz plane) [63, 99, 100]. If a similar behaviour occurs in the baryonic modes, then the $\Xi_b^-(\Omega_b^-) \to phh'^-$ channels could be good candidates to make a first observation of CP violation in the baryon sector through either model-independent methods [109–112] or through model-dependent amplitude analysis.

All the potentially dominant processes ($\Xi_b^- \to pK^-K^-$, $\Xi_b^- \to pK^-\pi^-$ & $\Omega_b^- \to pK^-K^-$) proceed in the Standard Model (SM) via CKM-suppressed Charged Current (CC) weak interactions at tree level and through penguin-suppressed Flavour Changing Neutral Current (FCNC) at loop level. These decays are discussed in detail below:

- The Feynman diagrams for the $\Xi_b^- \to pK^-K^-$ decay are shown in Figure 4.1. The decay at tree-level is $b \to u\overline{u}s \propto V_{ub}V_{us}^*$ ($\mathcal{O}(\lambda^4)$ at leading order in the Wolfenstein parameter λ), and at loop-level it is $b \to su\overline{u} \propto V_{tb}V_{ts}^*$ ($\mathcal{O}(\lambda^2)$). Hence this should be the most prominent decay (largest branching fraction), and CP violation effects of $\mathcal{O}(5\,\%)$ are expected to be possible. Note that the resonant decays of $\Xi_b^- \to \Lambda^*K^-, \Sigma^*K^-$ to pKK final states also proceed via the same transitions.

- The Feynman diagrams for the nonresonant $\Xi_b^- \to pK^-\pi^-$ decay are shown in Figure 4.2. The decay at tree-level is $b \to u\overline{u}d \propto V_{ub}V_{ud}^*$ ($\mathcal{O}(\lambda^3)$), while at loop level it is $b \to du\overline{u} \propto V_{tb}V_{td}^*$ ($\mathcal{O}(\lambda^3)$. Hence the branching fraction of this decay should be lower compared to $\Xi_b^- \to pK^-K^-$ decay, but $\mathcal{O}(100\,\%)$ CP violating asymmetries are in principle possible.

- The Feynman diagrams for the nonresonant $\Omega_b^- \to pK^-K^-$ decay are shown in Figure 4.3. The decay at tree-level is $b \to u\overline{u}d \propto V_{ub}V_{ud}^*$ ($\mathcal{O}(\lambda^3)$), while at loop level it is $b \to du\overline{u} \propto V_{tb}V_{td}^*$ ($\mathcal{O}(\lambda^3)$). Hence the branching fraction of this decay should be lower compared to $\Xi_b^- \to pK^-K^-$ decay, but $\mathcal{O}(100\,\%)$ CP violating asymmetries are in principle possible. Compared to the Ξ_b^- modes, the yields of Ω_b^- decays are expected to be additionally suppressed due to the expected smaller fragmentation fraction.

- The $\Xi_b^- \to p\pi\pi$ decay is expected to be highly suppressed compared to Ξ_b^- decays to the

$pK\pi$ and pKK final states due to the extra requirement of an $s \to d$ transition. Similarly, Ω_b^- decays to the $pK\pi$ and $p\pi\pi$ final states should also be highly suppressed as the former (latter) requires one (two) extra $s \to d$ transitions.

Note here that the b-quark fragmentation fraction (f) to a particular b-hadron, X_b^i, is defined as

$$f_{\mathrm{X_b^i}} = \frac{\sigma(b \to X_b^i)}{\sum_i \sigma(b \to X_b^i)} \tag{4.1}$$

where $\sigma(b \to X_b^i)$ denotes the hadronisation rate of a b-quark to a bottom hadron such as B^- ($b\overline{u}$), $\overline{B}^0$ ($b\overline{d}$), $\overline{B}_s^0$ ($b\overline{s}$), B_c^- ($b\overline{c}$), Λ_b^0 (bdu), Ξ_b^0 (usb), Ξ_b^- (dsb), Ω_b^- (ssb), Σ_b^+ (uub), Σ_b^0 (udb) or Σ_b^- (ddb). The b-quark fragmentation fraction to B^- ($b\overline{u}$), $\overline{B}^0$ ($b\overline{d}$) and $\overline{B}_s^0$ ($b\overline{s}$) mesons are denoted, in literature, by f_u, f_d and f_s, respectively.

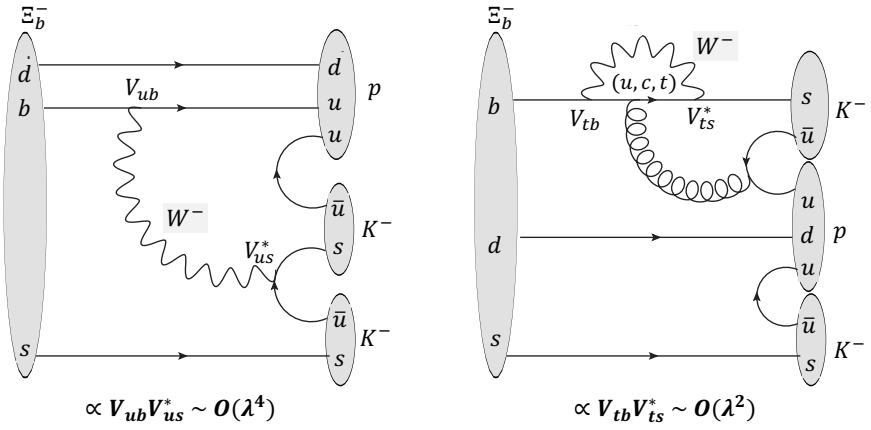

Figure 4.1: Feynman diagrams of $\Xi_b^- \to pK^-K^-$ decay channel.

The aim of the analysis described in this chapter is to make the first observation of the decay of the Ξ_b^- baryon to a three-body charmless final state that includes a proton and two like-charge mesons, *i.e.* $\Xi_b^- \to phh'^-$. In the same final state, it is also possible to search for $\Omega_b^- \to phh'^-$ decays. There are no published theoretical predictions or experimental limits on the branching fractions of the considered b-baryon decay channels. Therefore, we also report the relative branching fraction of these modes with respect to a suitable normalisation channel. The normalisation mode is chosen such that it is a high yield channel with precisely determined branching fraction and a parent particle with similar lifetime and similar decay topology as that of the signal channel, so that the common systematic uncertainties can be expected to cancel to first order. Since the largest yields are expected in the $\Xi_b^- \to pK^-K^-$ decay channel, the

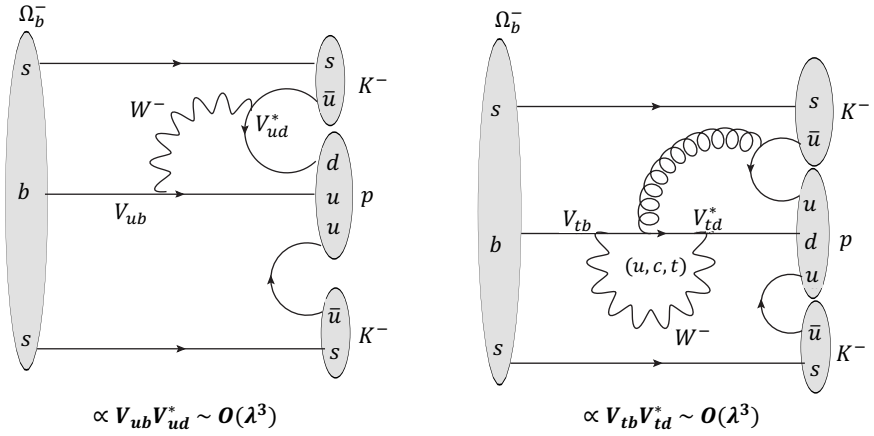

Figure 4.2: Feynman diagrams of $\Xi_b^- \to pK^-\pi^-$ decay channel.

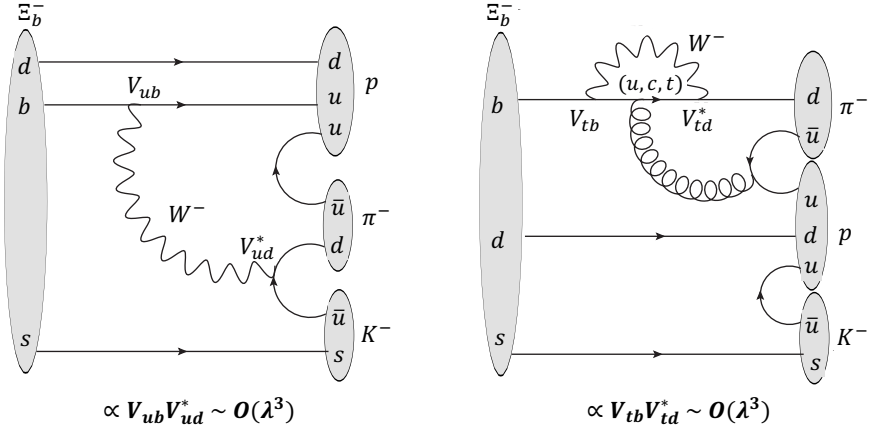

Figure 4.3: Feynman diagrams of $\Omega_b^- \to pK^-K^-$ decay channel.

normalisation channel is chosen to be $B^- \to K^-K^-K^+$. The properties of the Ξ_b^-, Ω_b^- and B^+ hadrons along with their mass and lifetime values used in this analysis are listed in Table 4.1 [113–116]. Therefore in this analysis we measure the ratio of fragmentation fractions times branching fraction ratios.

Table 4.1: Properties of the Ξ_b^-, Ω_b^- and B^+ hadrons [113–116].

	$I(J^P)$	Quark content	Mass (MeV/c^2)	Lifetime (ps)
Ξ_b^-	$\frac{1}{2}(\frac{1}{2}^+)$	bsd	$5797.72 \pm 0.46 \pm 0.31$	1.560 ± 0.040
Ω_b^-	$\frac{1}{2}(\frac{1}{2}^+)$	bss	6046.1 ± 1.7	1.66 ± 0.18
B^+	$\frac{1}{2}(0^-)$	$u\bar{b}$	5279.32 ± 0.14	1.638 ± 0.004

4.1.1 Analysis strategy

The strategy to observe the signal decays and measure their relative branching fraction with respect to the $B^- \to K^-K^-K^+$ normalisation channel can be broken down as follows:

- As the signal modes have not been previously observed, to avoid experimenter's bias entering the analysis, we have decided to remove or blind the $m(ph^-h'^-)$ invariant mass region around the Ξ_b^- and Ω_b^- masses until a signal selection procedure (to reduce background contributions) and nominal fit model (to extract the signal yield) are established.

- The signal selection procedure is aimed at reducing the background contributions by applying pre-selection requirements to the collision data. Further reduction of specific species of background can be conducted employing various techniques. In this analysis, we have trained neural networks using a chosen set of variables that provide good discriminating power between signal and combinatorial (from random combinations of tracks in an event) background candidates. A requirement on the output of the neural network is then optimised and used to reduce the combinatorial background contribution. An optimised requirement on the Particle Identification (PID) of the final state tracks has also been imposed, with the aim of reducing both combinatorial and cross-feed (from the mis-identification of one or more final state particles) backgrounds.

- In developing a fit model to extract the signal yield, the shapes of the signal and background components have been obtained using simulated (Monte-Carlo/MC) samples. Various potential sources of peaking and partially reconstructed background are considered. These are either vetoed, included in the final fit, or determined not to contribute in the region of interest, as appropriate.

- The signal selection procedure could introduce dependence of the signal selection efficiency on the phase-space of the signal decays, therefore the raw yields (*i.e.* those directly obtained from the fit to the selected data) of signal and normalisation modes have to be corrected

before evaluating the relative branching fraction. These phase-space dependent efficiencies are obtained from MC samples. Due to the spin-$\frac{1}{2}$ particles in the initial and final states the phase-space of the signal modes is five-dimensional, and can be represented by two Dalitz plot variables (*i.e.* invariant masses-squared of two different particle combinations) and three angular variables.

- Using the signal yields ($\mathcal{N}$) obtained from the fit to the b-baryon candidate mass distribution, together with phase-space dependent efficiencies (ϵ), we measure for all investigated channels the relative branching fractions ($\mathcal{B}$) multiplied by the ratio of fragmentation fractions for b-hadrons in the LHCb acceptance, as follows:

$$\frac{\mathcal{B}(\Xi_b^-(\Omega_b^-)\to hhp)}{\mathcal{B}(B^-\to KKK)}\times\frac{f_{\Xi_b^-(\Omega_b^-)}}{f_u}=\frac{\mathcal{N}_{\Xi_b^-(\Omega_b^-)}}{\mathcal{N}_{B^-}}\times\frac{\epsilon_{B^-}}{\epsilon_{\Xi_b^-(\Omega_b^-)}}\,. \tag{4.2}$$

  Since the fragmentation fraction ratio $f_{\Xi_b^-(\Omega_b^-)}/f_u$ is at present unknown, it will not be possible to convert the relative branching fraction into an absolute value. The systematic uncertainties introduced by the measurement strategy will also be evaluated.

- The significance of each signal channel will be investigated using Wilks' theorem [117]. For channels with no significant signal ($< 3\sigma$), an upper limit on the relative branching fraction will be set.

- In the event that significant signal is observed for $\Xi_b^-\to pK^-K^-$ (since it is the least suppressed signal mode in SM), we hope to also measure

$$\frac{\mathcal{B}(\Xi_b^-\to h^-\pi^-p)}{\mathcal{B}(\Xi_b^-\to pK^-K^-)} \tag{4.3}$$

  from the ratios of yields and efficiencies. This is advantageous as it does not involve any ratio of fragmentation fractions, and certain systematic uncertainties may also be expected to be reduced. we therefore also report these quantities for both $\Xi_b^-\to pK^-\pi^-$ and $\Xi_b^-\to p\pi^-\pi^-$ decays.

This analysis is performed with the LHCb Run 1 data sample comprising $1\,\text{fb}^{-1}$ collected during 2011 with pp collision energy $\sqrt{s}=7\,\text{TeV}$ and $2\,\text{fb}^{-1}$ collected during 2012 with $\sqrt{s}=8\,\text{TeV}$. The inclusion of charge-conjugate modes is implied throughout this chapter, unless otherwise stated.

4.2 Signal selection

In the data sample, the signal is buried under a large amount of background. One needs to identify and reduce the background contributions to reliably extract the signal yield with the smallest possible statistical uncertainty. In this section, I discuss the various procedures employed in the signal selection. In Subsec. 5.2.1.1, I highlight a set of pre-selection requirements imposed on the data sample before being made available for further processing, which are collectively referred to as 'Stripping' requirements. Also outlined in Subsec. 5.2.1.2 are certain requirements on how the selected candidates were triggered. In the last Subsec. 4.2.3, I discuss additional selection requirements imposed on the data sample to further reduce the background contributions.

4.2.1 Stripping

In this analysis, I use a data sample that consists of candidates that are reconstructed from three charged particles that have been assigned the kaon mass. This sample undergoes a set of pre-selection requirements, referred to here as 'Stripping' lines. The stripping line used in this analysis is `StrippingB2hhh_KKK_inclLine` and the set of requirements pertaining to this stripping line is shown in Table 4.2. [1] As seen from this table, the daughter candidates undergo loose requirements on kinematic and topological variables before b-hadron candidates are formed by simple four-momentum addition. Then the formed b-hadron candidates undergo 'combination requirements' to reduce the number subject to the full decay vertex fit. Following the vertex fit, further requirements ('mother requirements') are made on the b-hadron candidates.

The final state particles of these stripped candidates are then assigned various particle masses and all the parameters in the decay tree for a candidate are refit simultaneously using a package named "DecayTreeFitter". The refitting is performed using a Kalman fitter which conducts progressive refitting of the full decay tree for a given mass hypothesis of the daughter particles together with selected vertex constraints. The refitting results in, amongst other things, significantly improved invariant mass resolution of the particles 'upstream' in the decay. The estimated signal efficiency of the stripping requirements is about 30% for all the signal modes.

4.2.2 Trigger

To select the hadronic final states, at the hardware L0 stage, I require that either the L0 hadron trigger fired as 'Trigger on Signal' (TOS) or the global L0 trigger fired as 'Trigger Independently of Signal' (TIS) (see Sec. 3.5 for more details). At the first stage of the high level trigger,

[1] The corrected mass (`BPVCORRM`) is defined as $M^{\rm CORR} = \sqrt{M^2 + |P_{\rm T}^{\rm miss}|^2} + |P_{\rm T}^{\rm miss}|$, where M is the $B^{\pm}$ candidate mass and $P_{\rm T}^{\rm miss}$ is the missing momentum transverse to the line of flight of the B decay.

Table 4.2: Requirements used in the `StrippingB2hhh_KKK_inclLine` stripping line. The impact parameter, vertex and track quality are evaluated from the results of the Kalman fitter during the reconstruction.

Particle(s)	Requirements	Description
Daughter track requirements	$p_{\mathrm{T}} > 100\,\mathrm{MeV}/c$	Transverse momentum
	$p > 1500\,\mathrm{MeV}/c$	Momentum
	$\chi^2_{\mathrm{IP}} > 1$	χ^2 distance of particle's trajectory to PV
	χ^2/NDOF < 3	χ^2 per degree of freedom of the track fit
	Track_GHOSTPROB < 0.5	Ghost probability (See Sec. 3.3)
Combined four-momentum requirements	$5.05 < m(KKK) < 6.3\,\mathrm{GeV}/c^2$	Mass of combination of 3 charged kaons.
	DOCA < 0.2 mm	Distance of closest approach between all possible pairs of particles in 3D
Mother requirements (after full vertex fit)	Leading $p_{\mathrm{T}} > 1500\,\mathrm{MeV}/c$	Transverse momentum of the leading p_{T} track
	DIRA(θ) > 0.99998	Cosine of angle between B candidate momentum and direction from associated PV to decay vertex
	Flight $\chi^2 > 500$	χ^2 distance of B candidate from associated PV
	Vertex $\chi^2 < 12$	Vertex χ^2 of B candidate
	$\chi^2_{\mathrm{IP}} < 10$	IP χ^2 wrt PV for B candidate
	$p_{\mathrm{T}} > 1000\,\mathrm{MeV}/c$	B cand. p_{T}
	$\sum p_{\mathrm{T}} > 4500\,\mathrm{MeV}/c$	Vector p_{T} sum of daug. tracks of B candidate
	$\sum p > 20000\,\mathrm{MeV}/c$	Vector p sum of tracks in B candidate
	$\sum \chi^2_{\mathrm{IP}} > 500$	Sum of IP χ^2 wrt PV of tracks in B candidate
	$4 <$ BPVCORRM $< 7\,\mathrm{GeV}/c^2$	B candidate corrected mass under KKK hypothesis.

HLT1, candidates with high p_{T} and tracks with significant impact parameter (IP) that trigger the `Hlt1TrackAllL0Decision_TOS` line (see Sec. 3.5.2 for more details) are accepted. At the second stage of the high level trigger, HLT2, I require that one of the topological triggers was fired by the signal candidate (see Sec. 3.5.2 for more details). Hence the trigger requirement can be expressed as

```
(L0Global_TIS || L0HadronDecision_TOS) &&
    Hlt1TrackAllL0Decision_TOS &&
(Hlt2Topo{2,3}BodyBBDTDecision_TOS).
```

4.2.3 Offline selection

For the search of the so-far unobserved modes, the procedure chosen was to blind the $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal mass regions, to reduce the possibility of bias until the signal selection procedure and the fit model has been finalised. The blind regions lie within ± 120 MeV (*i.e.* $\pm 6\sigma$ where signal width $\sigma \sim 20\,\mathrm{MeV}$) of the nominal Ξ_b^- ($5794.4\,\mathrm{MeV}/c^2$) and Ω_b^- ($6049\,\mathrm{MeV}/c^2$) masses, corresponding to a range of $[5674.4, 5914.4]$ MeV/c^2 and $[5929, 6169]$ MeV/c^2, for the Ξ_b^- and Ω_b^- regions, respectively.[2] Consequently, the whole offline selection has been optimised with

[2] The mass values are not identical to those given in Table 4.1 as the blind windows were chosen at an early stage of the analysis. The differences in mass are small compared to the widths, and thus do not affect the analysis.

respect to the control channel $B^+ \to K^+K^-K^+$, which has high signal statistics and has the same topology as my signal modes.

In this sub-section, I describe the offline selection of the $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes, including its optimisation. The discussion starts with the procedure employed for background subtraction of the $B^+ \to K^+K^-K^+$ data sample. I then describe how this pure signal sample is used in the training of the multivariate analyser (MVA). Finally, I tackle the optimisation of MVA output and Particle Identification (PID) criteria to reduce contributions from various background categories.

The selection is optimised for the combined 2011 and 2012 datasets to increase signal statistics and hence signal significance. To simplify the analysis, I do not perform explicit optimisation for Ω_b^- decays but use the same selection optimised for the $\Xi_b^- \to phh'^-$ modes.

4.2.3.1 Background subtraction of $B^+ \to K^+K^-K^+$ data sample

Along with the trigger and stripping requirements discussed previously, I also apply a set of pre-selection requirements to the $B^+ \to K^+K^-K^+$ control mode before the invariant mass distribution is parametrised. The fit range for $m(KKK)$ is chosen to be $\{5220, 5600\}\,\mathrm{MeV}/c^2$ as this removes most of the contribution from partially reconstructed background, thus avoiding the need to parametrise its distribution. The fit to full range of $B^+ \to K^+K^-K^+$ mode can be seen in Refs [63, 99, 100]. I also apply a requirement on all the final state particles not to be Muons (*i.e.* `IsMuon`$= 0$ [3]) to exclude cases where a muon fakes a light meson, for example from $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$. Since we do not have any particle identification power in certain low and high values of the kinematic variables, we simply remove such regions by a set of kinematic requirements on the momentum, transverse momentum and track multiplicity. The PID requirements on the three kaon tracks are based on the PID selection developed for the $B^+ \to K^+K^-K^+$ analysis [63]. A summary of all these pre-selection requirements is shown in Table 4.3.

With the aforementioned set of requirements applied, I perform an unbinned extended maximum likelihood fit to the $B^+ \to K^+K^-K^+$ candidates' invariant mass distribution including the dominant background components. Using the signal, background and total Probability Density Functions (PDF), a set of weights are extracted for each component (referred to as *sWeights*) using the *sPlot* method as described in Ref. [118]. The sum of *sWeights* for each component is equal to the total number of candidates for that component extracted from the fit to the $m(KKK)$. Variables that show negligible correlation with the $m(KKK)$ invariant mass can be weighted with the signal *sWeights* to obtain background subtracted distributions.

[3] The `IsMuon` binary selection is defined according to the number of muon stations where a hit is found within a field of interest that is defined around the track extrapolation.

Table 4.3: Preselection requirements on $B^+ \to K^+K^-K^+$ data sample for the background subtraction procedure. Here h1, h2 and h3 refer to the three final state particles. Please refer to Sec. 3.4 for details on the 'ProbNN' PID variables.

Variable	Selection Requirements
Mass of $B^\pm$ candidate	$5220 < m(pKK) < 5600\,\mathrm{MeV}/c^2$
Daughters PID	{h1,h2,h3}_PROBNNK>0.2
Not muon requirement	{h1,h2,h3}_IsMuon= 0
Daughter momentum	{h1,h2,h3}_P $< 100\,\mathrm{GeV}/c$
Daughter transverse momentum	{h1,h2,h3}_PT $< 30\,\mathrm{GeV}/c$
Track multiplicity	nTracks < 700

In the fit to the $m(KKK)$ the signal shape is described by double Crystal Ball function (DCB) [119] obtained from a fit to the $B^+ \to K^+K^-K^+$ MC sample. The Crystal Ball (CB) function is defined as a central Gaussian function with a power-law tail [119],

$$\mathrm{CB}(t;n,\alpha,\sigma) = \mathcal{N}\cdot\begin{cases} \exp(-t^2/2\sigma^2) & \text{if } t/\sigma > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \left(\frac{n-\alpha^2}{|\alpha|} - \frac{t}{\sigma}\right)\exp(-\alpha^2/2) & \text{if } t/\sigma \leq -\alpha, \end{cases} \tag{4.4}$$

where $t = m - \mu$ is related to the reconstructed mass m, μ and σ are the mean and standard deviation of the Gaussian function, respectively, and $\mathcal{N}$ is a normalisation factor. The power law in the tail is controlled by n and the sign of α determines its size relative to the mean of the central Gaussian. Instead of the independent normalisation factors, in the DCB PDF a parameter, $f = \frac{\mathcal{N}_1}{\mathcal{N}_1+\mathcal{N}_2}$ is introduced, which indicates the relative sizes of the two CB functions *i.e.* $\mathrm{DCB} = \mathrm{CB}_1 + f \times \mathrm{CB}_2$. The left tail of the DCB primarily describes the radiative tail of the mass distribution, which has been shifted to lower values by the radiation of one or more photons from the final state particles. Final state radiation is expected to be a smaller effect in the modes containing one or more kaons, owing to their higher mass. The right tail accounts for non-Gaussian tracking imperfections and other related stochastic detector effects. In the subsequent fit to data, all tail parameters of the DCB are fixed, as are the relative yields of the two Crystal Ball functions [102, 104].

Major background contributions can arise from random combination of tracks that fake the KKK candidate (referred to as combinatorial background) and from other $B^- \to h^-h'^-h'^+$ decays due to misidentification (mis-ID) of one or more final state particles (referred to as cross-feed background). The shape of cross-feed background from $B^+ \to K^+\pi^-K^+$ decays is obtained by fitting a double Crystal Ball function to the KKK invariant mass distribution obtained when

reconstructing a $B^+ \to K^+\pi^-K^+$ MC sample in the KKK final state. The parameters of this function have similar constraints applied as in the case of $B^+ \to K^+K^-K^+$ signal component, when fitting data. The combinatorial background is modelled with an exponential function. The MC samples have had the same pre-selection requirements imposed that were applied to the data sample before extracting the shapes.

Figure 4.4 (top) shows the fit result overlaid on the data. In the same figure (bottom) I show the signal *sWeight* as a function of KKK invariant mass. It is worth noting here that the yields of all three components were left floating during the fit, as required by the *sPlot* method. It is evident from the pull distribution of the fit in Figure 4.4 that the fit model could be improved; similarly the cross-feed yield is found to be about a factor of 2 larger than expected based on the known branching fractions and misidentification probabilities. However the mis-modelling introduces a very negligible second order effect in the MVA performance, where the MVA is trained using *sWeights* extracted from the fit (only signal *sWeight* are required; discrimination between cross-feed and combinatorial background does not affect the training). I confirm this by examining the agreement between the *sWeighted* data and $B^+ \to K^+K^-K^+$ signal MC distributions of the input variables for the training of the Neural Networks (NN), which will be investigated in the next subsection. The fitted yields for signal, cross-feed and combinatorial background are shown in Table 4.4.

Table 4.4: Yields extracted from the fit to the $B^+ \to K^+K^-K^+$ control channel used to determine *sWeights*.

Signal yield	Cross-feed yield	Combinatorial bkgd yield
91413 ± 364	2883.9 ± 231	21470 ± 327

4.2.3.2 Training of multivariate analyser

As previously mentioned, the approach is to train an MVA, in this case neural networks using the NEUROBAYES algorithm. To train the MVA classifier, the same $B^+ \to K^+K^-K^+$ data sample is used as signal and background training samples but in each case the data is weighted by signal and background *sWeights*, respectively. Note here that the background *sWeight* for each candidate is just 1 minus the signal *sWeight*.

The neural network is made of one hidden layer with a sigmoid activation function with as many hidden nodes as the number of input variables used to train the neural network. Adding or removing the nodes in the hidden layer did not lead to significant improvement in the neural net performance and neither did usage of a different activation function. I have chosen nine

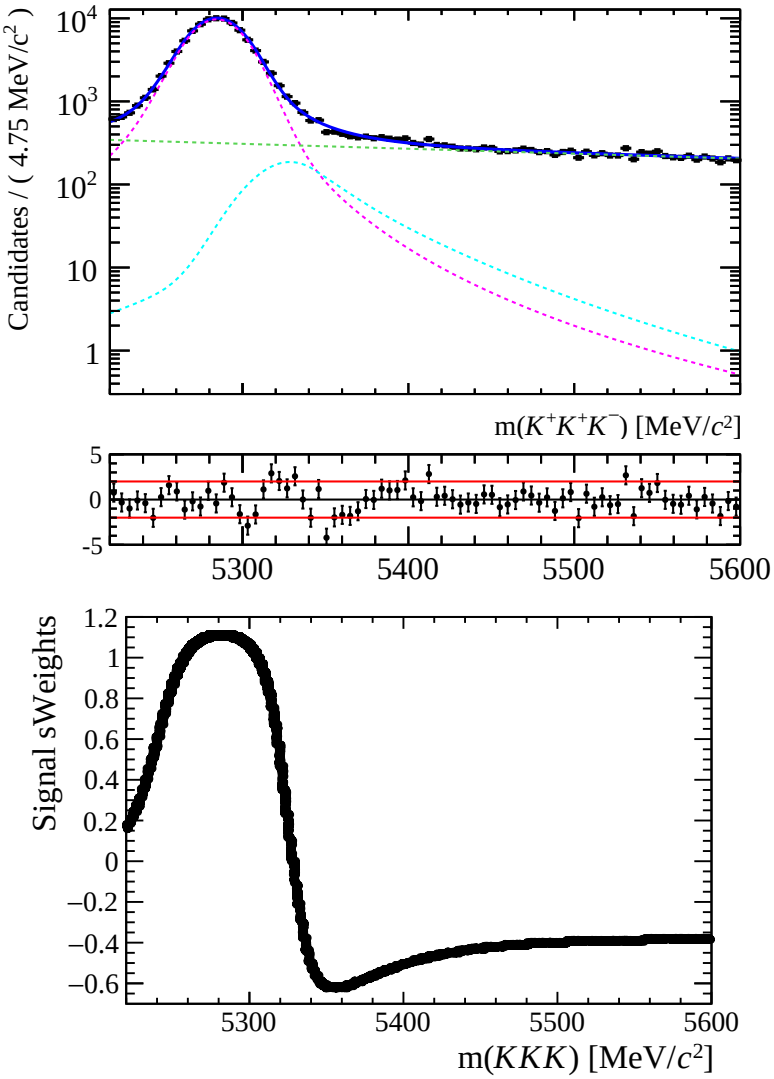

Figure 4.4: (Top) Fit to the $B^+ \to K^+K^-K^+$ control channel used for *sWeight* extraction. The dashed cyan line is $B^+ \to K^+\pi^-K^+$ cross-feed, the dashed pink line is signal mode and the dashed green line is combinatorial background. The pull defined as $\frac{N_{fit}-N_{data}}{\sigma_{N_{data}}}$, where N_{fit} is the yield obtained from the fit and $N_{data} \pm \sigma_{N_{data}}$ are the number of candidates with associated uncertainty in a given bin of the invariant mass distribution, is also shown. (Bottom) *sWeights* as a function of KKK invariant mass.

discriminating variables as inputs to the neural network training. These variables are listed in Table 4.5 according to their ranking/significance. The criteria used for choosing these input variables can be summarised as follows:

- They must have different signal and background distributions.

- They must have the same signal distribution in $B^+ \to K^+K^-K^+$ and $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes.
- They must have the same distribution in sWeighted $B^+ \to K^+K^-K^+$ data and $B^+ \to K^+K^-K^+$ signal MC samples. I show the data/MC discrepancy plots later in the section.
- Have relatively low degree of correlation with $m(pKK)$, with phase-space variables and are considered as significant variables by the NEUROBAYES algorithm.

Table 4.5: Input variables used in the training of the Neural Networks with the NEUROBAYES algorithm. For the definition of angle $\theta_{\rm DIRA}$ refer to Table 4.2.

Rank	Input variables	Description
1	Bu_POINTING	$\frac{\lvert p_B\rvert\times\sin(\theta_{\rm DIRA})}{\lvert p_B\rvert\times\sin(\theta_{\rm DIRA})+\sum_{\rm tracks} p_{\rm T}}$
2	sum_h_PT	$\sum p_{\rm T}$ of all tracks
3	IPCHI2_LEADING_DAUG	$\chi^2_{\rm IP}$ of the leading $p_{\rm T}$ track
4	Bu_ENDVERTEX_CHI2NDOF	B cand vertex χ^2/NDOF
5	Bu_DIRA_OWNPV	$\cos(\theta_{\rm DIRA})$
6	sum_h_IPCHI2_OWNPV	$\sum \chi^2_{\rm IP}$ for all tracks
7	PT_LEADING_DAUG	$p_{\rm T}$ of the leading daughter track
8	Bu_PTASYM_1_5	$p_{\rm T}$ asymmetry for a cone in η–ϕ plane with radius 1.5[4]
9	Bu_IPCHI2_OWNPV	$\chi^2_{\rm IP}$ of B candidate wrt PV

All the chosen input variables satisfy the above mentioned conditions. The agreement of MVA input variables between the $B^+ \to K^+K^-K^+$ control channel (denoted by slanted blue lines) and $\Xi_b^- \to pK^-K^-$ (denoted by slanted pink lines) signal mode only can be seen in Figures 4.5 and 4.6. These figures also show MVA input variable distributions for signal *sWeighted* $B^+ \to K^+K^-K^+$ data (denoted by red markers). Also shown in these figures are the distributions of the input variables for background candidates which lie in the Right Side Band (RSB) region of $m(KKK)$, which is defined as $m(KKK) > 5400\,{\rm MeV}/c^2$. The distribution of the Bu_PTASYM_1_5 variable is a bit different for $\Xi_b^- \to pK^-K^-$ and $B^+ \to K^+K^-K^+$ MC samples. A similar scenario is observed in data, after unblinding, through comparison of sWeighted data distributions of Bu_PTASYM_1_5 variable for $\Xi_b^- \to pK^-K^-$ and $B^+ \to K^+K^-K^+$ decay modes.

The classification response of the neural network output (NN), trained using the NEUROBAYES algorithm, on the test sample is shown in Figure 4.7 (top). In order to have a

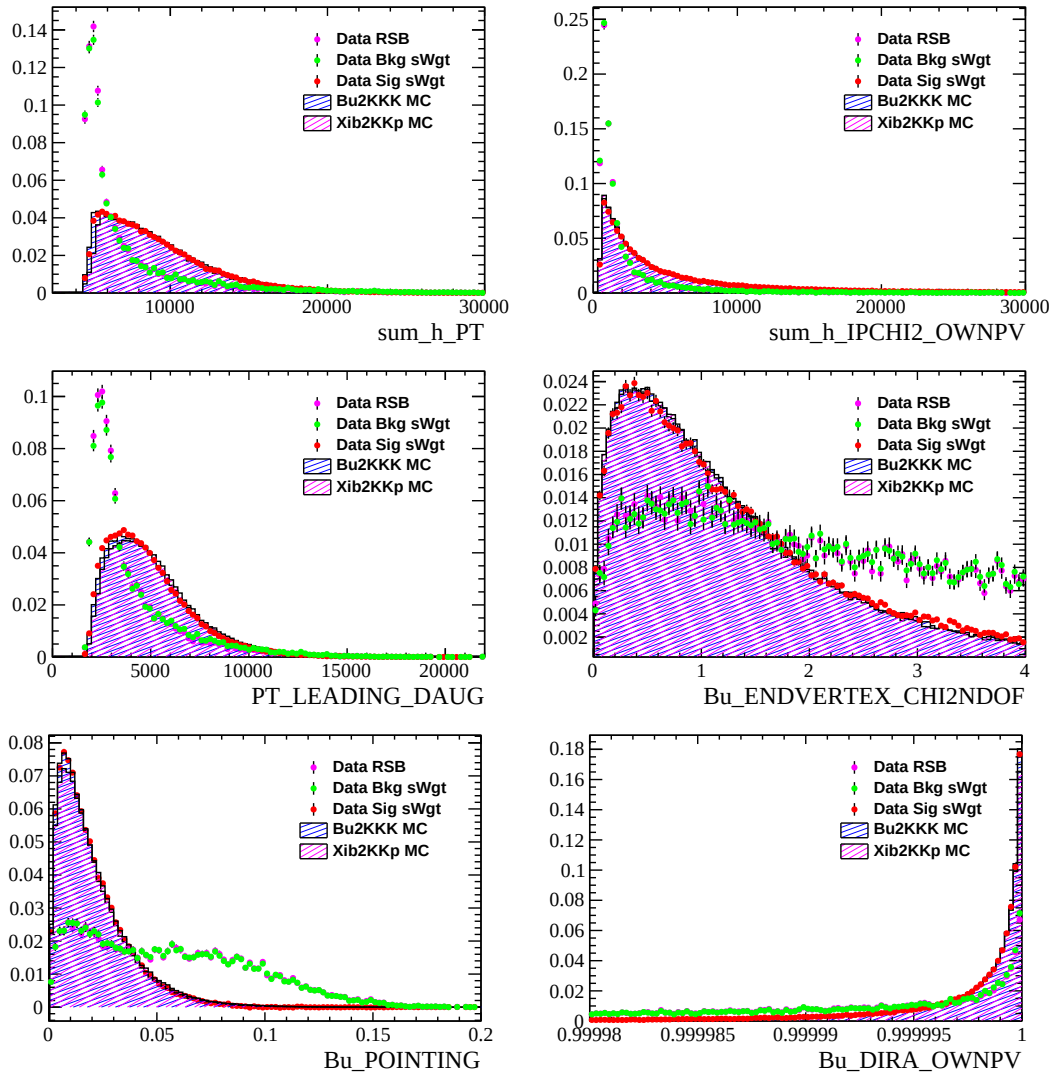

Figure 4.5: NEUROBAYES input variable distributions. Pink markers indicate Right Side Band (RSB) of the data, green markers indicate background *sWeighted* data, red markers signal *sWeighted* data, slanted blue lines indicate $B^+ \to K^+K^-K^+$ MC and slanted red lines indicate $\Xi_b^- \to pK^-K^-$ MC distributions.

NEUROBAYES output that can be interpreted as a probability that a given candidate is signal, the purity has to be a perfect linear function of the NEUROBAYES output (*i.e.* $P = 0.5 \times (\mathrm{NN} + 1)$). This linear relation between signal purity and the MVA output proves that the NEUROBAYES is well calibrated in the training. For this training, the purity in bins of NEUROBAYES output is shown in Figure 4.7 (bottom).

In Figure 4.7 it can be seen that the MVA output has quite a lot of peaks. This is because,

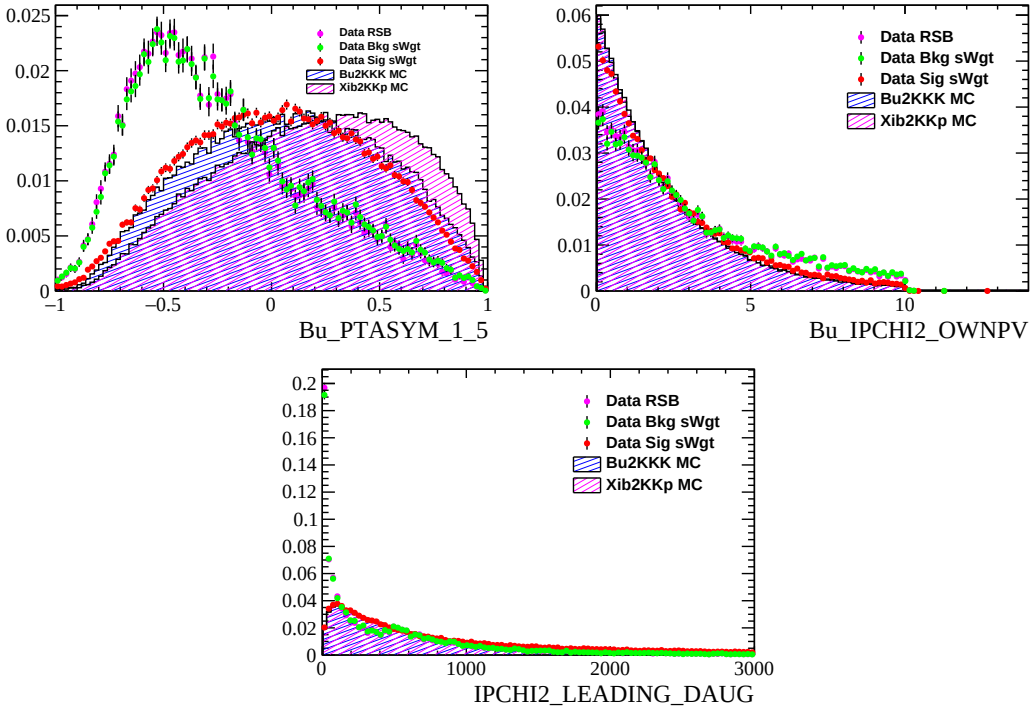

Figure 4.6: NEUROBAYES input variable distributions. Pink markers indicate Right Side Band (RSB) of the data, green markers indicate background *sWeighted* data, red markers signal *sWeighted* data, slanted blue lines indicate $B^+ \to K^+K^-K^+$ MC and slanted red lines indicate $\Xi_b^- \to pK^-K^-$ MC distributions.

at the pre-processing stage of the training, a linear model is constructed that is a single linear combination of the input variables. Then a monotonic transformation is carried out so that the NEUROBAYES output can be interpreted as probability. It can happen that several values of input variables transform to the same output, which gives us the peaks that are observed in the MVA output.

4.2.3.3 Optimisation of selection for background rejection

The general idea behind selection optimisation is to try to maximise some "figure-of-merit" (FoM), such as the signal to background ratio. For decays with unknown branching fractions, it is useful to use the so-called Punzi FoM [120], which is independent of the signal decay rate, and is given by the following formula:

$$\text{FoM} = \frac{\epsilon}{a/2 + \sqrt{B}}, \tag{4.5}$$

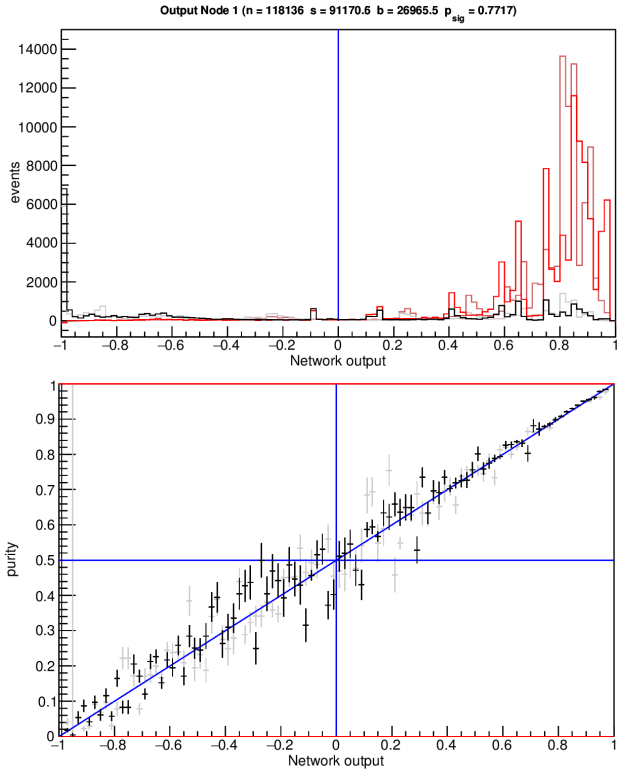

Figure 4.7: NEUROBAYES output distribution (top), where red corresponds to signal and black is background, and purity vs NEUROBAYES output (bottom). The faded red and black lines and black markers in the above figures indicate distributions in the first iteration of the boosting procedure whereas the solid red and black lines and black markers indicate the distributions for the final iteration.

Here ϵ is the signal efficiency for a given selection requirement, which is obtained from the MC sample, a is the aimed-for significance of the measurement (a 5σ observation is aimed for in this analysis, and thus $a = 5$) and B is the expected number of background candidates passing the given requirement in the signal region.

I perform a simultaneous optimisation of the PID and NEUROBAYES output requirements, with the intention of reducing the combinatorial background (cross-feed backgrounds are considered separately). For PID, the neural network-based PROBNN variables are used, whose values vary between 0 and 1.

To evaluate the signal efficiency for a given NEUROBAYES output requirement is straightforward. However, the signal efficiency of a given PID requirement cannot be evaluated by imposing it directly on the MC sample, because of the well-established disagreement [121] between simulation and data in the distributions of PID variables. One possible approach is to use

a data-driven calibration of MC PID variables using data samples, provided by the `PIDCalib` package [121], which have been fully reconstructed using only the kinematic variables with no *a priori* knowledge of PID variables. The efficiency could then be determined for a particular PID selection requirement on a candidate-by-candidate basis as a function of the kinematics of the MC track. The efficiencies are evaluated as functions of momentum (p), transverse momentum ($p_{\rm T}$) and track multiplicity (`nTracks`). Another important thing to note here is that the track multiplicity (`nTracks`) is not correctly simulated in the MC samples. Therefore, for all the MC samples, prior to the calculation of PID and mis-identification efficiencies, the `nTracks` distribution has been randomly sampled from that obtained from the signal *sWeighted* $B^+ \to K^+K^-K^+$ data sample.

Before optimising a requirement on the $\Xi_b^-\,(\Omega_b^-) \to phh'^-$ signal modes, I impose the trigger requirements discussed previously and also apply pre-selection requirements, similar to those discussed for the $B^+ \to K^+K^-K^+$ control mode (we show these requirements later). I have vetoed the possible intermediate charm contribution from the $\Xi_b^- \to \Xi_c^0(ph)h'$ decay (even though the decay $\Xi_c^0 \to pK^-$ is isospin-suppressed and not yet observed). In the case of a considerable contribution from the decay with intermediate charm, I could reverse the veto to investigate this mode further. These pre-selection requirements are shown in Table 4.6. Since we do not have any particle identification power in certain low and high values of the kinematic variables, we simply remove such regions by a set of kinematic requirements on the momentum, transverse momentum and track multiplicity.

Table 4.6: Pre-selection requirements on $\Xi_b^-\,(\Omega_b^-) \to phh'^-$ signal modes. Here the three final state particles are given a generic name h1, h2 and p. I also apply the same pre-selection requirements to the normalisation mode (except for the minimum proton momentum requirement `p_P` $> 10\,\text{GeV}$ and the $m(pKK)$ requirement). Refer to the text as to why these requirements have been chosen.

Variable	Selection requirements
Mass of Ξ_b^- (Ω_b^-)	$5500 < m(pKK) < 6400\,\text{MeV}/c^2$
Veto $\Xi_b^- \to \Xi_c^0(ph)h'$	$2446 < m(Kp) < 2496\,\text{MeV}/c^2$
Not muon requirement	{`h1,h2,p`}`_IsMuon`=0
Daughters momentum	{`h1,h2,p`}`_P` $< 100\,\text{GeV}/c$ && `p_P` $> 10\,\text{GeV}/c$
Daughters transverse momentum	{`h1,h2,p`}`_PT` $< 30\,\text{GeV}/c$
Track multiplicity	`nTracks` < 700
Initial PID requirement	$\sqrt{\texttt{PROBNNK}^2 + \texttt{PROBNNpi}^2} > 0.02$

Combinatorial background rejection: The combinatorial background, which arises from random combinations of tracks in an event, can be reduced by requirements on both the NEUROBAYES output (NN) and the PID variables of the final state tracks. The PID variables are effective since most random tracks are pions, and therefore if the final state contains protons and kaons, much of the background can be rejected with little signal loss. Hence the strategy of simultaneous optimisation of NN and PID selection requirements. It will be shown later that this is the dominant background for all of my signal modes.

The proton PID requirement is, however, considered separately in order to reject the dominant cross-feed background from $B^+ \to K^+K^-K^+$. Therefore I apply a requirement $\texttt{PROBNNp} \times (1 - \texttt{PROBNNK}) > x$ to identify a proton where x is an optimal value of the PID requirement. When this study was conducted with a preliminary requirement on the NN output (NN > 0.7), the maximum of the Punzi FoM preferred a hard requirement on the proton PID (*i.e.* $x > 0.8$). However, it was found that around the maximum of the FoM (see Fig. 4.8) there were some fluctuations that resulted from a low amount of cross-feed background in the signal region. Therefore, rather than taking the absolute maximum of the FoM, I opt to take the value just before the maximum where the FoM can be extracted with a low relative uncertainty ($< 20\%$). This still leads to a relatively tight requirement of $\texttt{PROBNNp} \times (1 - \texttt{PROBNNK}) > 0.5$ with a corresponding signal efficiency of $\sim 50\%$.

We now need a strategy to classify the like-charged meson tracks as either pion or kaon. This strategy is give below:

$$\text{Kaon: } \texttt{PROBNNK} - \texttt{PROBNNpi} > y\,, \qquad \text{Pion: } \texttt{PROBNNK} - \texttt{PROBNNpi} < y\,.$$

where y is the optimal value that classifies a given track as kaon or pion. The motivation for such a strategy is that, we have three different final state samples to investigate pKK, $pK\pi$ & $p\pi\pi$ and a given candidate with a specific final state needs to be only present in one of these 3 samples. The strategy employed designates uniquely a candidate to a particular sample. It would, however, be possible to choose a different PID requirement for each final state particle in each decay, as done in several other analyses. However, my choice simplifies the analysis, since there is a single boundary defined by 'y' to classify K & π tracks.

After imposing the proton PID requirement, discussed previously, I perform 2D scans of the FoM for NEUROBAYES output (NN) and PID offset (y) for the $K\pi p$ sample in the ranges -1 to 1 and 0 to 1 respectively, with a step size of 0.05 for both requirements. The optimal values for selection requirement determined here for NN and y will be applied for all hhp samples accordingly. Note here that I have an additional requirement of $\sqrt{\texttt{PROBNNK}^2 + \texttt{PROBNNpi}^2} > 0.02$ on PID variables of the mesons in the $p\,h^-h'^-$ data sample to reject background candidates

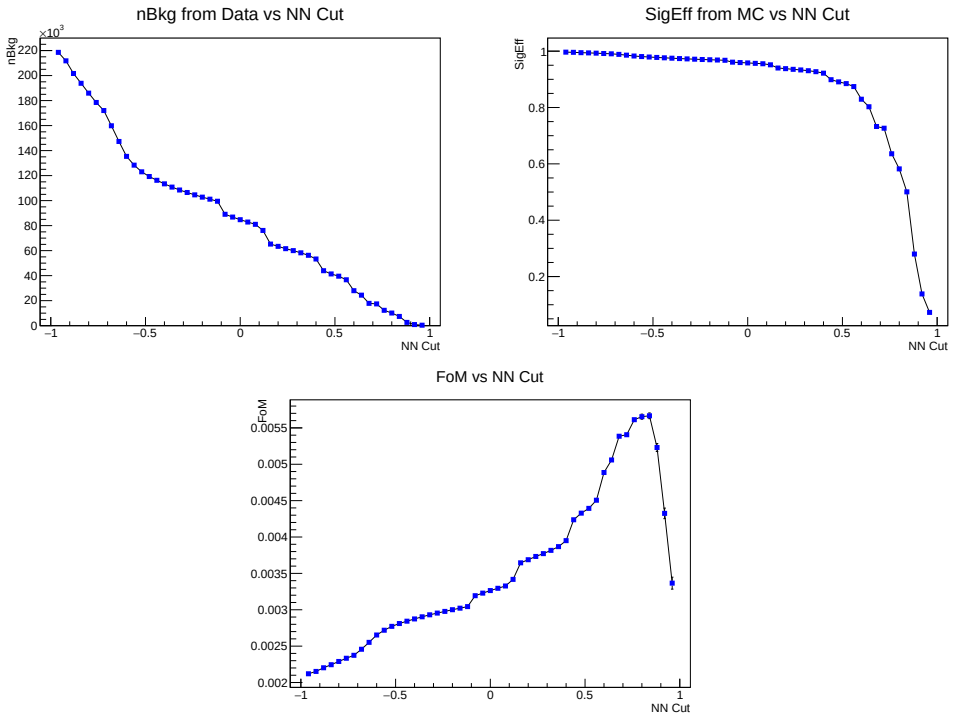

Figure 4.8: The number of combinatorial background events in the signal region (top left), signal efficiency from MC (top right) and Punzi FoM as a function of NN cut (bottom) for $\Xi_b^- \to pK^-K^-$ mode.

where a track is not well identified as either a π or a K.

The efficiency of the signal selection in the expression for the Punzi FoM (Eq. 4.5), in the signal region defined as $m(pK\pi) \in [5674.4, 5914.4]\,\mathrm{MeV}/c^2$, is obtained from the PID weighted MC sample, where candidate-by-candidate PID weights have been calculated using the `PIDCalib` package, along with the selection requirement on NN. The number of combinatorial background candidates lying in this signal region is obtained by conducting a linear fit to the right sideband ($m(pK\pi) \in [6169, 6400]\,\mathrm{MeV}/c^2$) of the invariant mass distribution and extrapolating the yield in the signal region.

Figure 4.9 shows the Punzi FoM values (top left) with errors (top right), signal efficiencies (middle) with errors (middle right) and the numbers of combinatorial background candidates (bottom left) with errors (bottom right) for various NN requirements and PID offset (y) values. It can be seen that the two requirements are not strongly correlated. Figure 4.10 shows distributions, of the Punzi FoM as a function of NN requirement (top left) and as a function of PID offset (top right) for the optimal values of the PID offset and NN requirement, respectively. It also shows the

signal efficiency (in the signal region) as a function of NN requirement (bottom left) and of PID offset (bottom right) in the corresponding bins of PID offset and NN requirement, respectively.

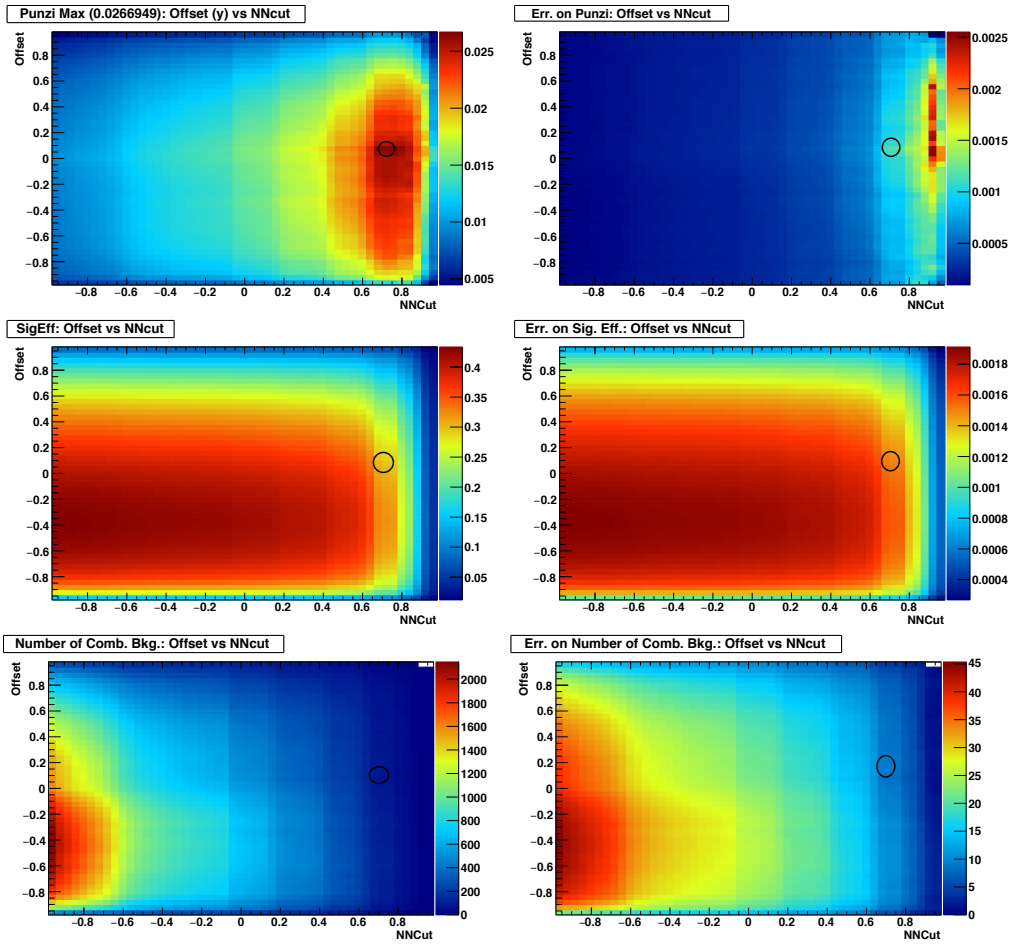

Figure 4.9: Scans of Punzi FoM (top left) with associated errors (top right), efficiency of the selection in the signal region (middle left) with errors (middle right) and number of combinatorial background candidates (bottom left) with errors (bottom right) as function of NEUROBAYES output (NN) requirement and PID offset (y). The black circle indicates the optimal point.

From these plots, I choose requirements of PID offset $y = 0.1$, and a NN requirement value of 0.7. All of the offline selection requirements (MVA and PID) until this stage are summarised in Table 4.7.

Cross-feed rejection: The possible cross-feed contribution, which comes from the mis-identification of one or more final state tracks can be reduced by imposing PID requirements on the final state tracks. As already discussed, the proton PID requirement (`PROBNNp`$\times$`(1-PROBNNK)`

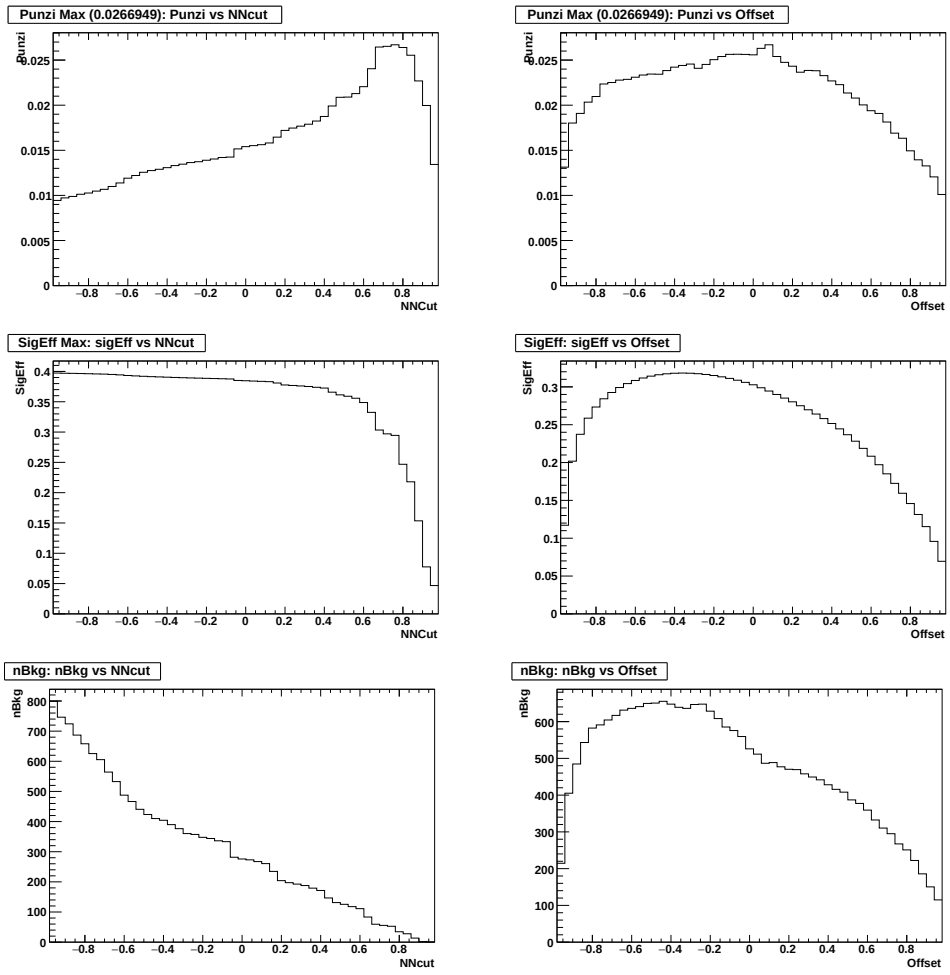

Figure 4.10: Projection of the slices of PID offset and NN requirement containing maximum Punzi value showing Punzi FoM (top), efficiency of the selection in the signal region (middle) and number of background candidates (bottom) as a function of NN requirement (left) and PID offset, y (right) respectively.

> 0.5), is intended to curb possible $B^- \to h^- h'^- h'^+$ cross-feed into the pKK sample. Hence the major cross-feed contribution with single and double mis-identification of the mesons towards my unobserved $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal modes comes from each other.

We expect that all the $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes apart from the $\Xi_b^- \to pK^-K^-$, $\Xi_b^- \to pK^-\pi^-$ & $\Omega_b^- \to pK^-K^-$ decay channels are highly suppressed, thus we do not need to consider cross-feeds from them. Moreover, with the PID requirements obtained in the previous subsection, K tracks are selected with a stronger π rejection than vice-versa. Consequently, the single misidentification cross-feed contribution of $\pi \to K$ is lower than the $K \to \pi$ contribution.

Table 4.7: Optimal selection requirements chosen for $p\,h^-h'^-$ samples. The kaon and pion tracks also have an additional requirement $\sqrt{\texttt{PROBNNK}^2 + \texttt{PROBNNpi}^2} > 0.02$.

Description	Requirements
NEUROBAYES output requirement	NN > 0.7
Kaon(K) PID requirement	$\texttt{PROBNNK} > \texttt{PROBNNpi} + 0.1$
Pion(π) PID requirement	$\texttt{PROBNNpi} > \texttt{PROBNNK} - 0.1$
Proton(p) PID requirement	$\texttt{PROBNNp} \times \texttt{(1-PROBNNK)} > 0.5$

Therefore, considering only signal modes for which I expect non-zero signal yield, the cross-feed contribution of $\Xi_b^- \to pK^-\pi^-$ mode towards $\Omega_b^- \to pK^-K^-$ signal mode will be suppressed compared to the $\Omega_b^- \to pK^-K^-$ cross-feed contribution towards $\Xi_b^- \to pK^-\pi^-$ signal mode. However the production (or fragmentation) factor, $f_{\Omega_b^-}$, of $\Omega_b^- \to pK^-K^-$ decay mode is low compared to its Ξ_b^- counterpart (as a rough estimation, it could be assumed to be similar to $\frac{f_s}{f_d}$ as in the mesonic sector [122] *i.e.* $\sim \frac{1}{4}$). Hence, the cross-feed contribution of $\Omega_b^- \to pK^-K^-$ towards $\Xi_b^- \to pK^-\pi^-$ signal mode is also small. Due to these reasons, the PID requirement on the like-charged mesons is only optimised for combinatorial background rejection and the PID offset (y) that was picked during the optimisation procedure are kept as the final selection requirements. For a full discussion on the background contributions please refer to Sec. **??** for more details.

4.2.3.4 Multiple candidates

Here I investigate the fraction of events that have multiple candidates. The number of events (in percent) with multiple candidates in the data sample after the full selection is given in Table 4.8. This table shows that only KKK and $p\pi\pi$ final states contain events with 2 candidates. However, the fraction of events that contain multiple candidates in these samples is not high when compared to other analyses where this posed a problem [102, 104] and hence I do not take any action *i.e.* all candidates are kept and treated equally.

4.3 Efficiency of the signal selection

To measure the relative branching fraction of $\Xi_b^-(\Omega_b^-) \to phh'^-$ with respect to $B^+ \to K^+K^-K^+$ decays, knowledge of the relative efficiency is required as can be seen in Eq. 4.2. In this section I present the efficiency of the selection procedure that has been outlined in the previous section.

Table 4.8: Fraction of events containing multiple candidates (in percent) for each $p\ h^-h'^-$ final state.

Channel	Fraction of events that contain 1 candidate (%)	Fraction of events that contain 2 candidate (%)	Total number of events
pKK	100.00	0.00	408
$pK\pi$	100.00	0.00	1733
$p\pi\pi$	99.95	0.05	2107
KKK	99.98	0.01	55965

The total efficiency can be factorised into components as

$$\epsilon^{\text{tot}} = \epsilon^{\text{geom}} \times \epsilon^{\text{sel|geom}} \times \epsilon^{\text{PID|sel\&geom}}, \tag{4.6}$$

where

- ϵ^{geom} is the geometrical efficiency, and is determined from MC samples with requirements on the initial state particles (b-hadron) to be within the LHCb acceptance. This efficiency varies as a function of track kinematics and hence depends on the position in phase space. This efficiency is presented for the signal and normalisation modes in Sec. 4.3.1.
- $\epsilon^{\text{sel|geom}}$ is the analysis selection efficiency given the geometrical efficiency and is calculated from the standard MC samples, which are generated with the requirement that all the initial and final state particles are within the LHCb acceptance. It includes contributions from reconstruction, stripping, trigger and offline selections. This efficiency is presented in Sec. 4.3.2 for signal and normalisation modes.
- $\epsilon^{\text{PID|sel\&geom}}$ is the Particle Identification (PID) efficiency given selection along with geometrical efficiency and is determined from data using tools and samples provided by the `PIDCalib` package [121]. This efficiency is presented in Sec. 4.3.3 for signal and normalisation modes.

An ideal analysis would select candidates evenly from any region of the phase space, but in reality the geometry, analysis, trigger and PID all introduce small biases to favour some parts of the phase space over others. To correct for this, the efficiencies of individual sources are calculated as a function of the phase space position.

Due to the presence of spin-$\frac{1}{2}$ particles in the initial and final state, the kinematics of the signal decays that are being studied are in principle described by a five-dimensional phase space that can be described by two Dalitz plot variables and three angular variables, which describe the orientations of different decay planes. The angular variables, shown in Figure 4.11, are described in the Ξ_b^- (Ω_b^-) rest frame, and defined by the axes

- $\hat{x}\prime$: direction of Ξ_b^- (Ω_b^-) ($p_{\Xi_b^-(\Omega_b^-)}$) in the lab frame,
- $\hat{z}\prime$: cross product of the beam axis and $\hat{x}$ in the lab frame,
- $\hat{y}\prime$: cross product of $\hat{z}$ and $\hat{x}$ in the lab frame,

where the beam axis is taken to be in the positive $\hat{z}_{\text{lab}}$ direction in the lab frame. The Euler angles that define how the decay plane of the 3-body decay is rotated in the rest frame of the Ξ_b^- (Ω_b^-) are then

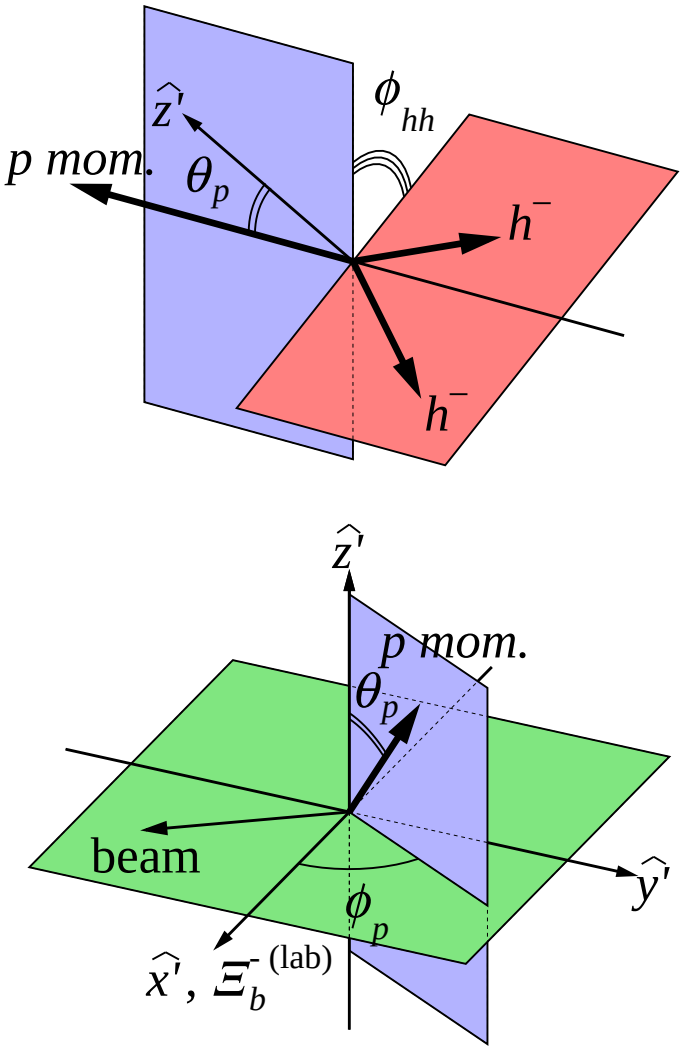

Figure 4.11: Pictorial representation of angular variables which are defined in the ($x\prime$, $y\prime$,$z\prime$) frame that has been defined in the text.

- $\cos(\theta_p)$: cosine of the polar angle of the p momentum in the above reference frame,
- ϕ_p: azimuthal angle of the p momentum in the above reference frame,
- $\phi_{hh'}$: the angle between the plane formed by the h and h' momenta, and the plane formed by the p momentum and $\hat{z}$.

Note that the choice of the five variables to characterise the phase space of the decays under study is dependent on the analysis. We choose here the above mentioned five variables as the phase space variables (one can also define a different set of angular variables as discussed in the next chapter in Sec. 5.4 also).

A feature of Dalitz plots describing decays of heavy flavoured particles is that the signal candidates lie at the corners of the DP regions. These are also the regions where the largest efficiency variations are often expected, as the edges of the Dalitz plot correspond to at least one of the final state particles having low momentum. Large variations occurring over small areas of the Dalitz plot are difficult to describe in detail, hence the Dalitz plot variables (m^2_{13} & m^2_{23}) are transformed to Square Dalitz Plot (SDP) variables [123, 124] (m' and θ', which are both defined in the ranges 0 to 1) as shown in Eq. 4.8. This enables accurate modelling of the efficiency variation and background PDF at the curved edges of the DP. For the case of $\Xi_b^-(\Omega_b^-) \to h^-h'^-p$ where h^-, $h'^- \in \{\pi, K\}$, the transformation of the volume element from conventional DP to SDP is given as follows:

$$dm^2(h^-p)\, dm^2(h'^-p) \to |det\mathcal{J}| dm'\, d\theta' \,. \tag{4.7}$$

Here $\mathcal{J}$ is the Jacobian of the transformation, with m' and θ' given by

$$m' = \frac{1}{\pi}\arccos\left(2\frac{m_{h^-h'^-} - (m_{h^-} + m_{h'^-})}{m_{\Xi_b^-(\Omega_b^-)} - (m_{h^-} + m_{h'^-} + m_p)} - 1\right), \tag{4.8}$$

$$\theta' = \frac{1}{\pi}\arccos\left(\frac{m^2_{h^-h'^-}(m^2_{h'^-p} - m^2_{h^-h'^-}) - (m^2_{h'^-} - m^2_{h^-})(m^2_{\Xi_b^-(\Omega_b^-)} - m^2_p)}{\sqrt{(m^2_{h^-h'^-} + m^2_{h^-} - m^2_{h'^-})^2 - 4m^2_{h^-h'^-}m^2_{h^-}}\sqrt{(m^2_{\Xi_b^-(\Omega_b^-)} - m^2_p - m^2_{h^-})^2 - 4m^2_{h^-h'^-}m^2_p}}\right). \tag{4.9}$$

Note that $\theta' = \frac{1}{\pi}\theta$, where θ is the helicity angle of the $h^-h'^-$ system, *i.e.* the angle between the h^- and p direction vectors in the $h^-h'^-$ rest frame. This particular choice of SDP definition is made so that the symmetrisation of the $\pi^-\pi^-p$ and K^-K^-p final states is trivially achieved by folding the SDP at the line $\theta' = 0.5$, *i.e.* restricting the range to $0 \leq \theta' \leq 0.5$. In Figure 4.12, I illustrate how the momentum of the final state particles is correlated to the DP plot position and how different areas of the DP distribution get transformed into the SDP.

A study was carried out to investigate correlations between the five phase-space variables

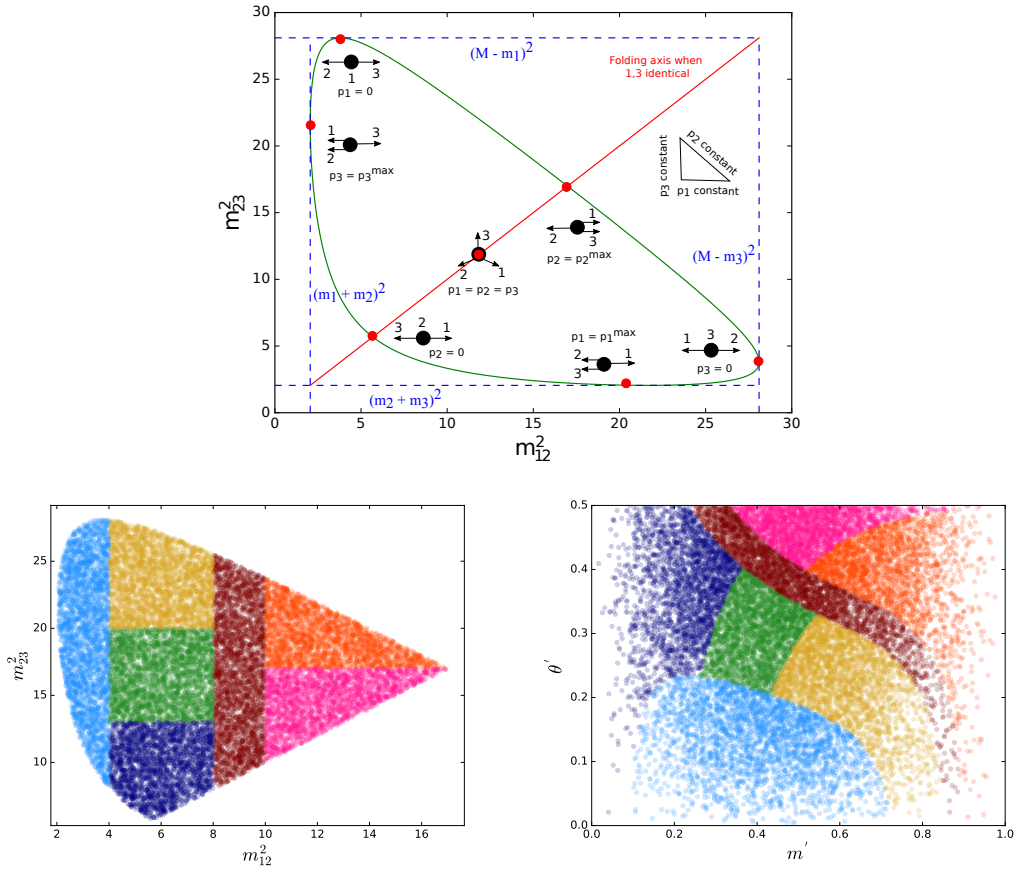

Figure 4.12: (Top) DP areas that correspond to different daughter particle kinematics in an unfolded DP. (Bottom left) Folded DP (along the folding axis) and (bottom right) folded SDP distributions, with colours indicating how different regions (regions with veto and regions with resonances) are mapped by the transformation. The DP is folded when particle 1 and 3 are identical. In the top plot M refers to the mass of the decaying particle and $p_{1,2,3}$, $m_{1,2,3}$ refer to the momenta and masses of the daughter particles, in the rest frame of the parent particle.

for $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes. The Pearson's correlation coefficients for these five phase-space variables, obtained from the signal MC for the $\Xi_b^- \to pK^-K^-$ mode, are shown in Figure 4.13. The square Dalitz plot variables, m' and θ' show a reasonable degree of correlation, as expected, and therefore these will be considered as a 2D efficiency map. It can also be seen that the angular variables have very small correlations with the square Dalitz Plot variables and among themselves. Hence, the efficiency dependence on these angular distributions can be factorised,

$$\epsilon(m', \theta', \cos(\theta_p), \phi_{hh'}, \phi_p) = \epsilon(m', \theta') \times \epsilon(\cos(\theta_p)) \times \epsilon(\phi_{hh'}) \times \epsilon(\phi_p)\,. \tag{4.10}$$

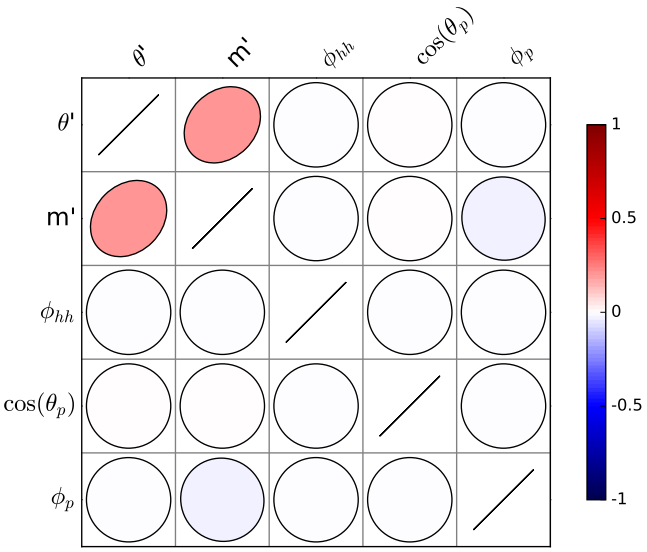

Figure 4.13: Pearson's correlation coefficient between phase-space variables obtained from the $\Xi_b^- \to pK^-K^-$ MC sample.

Furthermore, if the decaying Ξ_b^- (Ω_b^-) is unpolarised, everything should be spherically symmetric. The polarisation of Λ_b^0 baryons produced in $\sqrt{s} = 7\,\text{TeV}$ pp collisions has been measured to be small and consistent with zero [125]; it is likely that the Ξ_b and Ω_b^- polarisation are similarly small but these have not yet been measured. However, it was shown in Ref. [126] which studies decays of $\Lambda_b^0 \to Dp\pi$ (with similar spin structure to our signal $\Xi_b^- \to phh'^-$ decays) that the effect of polarisation on the observables under study would be of second order and for this reason, I do not consider $\epsilon(\cos(\theta_p))$, $\epsilon(\phi_{hh'})$ and $\epsilon(\phi_p)$, and in the following sections will investigate only $\epsilon(m', \theta')$.

4.3.1 Geometrical efficiency

The generator-level efficiency corresponds to the requirement that each of the charged tracks of a given decay must fall within the $10 < \theta < 400\,\text{mrad}$ acceptance of the LHCb detector. The efficiency can in principle vary between the magnet polarities since the beams have a different crossing angle in each of the magnet configurations. The value may also vary between $\sqrt{s} = 7$ and $8\,\text{TeV}$ corresponding to the conditions for 2011 and 2012 data taking, respectively. It can also vary between different b-hadron decays, and vary across the phase space of any given multibody decay.

As indicated in Eq. 4.2, the results of this analysis will be the product of the hadronisation

fraction ratio $\frac{f_{\Xi_b^-(\Omega_b^-)}}{f_u}$ and the ratio of branching fractions $\frac{\mathcal{B}(\Xi_b^-(\Omega_b^-)\to hhp)}{\mathcal{B}(B^-\to KKK)}$. I choose the former ratio to be defined within the LHCb acceptance and hence not to correct for the efficiency for the different b-hadrons to fall within that acceptance. I do, however, need to consider the ratio of efficiencies for the decay products of the b-hadrons to then be found within the acceptance of the detector.

Table 4.9 reports the fraction of accepted b-hadron candidates (*i.e.* b-hadron candidates already in the acceptance region, $0 < \theta < 400\,\mathrm{mrad}$) whose charged daughters all lie in the acceptance region of $10 < \theta < 400\,\mathrm{mrad}$ (ϵ^{geom}) for both configurations of magnet polarity for both 2011 and 2012 Monte Carlo. It also shows the weighted average of this number for both year and both magnet polarity configurations, weighted according to the number of candidates in each sub-sample.

Table 4.9: The fraction of accepted b-hadron candidates whose charged daughters are all in the acceptance region of $10 < \theta < 400\,\mathrm{mrad}$ for signal and normalisation modes (ϵ^{geom}).

Mode	PYTHIA version	ϵ^{geom} (%)				Weighted average ϵ^{geom} (%)
		2011		2012		
		MagUp	MagDown	MagUp	MagDown	
$\Xi_b^- \to pK^-K^-$	8	53.53 ± 0.28	53.53 ± 0.28	53.64 ± 0.28	53.64 ± 0.28	53.61 ± 0.16
$\Xi_b^- \to pK^-\pi^-$	8	52.35 ± 0.28	52.35 ± 0.28	51.99 ± 0.28	51.99 ± 0.28	52.09 ± 0.15
$\Xi_b^- \to p\pi^-\pi^-$	8	51.05 ± 0.27	51.05 ± 0.27	50.94 ± 0.27	50.94 ± 0.27	50.97 ± 0.15
$\Omega_b^- \to pK^-K^-$	8	53.23 ± 0.28	53.23 ± 0.28	52.74 ± 0.29	52.74 ± 0.29	52.87 ± 0.16
$\Omega_b^- \to pK^-\pi^-$	8	51.86 ± 0.28	51.86 ± 0.28	51.55 ± 0.29	51.55 ± 0.29	51.64 ± 0.16
$\Omega_b^- \to p\pi^-\pi^-$	8	50.61 ± 0.27	50.61 ± 0.27	50.10 ± 0.27	50.10 ± 0.27	50.23 ± 0.15
$B^+ \to K^+K^-K^+$	8	53.63 ± 0.28	53.63 ± 0.28	53.29 ± 0.28	53.29 ± 0.28	53.38 ± 0.15

The variation of the geometrical efficiency ϵ^{geom} across the phase space is investigated with dedicated Monte Carlo samples. The results are shown in Figure 4.14 for Ξ_b^- modes, Ω_b^- modes and for the normalisation mode. I show the interpolation (using a 2D cubic spline) of the efficiency as a function of SDP position to smooth out fluctuations due to limited statistics. The geometrical efficiency (for b-hadron decays within the LHCb acceptance) across the 2D SDP varies between 48%–60% ($\sim \pm$ 10% relative variation). Higher efficiency is found for b-hadron decays with higher m' values (corresponding, for the signal modes, to higher proton momentum in the rest frame of the b-hadron candidate).

4.3.2 Selection efficiencies

In this sub-section, I present the phase-space integrated selection efficiency ($\epsilon^{\mathrm{sel|geom}}$) for my signal and normalisation modes. These efficiencies are obtained from MC samples with the

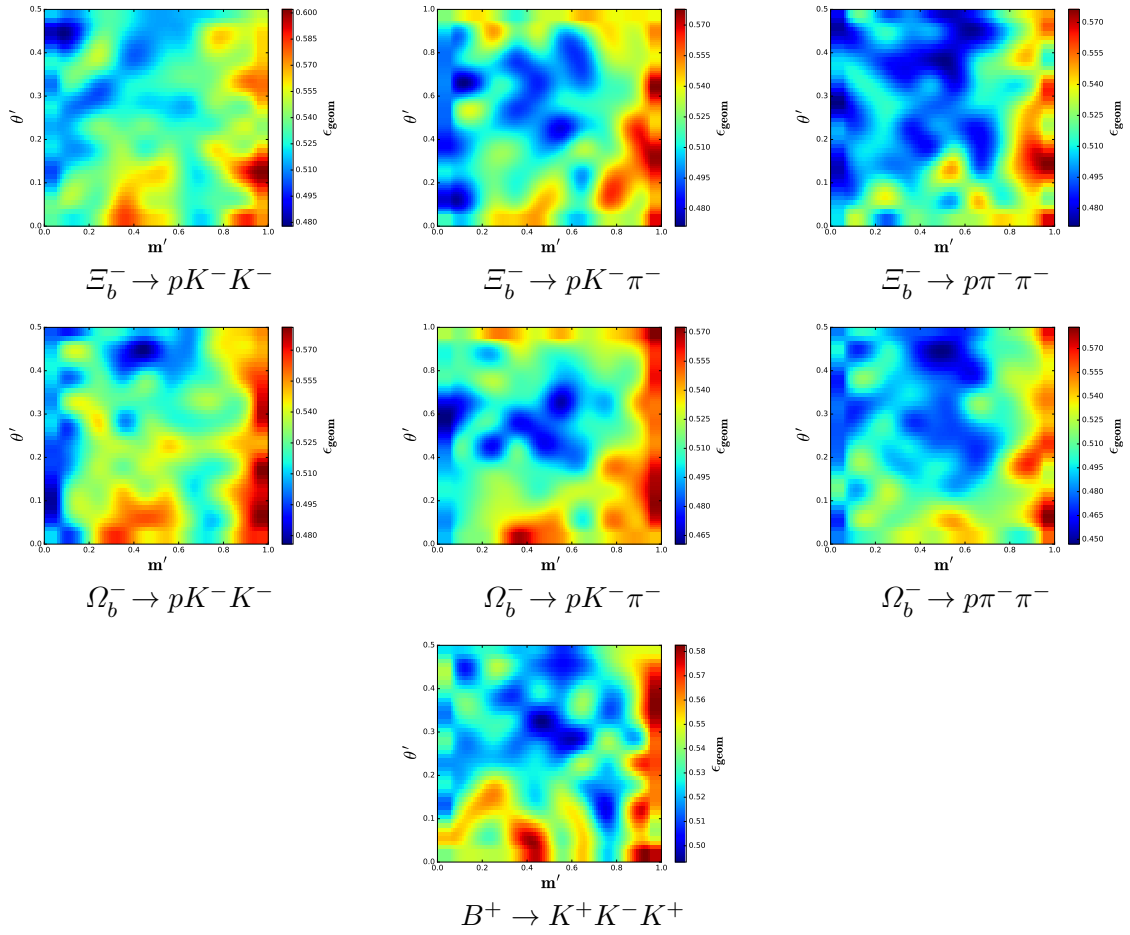

Figure 4.14: Cubic 2D spline interpolation of the geometrical efficiency (ϵ^{geom}) as a function of SDP position for signal and normalisation modes.

requirement that both initial and final state particles are in the LHCb acceptance. As mentioned previously, it includes contributions from reconstruction, stripping, trigger and offline selections, which comprise MVA requirement and any vetoes applied.

Table 4.10 shows the selection efficiencies for different sub-samples and the weighted average (weighted according to the statistics available in each sub-sample).

Table 4.10: Phase-space integrated selection efficiencies ($\epsilon^{\text{sel|geom}}$) for signal ($\Xi_b^- (\Omega_b^-) \to phh'^-$) and $B^+ \to K^+K^-K^+$ modes. We also show the average $\epsilon^{\text{sel|geom}}$ weighted according to the integrated luminosity collected during the data taking period. Statistical uncertainties are also given.

Mode	$\epsilon^{\text{sel|geom}}$ (%)				Weighted average $\epsilon^{\text{sel|geom}}$ (%)
	2011		2012		
	MagUp	MagDown	MagUp	MagDown	
$\Xi_b^- \to pK^-K^-$	2.214 ± 0.015	2.229 ± 0.015	1.825 ± 0.010	1.814 ± 0.010	1.927 ± 0.006
$\Xi_b^- \to pK^-\pi^-$	2.191 ± 0.015	2.190 ± 0.015	1.789 ± 0.009	1.791 ± 0.009	1.896 ± 0.005
$\Xi_b^- \to p\pi^-\pi^-$	2.559 ± 0.022	2.574 ± 0.022	2.108 ± 0.014	2.096 ± 0.015	2.226 ± 0.009
$\Omega_b^- \to pK^-K^-$	1.765 ± 0.013	1.785 ± 0.013	1.414 ± 0.008	1.431 ± 0.008	1.517 ± 0.005
$\Omega_b^- \to pK^-\pi^-$	1.673 ± 0.013	1.661 ± 0.013	1.357 ± 0.008	1.365 ± 0.008	1.442 ± 0.005
$\Omega_b^- \to p\pi^-\pi^-$	1.747 ± 0.019	1.719 ± 0.018	1.423 ± 0.012	1.425 ± 0.012	1.505 ± 0.007
$B^+ \to K^+K^-K^+$	2.763 ± 0.023	2.725 ± 0.023	2.304 ± 0.015	2.290 ± 0.015	2.415 ± 0.009

The variation of $\epsilon^{\text{sel|geom}}$ over the SDP for $\Xi_b^- \to phh'^-$, $\Omega_b^- \to phh'^-$ signal Monte Carlo and for $B^+ \to K^+K^-K^+$ MC can be found in Figure 4.15. The main features seen in these figures are investigated and found to be due to vetoes, introducing significant efficiency drops (*i.e.* to zero). Drops in efficiency can also be seen in regions corresponding to one of the final state particles having low momentum in the b-hadron candidate rest frame (*e.g.* low m' corresponds to low proton momentum for the signal modes). The positive and negative uncertainties on the efficiencies due to limited MC statistics are evaluated using the Clopper-Pearson method and will be considered in the evaluation of the systematic uncertainty.

4.3.3 PID efficiency

The PID efficiency is calculated using calibration samples from the `PIDCalib` package, as described in Sec. 4.2.3.3. The distributions of PID efficiencies of identifying (top) a kaon as a kaon, (middle) a pion as a pion and (bottom) a proton as a proton as functions of the particle's p, p_{T} and `nTracks` (also referred to as Number of 'BestTracks' in an event)[5], obtained from

[5]For a given event, `nTracks` is defined as the number of tracks in the best container. The best container picks the best track type (out of four track types *i.e.* VELO, T, upstream and downstream; see Fig. 3.6) that a given track in an event can be assigned to.

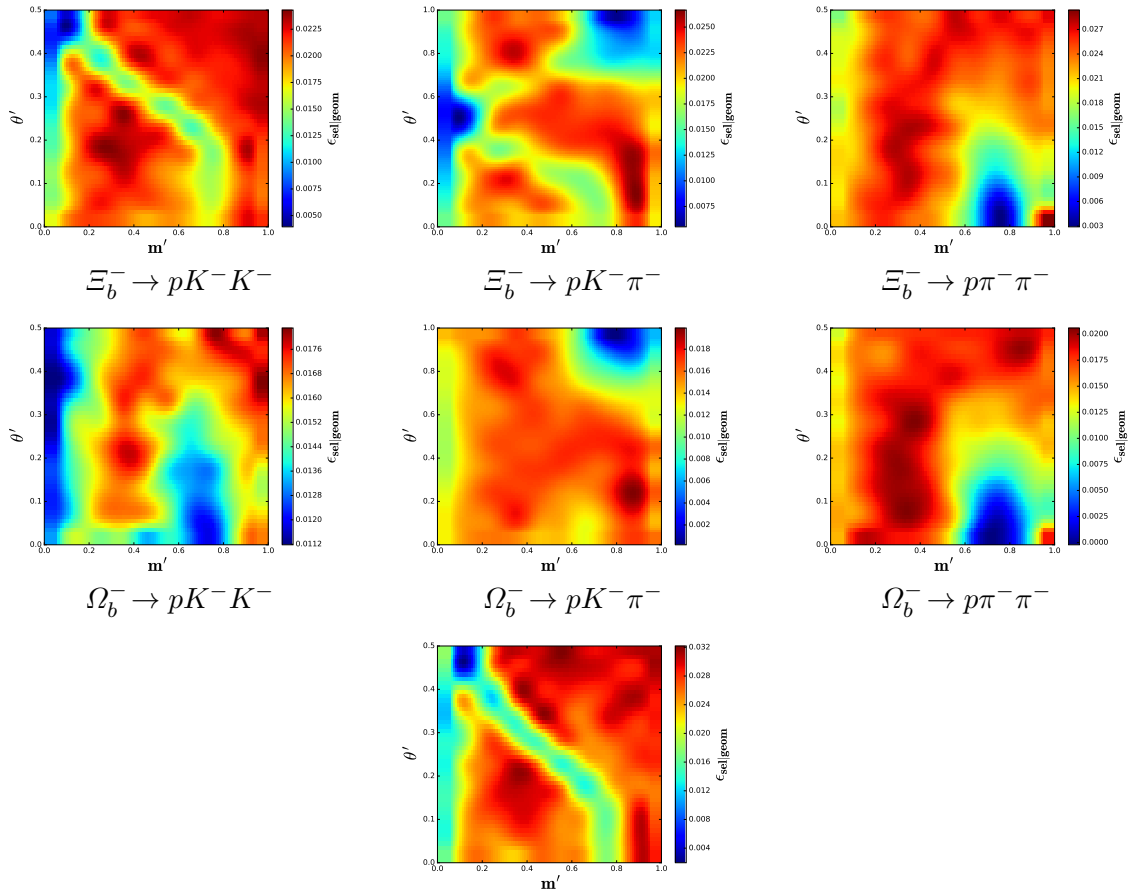

Figure 4.15: Cubic 2D spline interpolation of the selection efficiency ($\epsilon^{\mathrm{sel|geom}}$) as a function of SDP position for signal and normalisation modes.

the PIDCalib package for the optimal PID requirements discussed in Sec. 4.2.3.3, are shown in Figure 4.16. The single particle mis-identification rates associated with the selected PID requirements for mis-identifying $\pi \to K$, $K \to \pi$, $\pi \to p$ and $K \to p$ are approximately $0.8\,\%$, $9\,\%$, $0.2\,\%$ and $0.2\,\%$ respectively[6]. The rates associated with double particle mis-identification for $\pi\pi \to KK$, $KK \to \pi\pi$, $KK \to \pi p$, $\pi\pi \to Kp$ and $\pi K \to Kp$ are approximately $0.02\,\%$, $2\,\%$, $0.05\,\%$, $0.005\,\%$ and $0.05\,\%$ respectively.

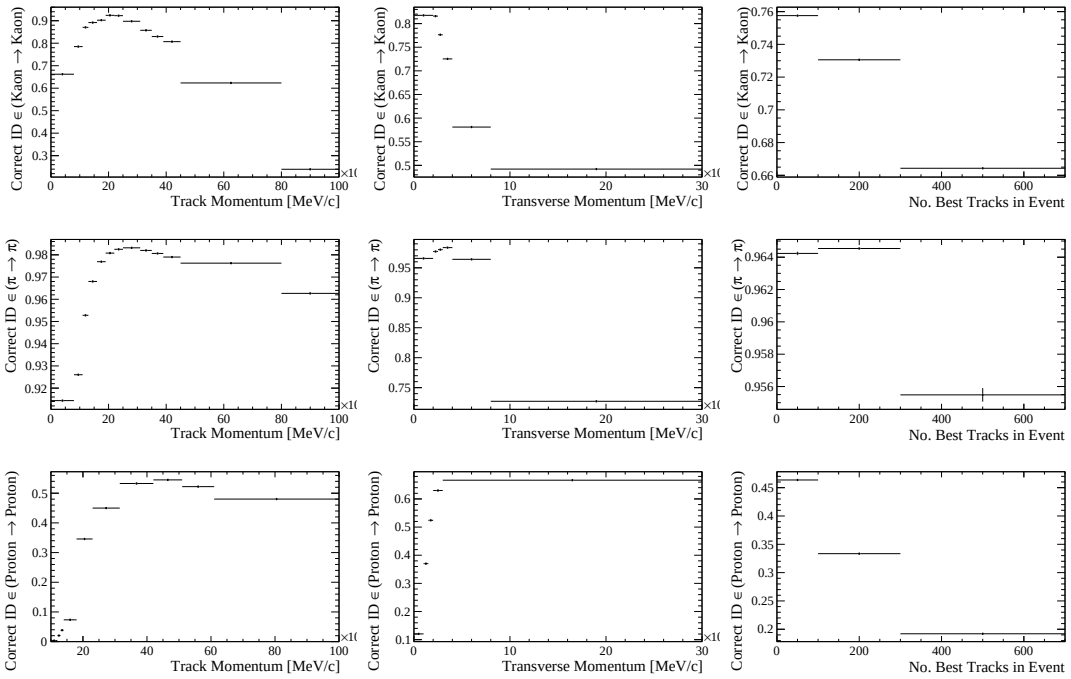

Figure 4.16: Distributions of PID efficiencies of kaon (top), pion (middle) and proton (bottom) tracks as a function of momentum (left), transverse momentum (centre) and track multiplicity of events (right) for calibration samples provided in the PIDCalib package for the value of $y = 0.1$.

Table 4.11 lists the combined (2011 + 2012) phase-space integrated and calibrated PID efficiencies for the finalised PID requirements, along with the muon identification veto applied to each final state charged track. In this table I also quote the PID efficiencies for the $B^- \to h^- h'^- h'^+$ modes because, as will be seen later, using these values I will be constraining the expected yields of the cross-feed from these modes in the fit to the normalisation and signal invariant mass distributions. This table also gives the breakdown of the statistical uncertainty in

[6]Note that these values are obtained using $\Xi_b^- \to phh'^-$ signal kinematics. Also only values which are used in the analysis are quoted here.

the one from calibration sample size (referred to in the table as 'CalibSampleStat') and that due to the size of the reference MC sample, from which PID efficiencies are obtained (referred in the table as 'RefSampleStat'). I also quote the combination in quadrature of the uncertainties.

To obtain the error due to calibration sample size on the phase-space integrated PID efficiency, I generate 1000 different efficiency histograms where the efficiency in each bin of p, p_{T} & `nTracks` is a random value drawn from a Gaussian where the mean (μ) and the width (σ) is the bin content and bin error of the default performance histogram produced using the `PIDCalib` package. I then calculate the average efficiency for my reference MC sample according to each of the smeared histograms. The width obtained from a χ^2 fit of a Gaussian to the resulting efficiency distribution is the statistical uncertainty due to the calibration sample size.

Uncertainties due to the reference MC sample size are estimated using the bootstrap resampling method [127] in which the `PIDCalib` efficiency is repeatedly evaluated on datasets of p, p_{T} and `nTracks` that are sampled (with replacement) from the original data set (about 1000 bootstrap samples were used).

The variation of $\epsilon^{\mathrm{PID|sel\&geom}}$ in the SDP for Ξ_b, Ω_b^- and B^+ modes can be found in Figure 4.17. For all modes, the PID efficiency varies between 12%–60% across the 2D SDP . In the PID efficiency maps for the signal and normalisation modes, larger efficiency for b-hadron candidates with higher m' values (corresponding to higher proton or kaon momentum in the rest frame of the b-hadron candidate) are observed.

Table 4.11: Phase-space integrated PID efficiencies for signal ($\Xi_b^-(\Omega_b^-) \to phh'^-$) and $B^- \to h^-h'^-h'^+$ modes. Quoted also are the statistical uncertainty from reference sample size ('RefSampleStat'), from the calibration sample size ('CalibSampleStat') and the combined uncertainties ('Combined').

Modes	ϵ_{PID} (%)	$\delta\epsilon_{\mathrm{RefSampleStat}}$ (%)	$\delta\epsilon_{\mathrm{CalibSampleStat}}$ (%)	$\delta\epsilon_{\mathrm{Combined}}$ (%)
$\Xi_b^- \to pK^-K^-$	32.60	0.06	0.04	0.07
$\Xi_b^- \to pK^-\pi^-$	40.12	0.07	0.06	0.09
$\Xi_b^- \to p\pi^-\pi^-$	50.33	0.10	0.06	0.12
$\Omega_b^- \to pK^-K^-$	33.02	0.08	0.05	0.09
$\Omega_b^- \to pK^-\pi^-$	40.09	0.08	0.05	0.10
$\Omega_b^- \to p\pi^-\pi^-$	49.86	0.13	0.07	0.14
$B^+ \to K^+K^-K^+$	46.74	0.06	0.02	0.06
$B^+ \to K^+\pi^-K^+$	56.94	0.07	0.02	0.07
$B^+ \to K^+\pi^-\pi^+$	71.13	0.06	0.02	0.06
$B^+ \to \pi^+\pi^-\pi^+$	88.90	0.02	0.01	0.02

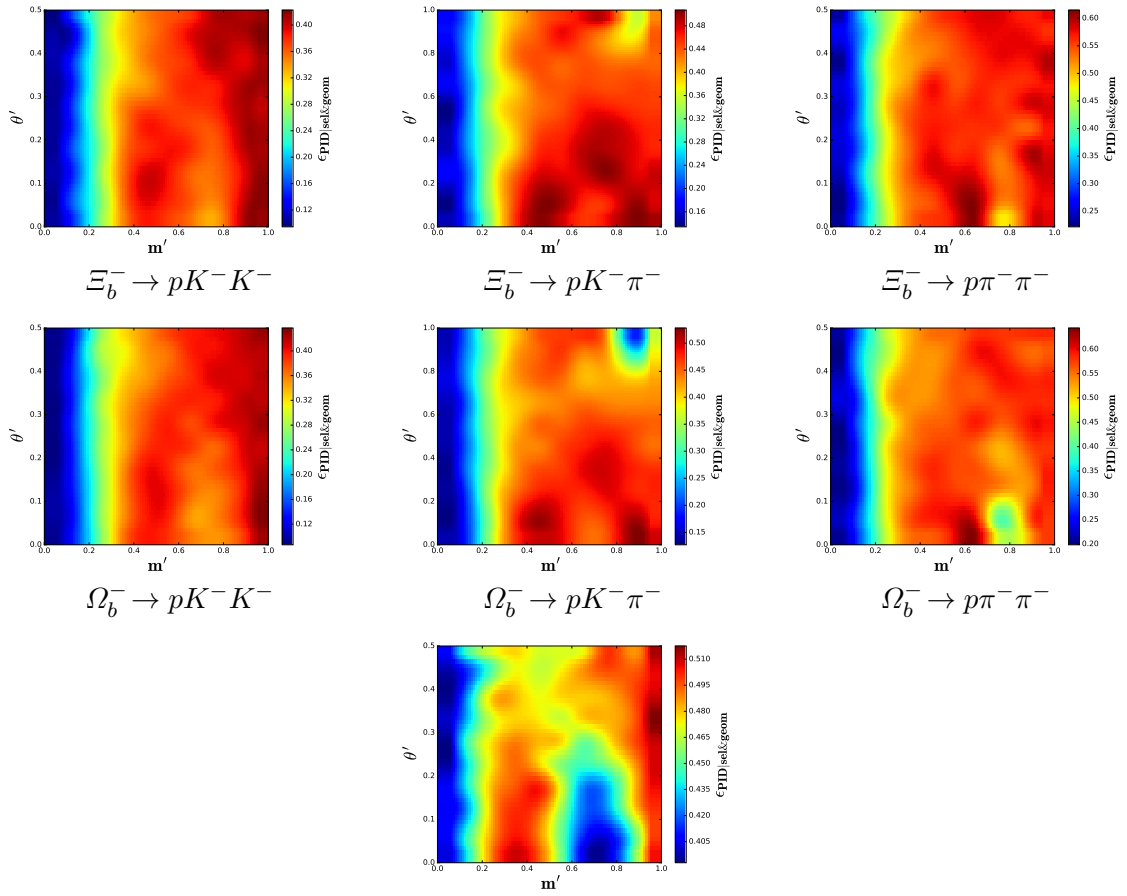

Figure 4.17: Cubic 2D spline interpolation of the PID efficiency ($\epsilon^{\mathrm{PID|sel\&geom}}$) as a function of SDP position for signal and normalisation modes.

4.3.4 Total efficiency

The phase-space integrated total efficiencies ($\epsilon^{\rm tot}$), evaluated using Eq. 4.6, for signal and normalisation modes are shown in Table 4.12.

The lifetimes for the Ξ_b^- and Ω_b^- used in the MC productions are

$$\begin{aligned} \tau(\Xi_b^-) &= 1.567\,\text{ps}\,, \\ \tau(\Omega_b^-) &= 1.135\,\text{ps}\,. \end{aligned}$$

Since for the lifetimes are quite different from the most up-to-date measurements (*i.e.* for Ξ_b^- and Ω_b^- the lifetimes measured are 1.560 ± 0.040 and 1.66 ± 0.18 respectively), I apply a correction to the efficiency and a corresponding systematic uncertainty related to the precision on the measured lifetime (see Sec. 4.6.1.1). Therefore, I also quote in this table the lifetime-corrected total efficiencies ($\epsilon^{\rm tot}_{\rm ltcorr}$).

Table 4.12: Phase-space integrated total efficiencies ($\epsilon^{\rm tot}$) for signal ($\Xi_b^-(\Omega_b^-) \to phh'^-$) and $B^+ \to K^+K^-K^+$ modes, with uncertainties due to MC statistics. The lifetime-corrected total efficiencies ($\epsilon^{\rm tot}_{\rm ltcorr}$) are also given. The uncertainty in the lifetime correction is accounted for separately (see Sec. 4.6.1.1).

Modes	$\epsilon^{\rm tot}$ (%)	$\delta\epsilon$ (%)	$\epsilon^{\rm tot}_{\rm ltcorr}$ (%)	$\delta\epsilon^{\rm tot}_{\rm ltcorr}$ (%)
$\Xi_b^- \to pK^-K^-$	0.3368	0.0016	0.3378	0.0016
$\Xi_b^- \to pK^-\pi^-$	0.3962	0.0019	0.3975	0.0019
$\Xi_b^- \to p\pi^-\pi^-$	0.5339	0.0029	0.5355	0.0029
$\Omega_b^- \to pK^-K^-$	0.2649	0.0014	0.3740	0.0019
$\Omega_b^- \to pK^-\pi^-$	0.2985	0.0016	0.4196	0.0022
$\Omega_b^- \to p\pi^-\pi^-$	0.3770	0.0023	0.5353	0.0033
$B^+ \to K^+K^-K^+$	0.6026	0.0030	0.6026	0.0030

The lifetime-corrected total efficiency variation over the SDP for $\Xi_b^- \to phh'^-$, $\Omega_b^- \to phh'^-$ and $B^+ \to K^+K^-K^+$ can be seen in Figure 4.18. For signal and normalisation modes, the total efficiency varies between 0.1%–0.9% across the 2D SDP. In the total efficiency maps the b-hadron candidates with higher m' values (corresponding to higher proton or kaon momentum) have a larger efficiency. This is due to the similar behaviour in geometrical, selection and PID efficiency over SDP variables which was discussed previously.

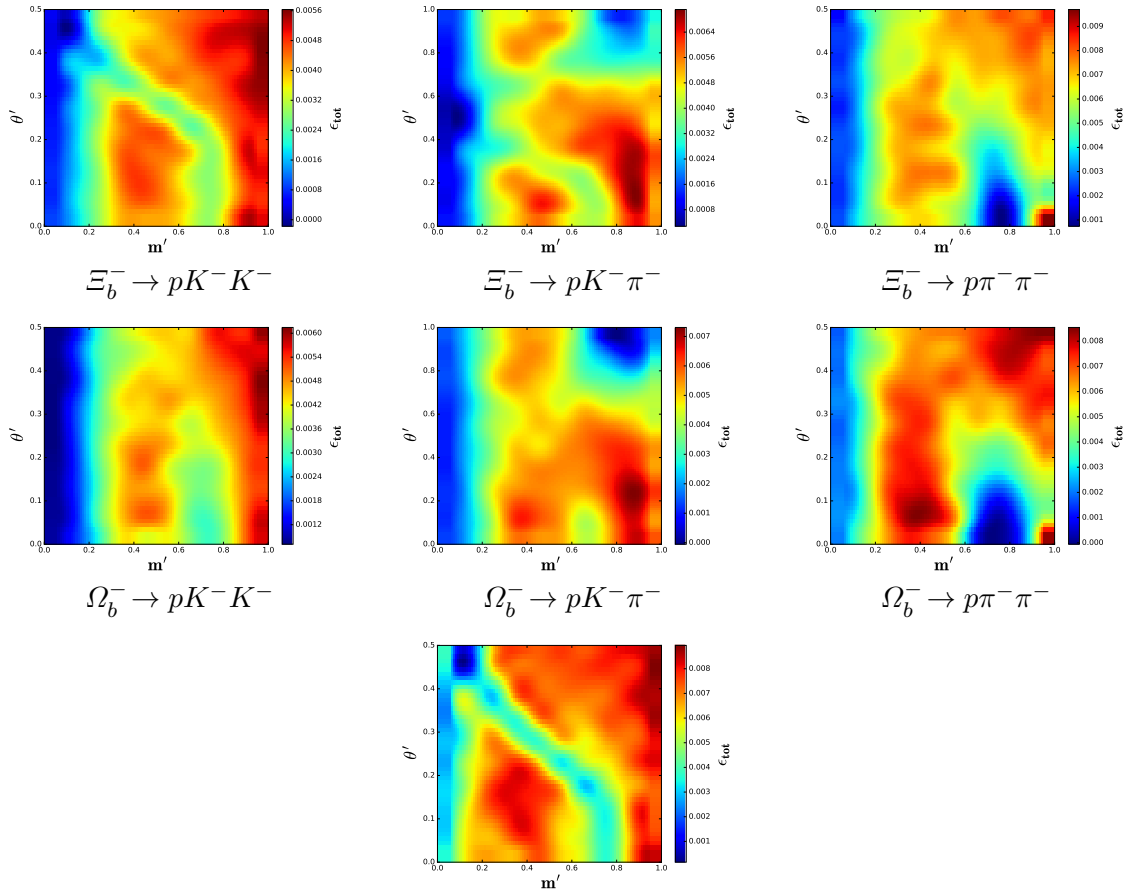

Figure 4.18: Cubic 2D spline interpolation of the lifetime corrected total efficiency ($\epsilon_{\rm corr}^{\rm tot}$) as a function of SDP position for signal and normalisation modes.

4.4 Fit model

To extract the yields for the six signal modes and the normalisation mode, requires building a robust fit model that exhibits negligible bias for all the $p\,h^-h'^-$ and KKK final states, where $h \in \pi, K$. The fit is one dimensional and is performed using the invariant mass variables denoted by $m(ph^-h'^-)$. In this section, I present the construction of such a fit model. In Subsec. 4.4.1, I discuss the choice of the PDF for each of the components to the fit model. I then discuss the construction of a likelihood function from the total PDF and the various constraints imposed on this PDF in Subsec. 4.4.2. I also discuss in Subsec. 4.4.3, the stability of the defined fit model, tested using pseudo experiments.

4.4.1 PDF components of the fit model

The fit model consists of signal and various background components. The various background contributions to the signal and normalisation modes are classed into separate categories as follows:

- **Peaking background**: These are background modes which have been fully reconstructed and which peak around the signal region in the invariant mass distribution of $m(ph^-h'^-)$. The decay modes that contribute towards this class of background typically proceed through an intermediate charm state (X_c) such as $\Xi_b^-(\Omega_b^-) \to X_c^0(ph^-)h'^-$. These backgrounds are vetoed by fully reconstructing the intermediate state from the tracks with 'correct' daughter hypothesis.

- **Partially reconstructed background**: These background modes decay into the same final states as the signal mode ($p\,h^-h'^-$) with an extra particle (X) in the final state that has not been reconstructed. Due to the fact these modes have been partially reconstructed they peak in the low-mass sideband region, shifted at least by the mass of the unreconstructed particle (m_X). Thus, for comparable yields the most problematic missing extra particles are the lightest ones: $\{\gamma, \pi^0, \pi^\pm\}$.

- **Cross-feed**: These contributions arise due to misidentification (mis-ID) of one or more final state particles as some other particle type. The mis-ID causes both a smearing and a shift of the mass peak, due to the change in hypothesis for the misidentified track. Thus, single mis-ID tends to be more of an issue than double mis-ID, not only because the mis-ID probability is low, but also because for multiple mis-ID the background is likely to be shifted outside the fit window and smeared sufficiently that the shape is hard to distinguish from combinatorial background.

- **Combinatorial background**: This arises due to the random combination of tracks that fake the phh candidate. The contributions from these backgrounds are reduced as described in Sec. 4.2.

The expected shapes of each of the components of the fit model will be discussed in Secs. 4.4.1.1– 4.4.1.5. They have been extracted from fits to MC samples with the same selection requirements as applied to the data sample. To account for the discrepancy between the MC and data PID response, I weight the individual distributions with the efficiency obtained from `PIDCalib` for the optimal PID criteria (The PID requirements can lead to significant change in the shape of the tail of the distribution.). I discuss below the PDF of various components of the fit model for signal modes and for the normalisation mode.

4.4.1.1 Parameterisation of the signal and normalisation channel components

The $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal and $B^+ \to K^+K^-K^+$ normalisation mass distributions are modelled as the sum of two Crystal Ball (DCB) functions with a common peak position and width. The Crystal Ball function has been defined in Eq. 4.4. The shape parameters of the DCB function are determined from unbinned likelihood fits to the signal MC, and then fixed in the fit to data. To illustrate the shape of the DCB, fits to the pKK signal MC sample are shown in Figure 4.19. The model parameters extracted from these fits can be found in Table 4.13. Similar fits are conducted to the $B^+ \to K^+K^-K^+$ and other $\Xi_b^-(\Omega_b^-) \to phh'^-$ mass distributions and the fit results are used in the fit to data.

4.4.1.2 Cross-feed components

Cross-feed contributions to the signal modes ($\Xi_b^-(\Omega_b^-) \to phh'^-$) arise from each other *i.e.* one signal mode serves as the cross-feed for the other *e.g.* $\Xi_b^-(\Omega_b^-) \to pK^-K^-$ is cross-feed for $\Xi_b^-(\Omega_b^-) \to pK^-\pi^-$ where $K \to \pi$. Also expect contributions from $B^- \to h^-h'^-h'^+$ modes, where either a kaon or a pion is misidentified as a proton are considered. In the fit model, I include cross-feed components from $\Xi_b^-(\Omega_b^-) \to phh'^-$ and $B^- \to h^-h'^-h'^+$ modes. These are obtained from MC samples weighted by the corresponding mis-ID weights calculated with the `PIDCalib` package. I do not include cross-feed components from $\Xi_b^-(\Omega_b^-) \to p\pi^-\pi^-$ to the pKK final state, because a very low yield for these modes is expected.

One thing to note here is that all the MC samples have been generated flat in the square Dalitz Plot (SDP), *i.e.* the allowed phase space regions of m' and θ' are populated uniformly with candidates. Generally, three-body decay channels exhibit very rich dynamics and this should be accounted for when the cross-feed shapes are obtained. For the $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes the phase space occupancy is not known *a priori*, however the known distributions for the

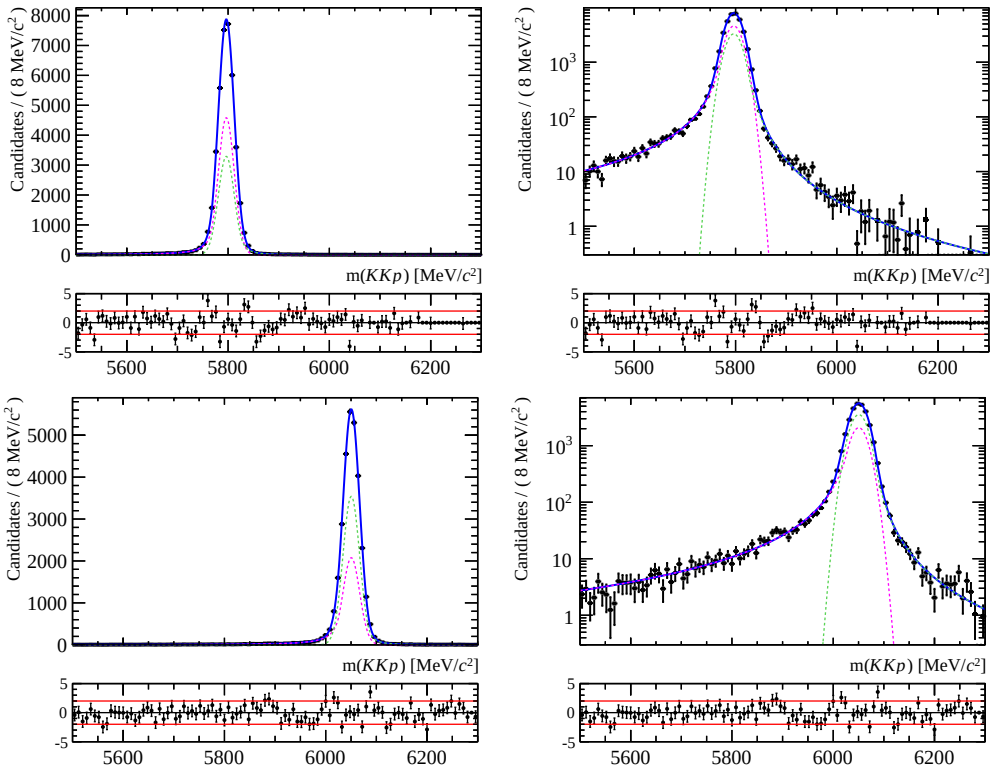

Figure 4.19: Plots with (left) linear and (right) logarithmic scale in y axis for (top) $\Xi_b^- \to pK^-K^-$ and (bottom) $\Omega_b^- \to pK^-K^-$ each modelled with a DCB PDF. The green and pink dashed curves are the individual CB PDFs. The χ^2 per degree of freedom for $\Xi_b^- \to pK^-K^-$ fit is 1.62 and that for $\Omega_b^- \to pK^-K^-$ is 1.51. Similar fits are conducted to the other $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes.

$B^- \to h^-h'^-h'^+$ modes can be taken into account. I do this by obtaining the SDP distributions of $B^- \to h^-h'^-h'^+$ decays from data with the help of signal *sWeights*. I then weight the corresponding MC sample to match the SDP distribution of signal *sWeighted* data. The weights obtained from this procedure are subsequently referred to as 'Dalitz Plot weights' or just 'DP weights'. To check that this procedure gives the correct signal distribution, I plot the SDP variable distributions (m' and θ') for signal *sWeighted* data, an MC sample without any weights, an MC sample with only PID weights from the `PIDCalib` package and an MC sample with PID & DP weights. I also plot the invariant mass of the misidentified $h^-h'^-h'^+$ final state, $m(h^-h'^-h'^+)$, for $B^- \to h^-h'^-h'^+$ MC samples to see how much difference the PID & DP weights make. Figure 4.20 shows these plots for the $B^+ \to K^+K^-K^+$ MC sample only. It can be seen that the m' and θ' distributions for signal *sWeighted* data and the MC sample with PID & DP weights are in good agreement. It can also be seen in this figure that applying PID weights to the MC

Table 4.13: Values of the DCB parameters obtained from fits to pKK MC samples. Similar fit results are obtained for $p\,h^-h'^-$.

Parameters	Values (2011 + 2012)
$\alpha_0(\Omega_b^- \to pK^-K^-)$	1.40 ± 0.18
$\alpha_0(\Xi_b^- \to pK^-K^-)$	1.71 ± 0.10
$\alpha_1/\alpha_0(\Omega_b^- \to pK^-K^-)$	-1.45 ± 0.24
$\alpha_1/\alpha_0(\Xi_b^- \to pK^-K^-)$	-1.06 ± 0.14
$f(\Omega_b^- \to pK^-K^-)$	0.41 ± 0.10
$f(\Xi_b^- \to pK^-K^-)$	0.59 ± 0.10
$\mu(\Omega_b^- \to pK^-K^-)$	6050.50 ± 0.12 MeV/c^2
$\mu(\Xi_b^- \to pK^-K^-)$	5796.50 ± 0.10 MeV/c^2
$n_0(\Omega_b^- \to pK^-K^-)$	1.72 ± 0.10
$n_0(\Xi_b^- \to pK^-K^-)$	1.55 ± 0.07
$n_1/n_0(\Omega_b^- \to pK^-K^-)$	1.34 ± 0.14
$n_1/n_0(\Xi_b^- \to pK^-K^-)$	1.56 ± 0.13
$\sigma(\Omega_b^- \to pK^-K^-)$	16.42 ± 0.11 MeV/c^2
$\sigma(\Xi_b^- \to pK^-K^-)$	15.74 ± 0.09 MeV/c^2

sample changes the shape of the Ξ_b^- (Ω_b^-) candidate mass distribution substantially but the DP weighting has a much smaller effect on the shape.

The backgrounds from mesonic modes and from baryonic modes are parametrised with exponential functions and with DCB PDFs, respectively. The fits to the cross-feeds that contribute towards the pKK final state are shown in Figure 4.21 and the parameters obtained from these fits can be seen in Table 4.14. Similarly fits to obtain shapes of cross-feeds contributing to the $pK\pi$ and $p\pi\pi$ final states have been conducted and used in the fit to data.

For the normalisation mode, I have included in the fit model the cross-feed contribution from single track ($\pi \to K$) mis-ID of $B^+ \to K^+\pi^-K^+$ and the double track mis-ID ($\pi\pi \to KK$) of $B^+ \to K^+\pi^-\pi^+$. The cross-feed contribution to $B^+ \to K^+K^-K^+$ from $B^+ \to K^+\pi^-K^+$ is modelled as a DCB PDF, whereas the cross-feed from the $B^+ \to K^+\pi^-\pi^+$ mode is modelled as a single CB PDF. I also investigated the possible cross-feed contribution from $\Xi_b^- \to pK^-K^-$, however it was found that this component can be accommodated in the combinatorial shape.

4.4.1.3 Partially reconstructed background components

Inspecting the $m(ph^-h'^-)$ distribution in data at different NEUROBAYES output requirements (see Figure 4.22), it can be seen that the shape changes as the MVA requirement is made tighter.

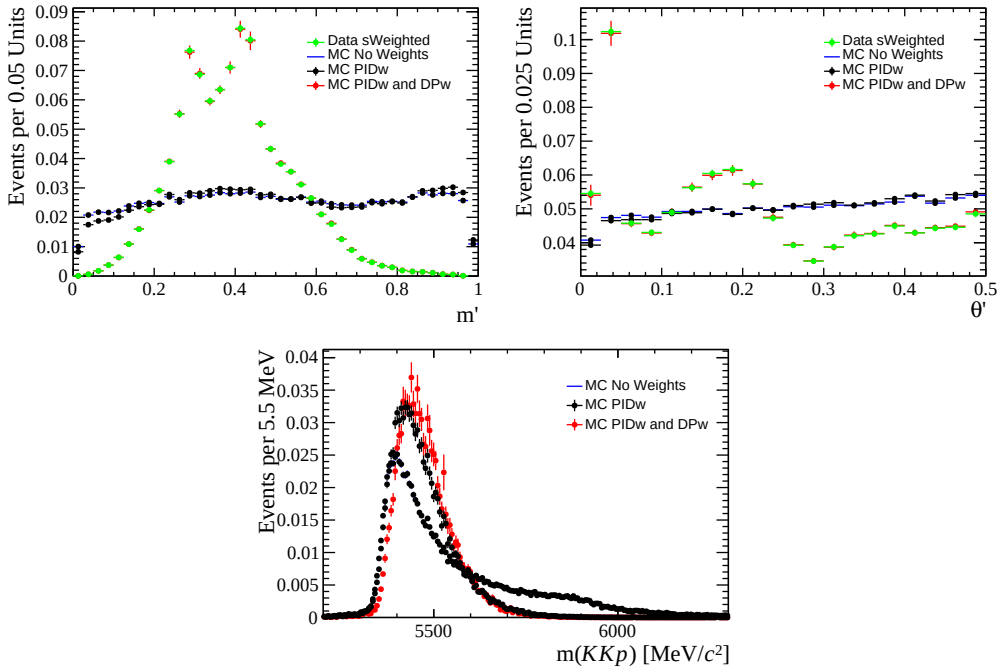

Figure 4.20: (Top) comparison in (left) m' and (right) θ' of the KKK signal *sWeight* data (green), the MC sample without any weights (blue), with PID weights (black) and with PID weights plus DP weights (red). (Bottom) the same for the reconstructed pKK invariant mass.

This implies the presence of Partially Reconstructed Background (PRB) in the data sample (since the peaks seem to be significant with tighter MVA output cuts).

Potential contributions to this category for all $p\,h^-h'^-$ final states could arise due to $\Xi_b^-(\Omega_b^-) \to \Delta(p\pi^0)h^-h'^-$ & $\Xi_b^-(\Omega_b^-) \to N(p\pi^0)h^-h'^-$ modes. There could also be possible contributions to final states containing at least one charged pion from $\Xi_b^-(\Omega_b^-) \to \rho^-(\pi^-\pi^0)hp$ decays, and from $\Xi_b^-(\Omega_b^-) \to K^{*-}(K^-\pi^0)hp$ decays to final states containing at least one charged kaon. Additionally, for each Ξ_b^- decay mode that contributes towards PRB due to a missing pion, there could also be a contribution from the corresponding decay of the isospin partner (Ξ_b^0); the PRB shapes for these decay modes would however be very similar as one cannot distinguish between losing a charged or a neutral pion from the final state cannot be distinguished. The PRB invariant mass distributions for Ξ_b^- decay to a given $p\,h^-h'^-$ final state look very similar and the same is true for PRB shapes from Ω_b^- decays. This motivates the inclusion of two PRB components, one per b-baryon, in the fit model for each of the $p\,h^-h'^-$ final states. However, in this case some of the yield parameters of Ξ_b^- and Ω_b^- PRB components in the fit to data hit negative values (though consistent with zero). Therefore, for simplicity, I only include

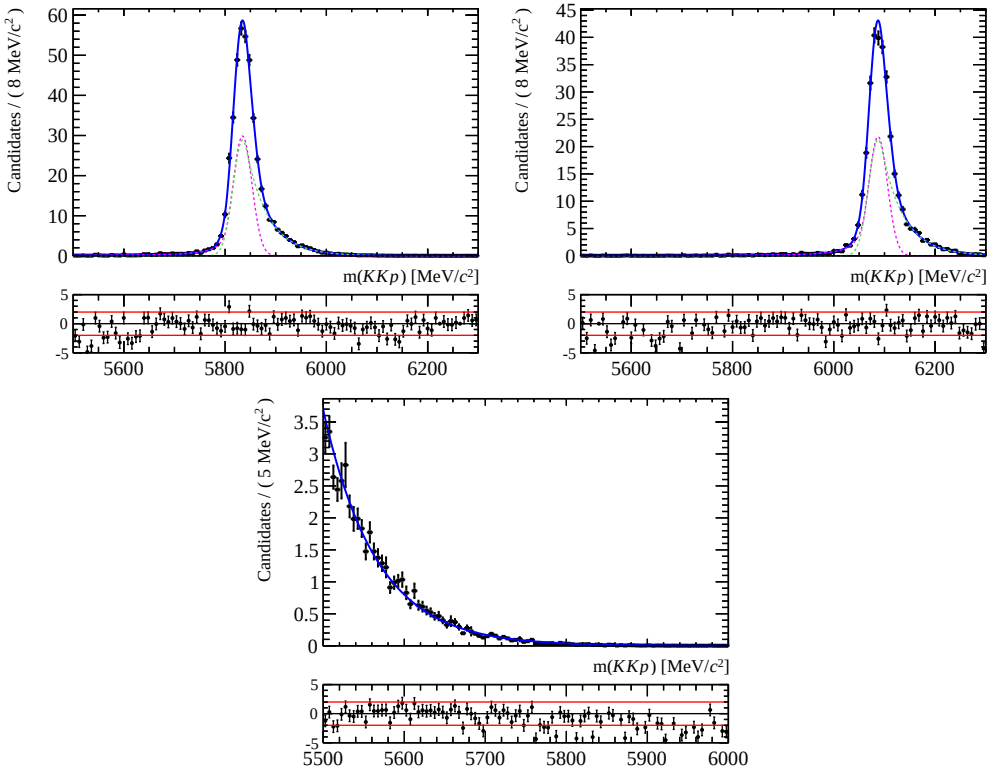

Figure 4.21: Cross-feeds to the pKK final state from the (top left) $\Xi_b^- \to pK^-\pi^-$ and (top right) $\Omega_b^- \to pK^-\pi^-$ modes, modelled as DCB PDFs. The green and pink dashed curves are the individual CB PDFs. The blue represents the fit and the data points is the data. The MC samples undergo the same selection as the data sample. (Bottom) the cross-feed from $B^+ \to K^+K^-K^+$ is modelled with an exponential PDF.

the shapes from $\Xi_b^- \to N(p\pi^0)h'^-h'^-$ modes in my fit model. The other modes will be used to quantify the systematic uncertainty arising from the choice of the fit model.

There is also a possibility of contamination from the $\Xi_b^-(\Omega_b^-) \to \Lambda_b^0(h^-h^+h^-p)\pi^-$ decay channel where two meson tracks are lost from the final state. However this potential background is highly suppressed as the Λ_b^0 decays weakly and so its vertex will be separated from the Ξ_b^- (Ω_b^-) decay vertex, and consequently it is expected to be rejected by the selection criteria. I also considered the contribution from $\Xi_b^-(\Omega_b^-) \to \Sigma(p\overline{K}^0)h^-h'^-$ decay mode. As the unreconstructed particle is $\overline{K}^0$, this contribution would peak well below the signal regions, but could contribute to the pKK final state, if a final state π is misidentified as K. However as this decay is already suppressed and receives an extra suppression from mis-ID, its contribution is very low.

Table 4.14: Values of the parameters for models describing cross-feeds to the pKK final state. The fit results of the cross-feed components contributing to other $p\ h^-h'^-$final state are extracted but not shown here. Note that although the n_1/n_0 parameters are poorly determined, this has little effect on the shape.

Parameters	Values (2011 + 2012)
$c(B^+ \to K^+K^-K^+)$	$(-1.52 \pm 0.22) \times 10^{-2}$
$\alpha_0(\Omega_b^- \to pK^-\pi^-)$	1.44 ± 0.46
$\alpha_0(\Xi_b^- \to pK^-\pi^-)$	1.65 ± 0.38
$\alpha_1/\alpha_0(\Omega_b^- \to pK^-\pi^-)$	-0.30 ± 0.12
$\alpha_1/\alpha_0(\Xi_b^- \to pK^-\pi^-)$	-0.26 ± 0.07
$f(\Omega_b^- \to pK^-\pi^-)$	0.43 ± 0.16
$f(\Xi_b^- \to pK^-\pi^-)$	0.42 ± 0.12
$\mu(\Omega_b^- \to pK^-\pi^-)$	6087.30 ± 1.99 MeV/c^2
$\mu(\Xi_b^- \to pK^-\pi^-)$	5834.50 ± 1.68 MeV/c^2
$n_0(\Omega_b^- \to pK^-\pi^-)$	1.77 ± 0.68
$n_0(\Xi_b^- \to pK^-\pi^-)$	1.41 ± 0.62
$n_1/n_0(\Omega_b^- \to pK^-\pi^-)$	24.99 ± 23.00
$n_1/n_0(\Xi_b^- \to pK^-\pi^-)$	24.98 ± 16.30
$\sigma(\Omega_b^- \to pK^-\pi^-)$	18.48 ± 2.16 MeV/c^2
$\sigma(\Xi_b^- \to pK^-\pi^-)$	18.37 ± 1.76 MeV/c^2

These partially reconstructed backgrounds (PRB) are modelled by an ARGUS [128] threshold function convoluted with a Gaussian to account for detector resolution, where the parameters are obtained from fits to MC samples. The generalised ARGUS function is defined via three parameters, $M_{\rm thr}$, c, and p,

$$P(m; M_{\rm thr}, c, p) = \frac{2^{-p}c^{2(p+1)}}{\Gamma(p+1) - \Gamma(p+1, c^2/2)} \times \frac{m}{M_{\rm thr}^2}\left(1 - \frac{m^2}{M_{\rm thr}^2}\right)^p \exp\left\{-\frac{1}{2}c^2\left(1 - \frac{m^2}{M_{\rm thr}^2}\right)\right\}, \tag{4.11}$$

when $m < M_{\rm thr}$, and zero elsewhere. The expressions $\Gamma(x)$ and $\Gamma(s, x)$ represent the gamma and upper incomplete gamma functions, $M_{\rm thr}$ is a threshold mass value, c governs the curvature of the function, and p controls the falling of the slope. In the fit to MC samples, the threshold mass is fixed to the value obtained from the difference in Ξ_b^- (Ω_b^-) mass (taken from Table 4.1) and the mass of the unreconstructed particle, which is π^0 in this case.

The shapes of the PRB for $\Xi_b^- \to N(p\pi^0)K^-K^-$ modes can be seen in Figure 4.23 and

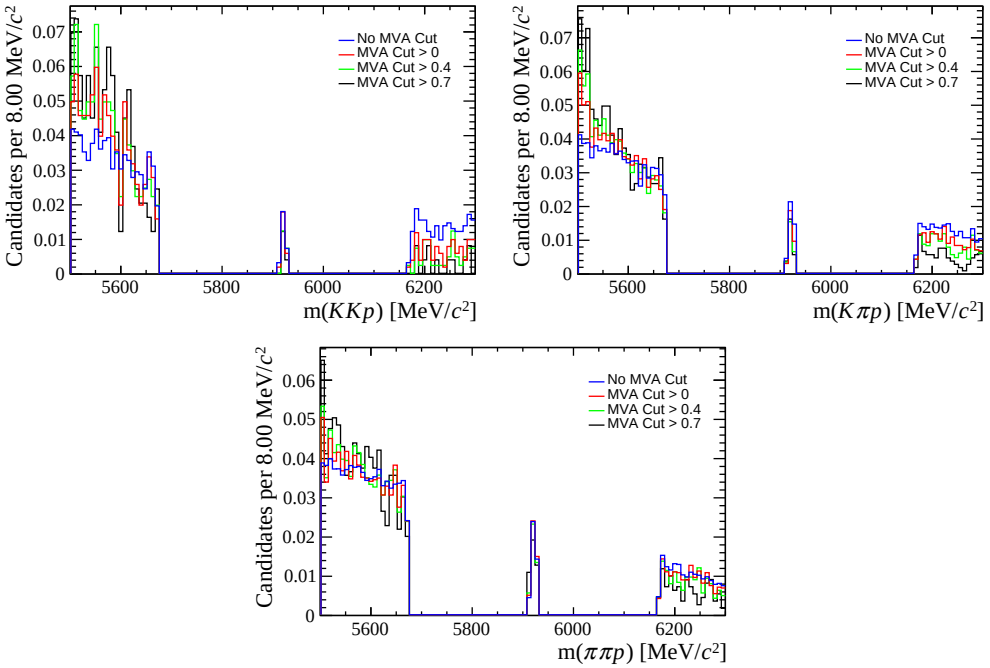

Figure 4.22: Normalised data background distributions in the pKK (top left), $pK\pi$ (top right) and $p\pi\pi$ (bottom) spectra with different MVA requirement.

the parameters obtained from these fits can be seen in Table 4.15. Fits are conducted with PRB components contributing to other $p\ h^-h'^-$ final states, but are not shown here.

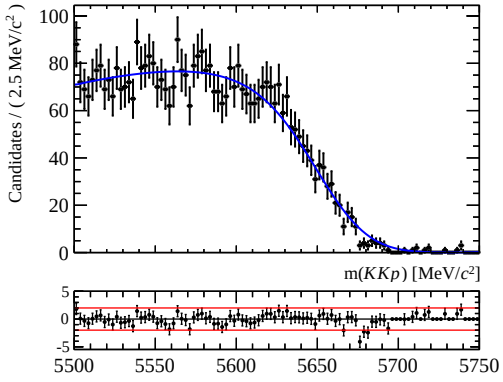

Figure 4.23: Partially reconstructed background shape for $\Xi_b^- \to N(p\pi^0)K^-K^-$ modelled as an ARGUS function convoluted with a Gaussian.

With regard to the PRB contribution towards the normalisation mode I have investigated

Table 4.15: Values of the parameters for the model describing the $\Xi_b^- \to N(p\pi^0)K^-K^-$ PRB contributions.

Parameters	Values (2011 + 2012)
$p(\Xi_b^- \to N(p\pi^0)K^-K^-)$	2.21 ± 0.45
$\sigma(\Xi_b^- \to N(p\pi^0)K^-K^-)$	20.62 ± 2.49 MeV/c^2
$c(\Xi_b^- \to N(p\pi^0)K^-K^-)$	-49.56 ± 14.3
$M_{\rm thr}(\Xi_b^- \to N(p\pi^0)K^-K^-)$	5659.40 MeV/c^2 (fixed)

not only the decay modes $B_s^0 \to \overline{K}^{*0}(K^-\pi^+)\phi(K^+K^-)$ & $B^+ \to K^{*+}(K^+\pi^0)\phi(K^-K^+)$, which will be included in the fit model, but also possible contributions from $B^+ \to K^{*+}(K^+\pi^0)K^-K^+$ which will be used for investigating systematic uncertainties and from $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $B_s^0 \to K^{*0}(K^+\pi^-)\overline{K}^{*0}(\pi^+K^-)$, $\Xi_b^0 \to pK^-K^-\pi^+$ & $B_s^0 \to f_0(\pi^-\pi^+)\phi(K^-K^+)$ modes, whose contributions are negligible and are accommodated in the combinatorial background shape in the normalisation fit model. The $B^+ \to K^{*+}(K^+\pi^0)\phi(K^-K^+)$ and $B_s^0 \to \overline{K}^{*0}(K^-\pi^+)\phi(K^+K^-)$ PRB contributions are modelled by ARGUS threshold functions convoluted with a Gaussian to account for detector resolution.

4.4.1.4 Peaking background components

The presence of like-charged mesons and a proton in the final state minimises contributions from any known fully reconstructed background that contribute to the signal region. There could in principle be a contribution from the previously unobserved $\Xi_b^- \to \Xi_c^0(ph)h'$ decay, but this is vetoed as described in Sec. 4.2.3.3. There could also be a contribution from $\Xi_b^0 \to p\pi^-, pK^-$ and $\Xi_b^-(\Omega_b^-) \to \Lambda(p\pi^-)h^-$, however these decays are not only suppressed but are removed by the NEUROBAYES output requirement due to their different topology to the signal modes.

For the normalisation mode, I veto contributions from $B^+ \to \overline{D}^0(h^+h'^-)h^+$ decays by removing the region $m(h^-h^+) \in \{1834, 1894\}\,\mathrm{MeV}/c^2$ where the nominal D^0 mass is $1864.84\,\mathrm{MeV}/c^2$ (Note that the $\sigma(D \to hh)$ is 15 MeV) [63]. This contribution can be seen in Figure 4.24 after all the selection requirements have been imposed on the data sample. Since for both signal and normalisation mode, this background component is vetoed, I do not include it in the fit model.

4.4.1.5 Combinatorial background for signal modes

For both signal and normalisation modes, the combinatorial background is described by an exponential function with floating slope parameter $c_{\text{comb. bkgd}}$. This parameter is shared between

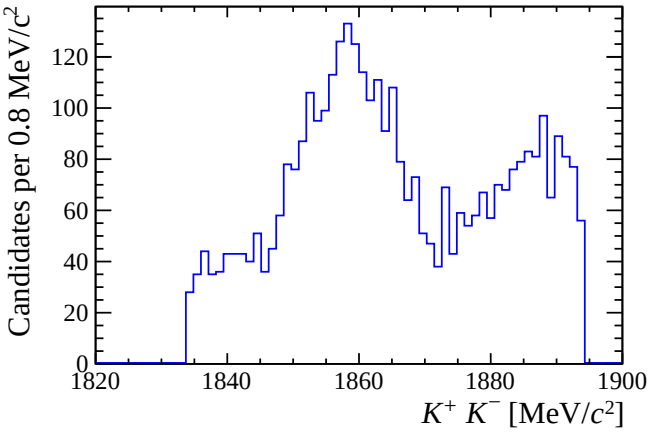

Figure 4.24: The D^0 meson contribution seen in the $m(K^-K^+)$ invariant mass spectrum on data, after all selection requirements (except the veto) have been imposed to the KKK data sample. I show here the vetoed region *i.e.* $m(h^-h^+) \in \{1834, 1894\}\,\mathrm{MeV}/c^2$.

each $p\,h^-h'^-$ spectrum in the fit to data to stabilise the fit and is floated separately in the fit to the normalisation channel.

4.4.1.6 Summary of the fit model

The fit models for the signal modes ($\Xi_b^-(\Omega_b^-) \to phh'^-$) and for the normalisation mode ($B^+ \to K^+K^-K^+$) are summarised in Table 4.16.

4.4.2 Likelihood construction and constraints on the fit model

To conduct the fit to the invariant mass (m) spectrum, I construct a total PDF. The form of this PDF is given below

$$\mathcal{P}_{\rm tot}(m, \vec{\theta}) = f_{\rm sig}\mathcal{P}_{\rm sig} + \sum_{bkg} f_{\rm bkg}\mathcal{P}_{\rm bkg}. \tag{4.12}$$

Here the subscripts sig and bkg denote the signal and individual background categories discussed in Sec. 4.4.1, respectively. The vector $\vec{\theta}$ denotes the parameters of the model that are extracted from the fit to the data sample. The symbol f denotes the fraction of the i^{th} fit component, which is related to the number of candidates that belong to the i^{th} fit component (N_i) and the total number of candidates in the sample (N_{tot}) by $f_i = N_i/N_{tot}$. The symbol $\mathcal{P}$ represents the PDFs of the components, which have been discussed in detail in Sec. 4.4.

The likelihood, which is a function of the parameters of the model given the data sample,

Table 4.16: Summary of the fit models for the $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal and $B^+ \to K^+K^-K^+$ normalisation modes. Here DCB refers to Double Crystal Ball PDF, SCB refers to Single Crystal Ball PDF and ArgusConvGauss refers to ARGUS PDF convoluted with Gaussian PDF.

Fit model for $\Xi_b^- \to phh'^-$ yield extraction		
Modes	Shape	Description
$\Xi_b^-(\Omega_b^-) \to phh'^-$	DCB	Signal component
$B^- \to h^-h'^-h'^+$	DCB	Cross-feed included
$\Xi_b^-(\Omega_b^-) \to phh'^-$	DCB	Cross-feed included, except for $\Xi_b^-(\Omega_b^-) \to p\pi^-\pi^- \to pKK$
$\Xi_b^-(\Omega_b^-) \to N(p\pi^0)h^-h'^-$	ArgusConvGauss	PRB included
$\Xi_b^-(\Omega_b^-) \to \Delta(p\pi^0)h^-h'^-$	ArgusConvGauss	PRB *not* included by default, but considered as a source of systematic uncertainty
$\Xi_b^-(\Omega_b^-) \to K^{*-}(K^-\pi^0)hp$	ArgusConvGauss	PRB *not* included by default, but considered as a source of systematic uncertainty
$\Xi_b^-(\Omega_b^-) \to \rho^-(\pi^-\pi^0)hp$	ArgusConvGauss	PRB *not* included by default, but considered as a source of systematic uncertainty
$\Xi_b^- \to \Xi_c^0(ph)h'$	-	Vetoed $m(ph^-) \in \{2446, 2496\}\,\mathrm{MeV}/c^2$
Combinatorial bkg	Falling exponential	Included
Fit model for $B^+ \to K^+K^-K^+$ yield extraction		
$B^+ \to K^+K^-K^+$	DCB	Signal component
$B^+ \to K^+\pi^-\pi^+$	DCB	Cross-feed included
$B^+ \to K^+\pi^-K^+$	SCB	Cross-feed included
$B^+ \to K^{*+}(K^+\pi^0)\phi(K^-K^+)$	ArgusConvGauss	PRB included
$B_s^0 \to \overline{K}^{*0}(K^-\pi^+)\phi(K^+K^-)$	ArgusConvGauss	PRB included
$B^+ \to K^{*+}(K^+\pi^0)K^-K^+$	ArgusConvGauss	PRB *not* included by default, but considered as a source of systematic uncertainty
$B^+ \to D^0(h^+h'^-)h^+$	-	Vetoed $m(h^-h^+) \in \{1834, 1894\}\,\mathrm{MeV}/c^2$
Combinatorial bkg	Falling exponential	Included

is constructed as follows:

$$\mathcal{L}(\vec{\theta}) = \prod_{i}^{N_{\rm tot}} \mathcal{P}_{\rm tot}(m_i, \vec{\theta}). \tag{4.13}$$

As the expected numbers of signal and background candidates are themselves random variables to be extracted from the fit to the data sample, I extend the likelihood by multiplying a Poisson term to Eq. 4.13 as follows

$$\mathcal{L}_E(\vec{\theta}) = \frac{\exp{(-N)} N^{N_{\rm tot}}}{N_{\rm tot}!} \times \prod_{i}^{N_{\rm tot}} \mathcal{P}_{\rm tot}(m_i, \vec{\theta}). \tag{4.14}$$

Here $N = N_{\rm sig} + \sum_{\rm bkg} N_{\rm bkg}$ is the total number of expected candidates, which can differ from $N_{\rm tot}$.

When the fit is conducted to the data sample, we aim to minimise the value of $-\ln(\mathcal{L})$ to obtain the best estimates of the parameters of the model ($\hat{\theta}$). The PDFs are implemented in the RooFit package [129] and the minimisation is carried out using the program MINUIT [130, 131], which first uses MIGRAD to minimise $-\ln(\mathcal{L})$ using the gradient descent method [132] to obtain the best estimates of the parameters of the model, and then uses HESSE to improve upon the numerical estimate of the covariance matrix obtained using MIGRAD.[7]

If some l parameters of the model are known ($\vec{\theta_k}$) with a given uncertainty ($l \times l$ covariance matrix V_k), it is common to incorporate such constraints into maximum likelihood fits by multiplying a multivariate Gaussian term to Eq. 4.14 as follows

$$\mathcal{L}_C(\vec{\theta}) = \frac{\exp(-(\vec{\theta_k} - \vec{\theta}) V_k^{-1} (\vec{\theta_k} - \vec{\theta})^T)}{(2\pi)^{l/2}\sqrt{|V_k|}} \times \mathcal{L}_E(\vec{\theta}). \tag{4.15}$$

This procedure from here on is referred to as applying a Gaussian constraint.

For signal modes I conduct an unbinned extended maximum likelihood fit to data, where the three spectra (pKK, $pK\pi$ and $p\pi\pi$) are fit simultaneously.[8]. The advantage of this is that in addition to being able to constrain parameters to those found in the MC simulations, one can also constrain the relative yields of various components, where appropriate as detailed in the next sub-section. For the normalisation mode, I conduct a separate fit to the KKK mass spectrum. Below I highlight the constraints imposed on the fits to both $p\ h^-h'^-$and KKK invariant mass spectra.

[7] The covariance matrix is determined numerically by calculating the matrix of second derivatives of $\ln(\mathcal{L})$ using the finite difference method [133] at the maximum likelihood estimates ($V_{ij}^{-1} = -\frac{\partial^2 \ln(\mathcal{L})}{\partial\theta_i \partial\theta_j}|_{\vec{\theta}=\hat{\theta}}$), and then inverting it. It can also be obtained from a $\ln(\mathcal{L})$ scan of the parameter space as is done using MINOS [130, 134].

[8] In a simultaneous fit, the product of likelihoods of each spectrum *i.e.* $L_C = L_C^{pKK} \times L_C^{pK\pi} \times L_C^{p\pi\pi}$ is maximised.

4.4.2.1 Constraints on the fit to $p\,h^-h'^-$ spectra

When conducting fits to the $m(ph^-h'^-)$ spectra using the data sample, I fix the shape parameters of the fit model and also Gaussian constrain certain parameters of the model that are known *a priori*. These various constraints on the fit model for signal modes are listed below:

- For all signal and cross-feed shapes, each of which is modelled with a DCB PDF with common mean (μ) and width (σ), the parameter f which indicates the relative sizes of the two CB functions (*i.e.* $f \times \mathrm{CB}_0 + (1-f) \times \mathrm{CB}_1$) and the tail parameters (n_0, n_1/n_0, α_0 and α_1/α_0) are fixed to the values obtained from fits to these MC samples. See Eq. 4.4 for the definition of various parameters of DCB.

- The widths of all DCB PDFs (signal & cross-feed) and the widths of all the PRB shapes (described by ARGUS threshold functions convolved with a Gaussian) are fixed to the values obtained from fits to the MC samples, with a common scaling fixed to the value obtained from the fit to the normalisation mode.

- The means of the DCB PDFs that parametrise the signal components are fixed to the known Ξ_b^- (Ω_b^-) mass values (see Table 4.1).

- The means of the DCB PDFs that parametrise cross-feed components are fixed to the values obtained from fits to the MC samples.

- The slope parameters (c) of the exponential PDFs that parametrise cross-feed shapes (mainly from $B^- \to h^-h'^-h'^+$) are also fixed to values taken from fits to MC samples.

- The slope parameters of the exponential functions describing combinatorial background are constrained to be the same in the pKK, $pK\pi$ and $p\pi\pi$ mass spectra.

- For PRB modelled as an ARGUS function convolved with a Gaussian PDF, the slope (c) and power (p) parameters are fixed to the values obtained from fits to the MC samples.

- The threshold parameters of ARGUS functions are fixed to the appropriate kinematic value obtained from the difference in the parent particle (Ξ_b^- (Ω_b^-)) mass and the mass of the unreconstructed particle (π^0).

- The cross-feed yields of the baryonic modes ($\Xi_b^-(\Omega_b^-) \to phh'^-$), *i.e.* from the other signal modes, are set to the corresponding yield in their signal spectrum, multiplied by the misidentification rate obtained by the calibrated PID information.

- The cross-feed yields from $B^- \to h^- h'^- h'^+$ modes are Gaussian constrained in the fit. To estimate the cross-feed yields from these modes, I first conduct fits to each of the $m(h^- h'^- h'^+)$ spectra to obtain the signal-to-background ratio (f_{sb}). I then deduce the cross-feed yield in each of the $p\,h^- h'^-$spectra ($N_{\rm cross-feed}$) using the relation

$$N_{\rm cross-feed} = \frac{\epsilon^{\rm mis-id}_{h^- h'^- h'^+ \to p h^- h'^-}}{\epsilon^{\rm corr-id}_{h^- h'^- h'^+ \to p h^- h'^-}} \times \frac{N_{h^- h'^- h'^+} \times f_{sb}}{(1 + f_{sb})}\,, \tag{4.16}$$

  where $N_{h^- h'^- h'^+}$ is the total number of data candidates lying in both $m(h^- h'^- h'^+)$ and $m(p h^- h'^-)$ fit ranges and $\frac{\epsilon^{\rm mis-id}_{h^- h'^- h'^+ \to p h^- h'^-}}{\epsilon^{\rm corr-id}_{h^- h'^- h'^+ \to p h^- h'^-}}$ is the efficiency ratio obtained `PIDCalib`.

4.4.2.2 Constraints on the fit to the KKK spectrum

As in the case of the signal, when conducting fits to the $m(KKK)$ spectra using the data sample, certain parameters of the model are either fixed or Gaussian constrained. The various constraints on the normalisation mode fit model are:

- For signal and cross-feed modelled as either a single or double CB, the tail parameters are fixed to values obtained from the fit to MC.
- The means and widths of all single and double CB PDFs are fixed to the corresponding values obtained from MC with a common floating shift and scale factor respectively to account for data/MC differences.
- For PRB modelled with an ARGUS function convolved with a Gaussian, the slope (c) and power (p) parameters are fixed to the values obtained from the fit to the MC sample.
- The ARGUS function convolved with a Gaussian PDF modelling PRB components and the DCB PDF which parametrise the signal component share the same width parameter.
- The threshold parameter of the ARGUS function, is fixed to the mean value of the DCB PDF that parametrise the signal component minus the mass of the unreconstructed (π^0) particle.
- The cross-feeds from $B^+ \to K^+\pi^-K^+$ and $B^+ \to K^+\pi^-\pi^+$ are constrained within uncertainties to the expected yield, which is obtained in a similar way as the signal modes.

4.4.3 Stability of the fit model

To test the stability of the fit model, I first conduct a blind fit to the data sample (where the signal yields, nuisance parameters and the signal regions are not revealed to the analyst) using the fit

model described in the previous subsections. The values of the nuisance parameters obtained from this fit are used to generate ensembles of pseudo-experiments (7500 pseudo-experiments per ensemble) with various different signal yields (*i.e.* signal yield varied between 0–100 candidates). These pseudo-experiments are then fitted using the same model used to fit the data sample. The generated and fitted yields are then compared to test the stability of the fit. The results of this study can be seen in Figure 4.25 and in Figure 4.26. Good linearity is seen for all yields. Small biases are evident in the case that the pKK signal yield is very small, which is as expected since the background is also very small in this final state. These biases, which are anyway only at the level of $\mathcal{O}(10\,\%)$ of the statistical uncertainty, disappear in the case of larger yields.

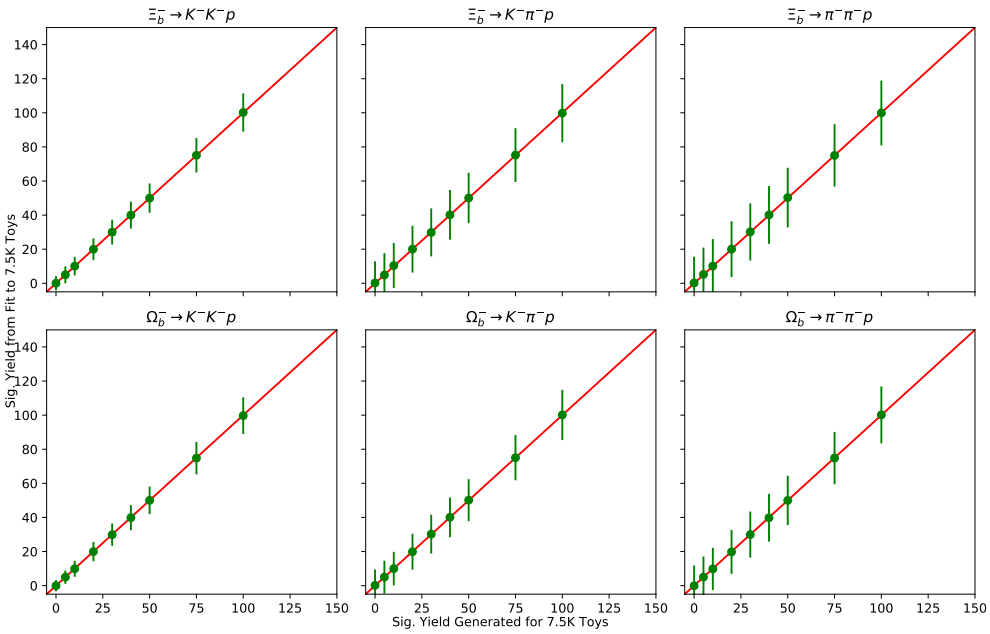

Figure 4.25: Mean fitted signal yield as a function of the generated signal yield, as obtained from ensembles of pseudo-experiments for each of the signal modes. A truly unbiased model shows a linear relationship between these to variables and this is indicated by the red line. The green dots indicate the mean of the distribution and the error bars indicate the width of the distribution.

4.5 Extraction of yields from the fit

In this section, I present the the results of the fit to the invariant mass spectra to extract the yields (see Subsec. 4.5.1) and discuss the strategy employed to calculate the efficiency-corrected yields that go into the branching fraction measurements (see Subsec. 4.5.2).

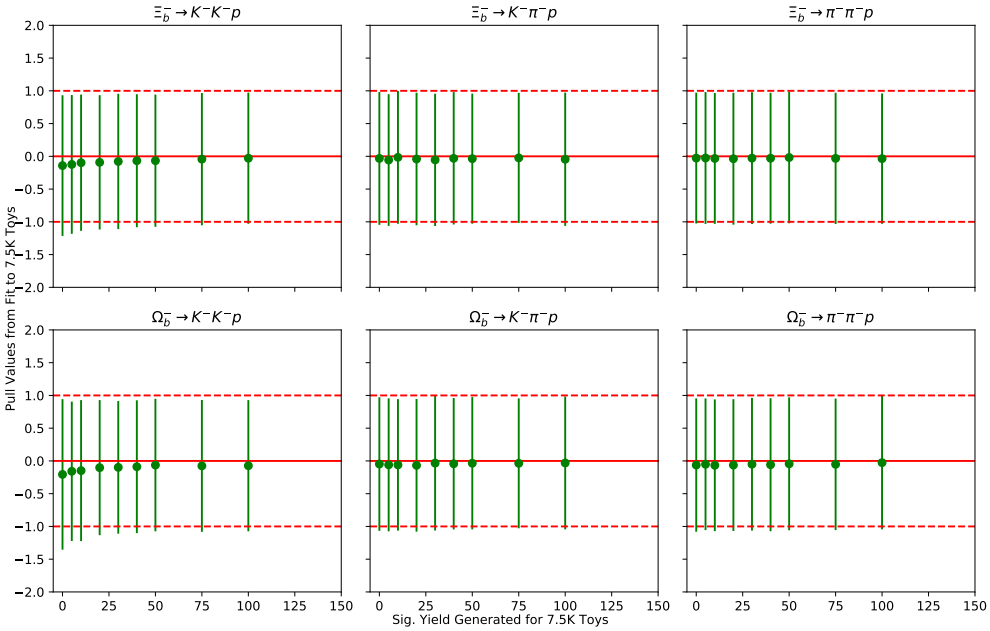

Figure 4.26: Pulls for each signal yield as a function of the generated signal yield, as obtained from ensembles of pseudo-experiments for each of the signal modes. A truly unbiased model would give a Gaussian distribution with mean of zero and width of one, as indicated by the solid and dashed red lines. The green dots indicate the mean of the pull distribution and the error bars indicate the width of the pull distribution.

4.5.1 Fit to the $p\,h^-h'^-$ and KKK spectra

I conduct, as mentioned in the previous section, a simultaneous unbinned extended maximum likelihood fit to $m(pKK)$, $m(pK\pi)$ and $m(p\pi\pi)$ invariant masses to extract the yields of all six signal $\Xi_b^-(\Omega_b^-) \to phh'^-$ modes. I also conduct a separate fit to the $m(KKK)$ mass spectrum to extract the yield of the normalisation $B^+ \to K^+K^-K^+$ mode. The fit projections to the Run I (2011–2012) LHCb data sample using the fit model described in the previous section are shown in Figure 4.27. The results of the fit to the $p\,h^-h'^-$ spectra are shown in Table 4.17 and for the KKK spectra in Table 4.18. Also shown is the correlation matrix of all the floated parameters of the simultaneous fit to $p\,h^-h'^-$spectra in Figure 4.28.

After unblinding the signal region, I see a strong signal in the $\Xi_b^- \to pK^-K^-$ mode at an approximate statistical significance of 11 Gaussian standard deviations. The significance is determined from $\sqrt{-2\Delta \ln \mathcal{L}}$, where $\Delta \ln \mathcal{L}$ is the change in log-likelihood between the nominal fit and a fit in which the relevant signal yield is fixed to zero. Evidence for the $\Xi_b^- \to pK^-\pi^-$ decay is found at a level of approximately 4σ. The $\Xi_b^- \to p\pi^-\pi^-$ decays and all Ω_b^- modes have a statistical significance below 3σ.

4.5.2 Efficiency corrected yields

The dynamics of a multibody decay govern the distribution of candidates in the phase space. In case the efficiency varies significantly over the phase space (as has been shown in Sec. 4.3), it is necessary to account for this correctly in the evaluation of Eq. (4.2). Since the dynamics of the signal modes are not known *a priori*, they must either be determined from data or some assumption must be made with a corresponding systematic uncertainty assigned. For modes observed with a significant signal I use the following procedure to calculate the efficiency-corrected yields:

- For each candidate in each final state, determine the signal *sWeight*, w_i, from the result of the fit to the signal modes.
- Use the efficiency maps given in Sec. 4.3 to obtain the appropriate efficiency value ϵ_i depending on the position of candidate i in the square Dalitz plane.
- The efficiency-corrected yield is then given by

$$N^{\rm corr}_{\Xi_b^-(\Omega_b^-)} = \frac{N_{\Xi_b^-(\Omega_b^-)}}{\epsilon_{\Xi_b^-(\Omega_b^-)}} = \sum_i^N \frac{w_i}{\epsilon_i} \tag{4.17}$$

  where N is the number of candidates for each final state. The statistical uncertainty on the

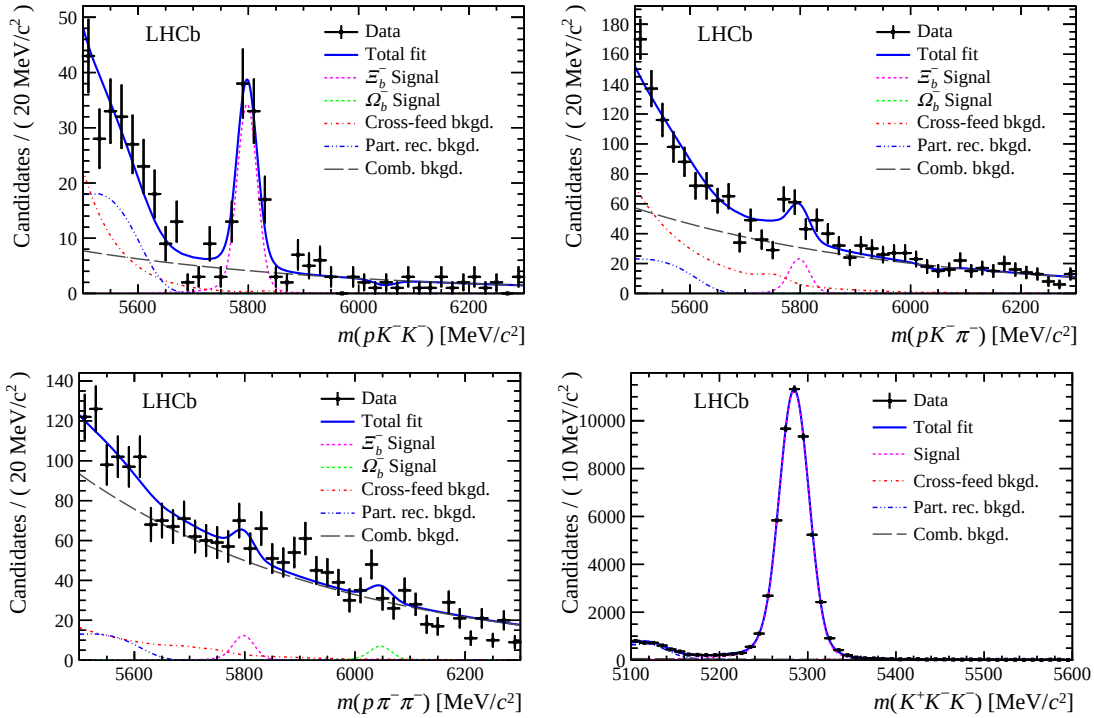

Figure 4.27: Mass distributions for b-hadron candidates in the (top left) pK^-K^-, (top right) $pK^-\pi^-$, (bottom left) $p\pi^-\pi^-$ and (bottom right) $K^+K^-K^-$ final states. Results of the fits are shown with dark blue solid lines. Signals for Ξ_b^- and B^- (Ω_b^-) decays are shown with pink (light green) dashed lines, combinatorial backgrounds are shown with grey long-dashed lines, cross-feed backgrounds are shown with red dot-dashed lines, and partially reconstructed backgrounds are shown with dark blue double-dot-dashed lines [135].

efficiency-corrected yield [136] is given by

$$\sigma^{\rm corr}(N^{\rm corr}) = \sqrt{\sigma(N^{\rm corr})^2 + \left(\frac{N^{\rm corr}}{N^{\rm shape}}\sigma(N^{\rm shape})\right)^2}, \tag{4.18}$$

where $N^{\rm shape} \pm \sigma(N^{\rm shape})$ is the yield and its uncertainty returned by a fit where the shape parameters are fixed and only yields are floated and $\sigma(N^{\rm corr}) = \sqrt{\sum_i^N (\frac{w_i}{\epsilon_i})^2}$. The procedure employed here is similar to the strategy followed in Ref. [136], to correct for the fact that the efficiency-corrected yield is determined using *sWeights*, which are necessarily obtained from a fit in which all shape parameters are fixed.

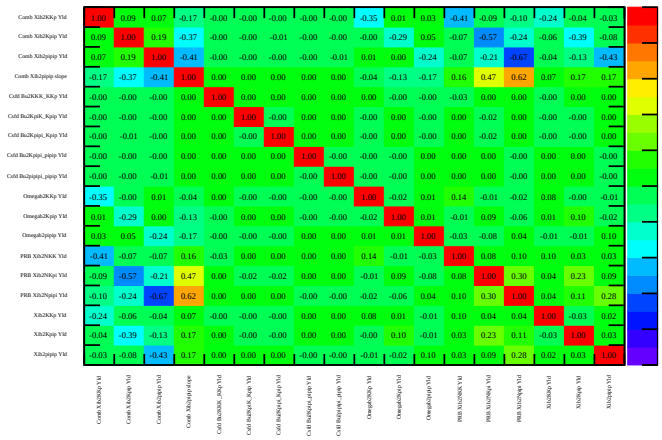

Figure 4.28: The correlation matrix of all the floated parameters obtained from the simultaneous fit to $p\ h^-h'^-$spectra.

Note that the average efficiency can then be written as

$$\bar{\epsilon}_{\Xi_b^-(\Omega_b^-)} = \frac{N_{\Xi_b^-(\Omega_b^-)}}{N^{\rm corr}_{\Xi_b^-(\Omega_b^-)}} \tag{4.19}$$

where $N_{\Xi_b^-(\Omega_b^-)}$ is the signal yield obtained from the fit. It is not, however, necessary to evaluate the average efficiency since the results for $N^{\rm corr}_{\Xi_b^-(\Omega_b^-)}$ can be used directly in Eq. (4.2).

In the absence of a significant signal ($< 3\sigma$), it is not possible to make a meaningful determination of the square Dalitz Plot (SDP) distribution of signal candidates. Hence the above method for the efficiency correction cannot be used. The following procedure is used:

- The binned distributions of the efficiencies as a function of the SDP variables are used to construct a 1D histogram of the variation in efficiency for each mode.
- The mean value from this histogram is taken as the central value for the average efficiency and is used in the calculation of the relative branching fraction. Note here that the MC samples have been generated uniform in the SDP variables and therefore this procedure can be used.
- The standard-deviation of the binned efficiency values is used to determine the associated systematic uncertainty. For non-significant modes, the statistical uncertainty on the yield is

Table 4.17: Parameters of the signal fit model obtained from the fit to $p\,h^-h'^-$ invariant mass spectra. Here N denotes the yield, c denotes the slope of the exponential parameter. The † symbol indicates Gaussian constrained parameters.

Parameter	Value
Parameter of interest	
$N_{\rm sig}(\Xi_b^- \to pK^-K^-)$	82.9 ± 10.4
$N_{\rm sig}(\Xi_b^- \to pK^-\pi^-)$	59.6 ± 16.0
$N_{\rm sig}(\Xi_b^- \to p\pi^-\pi^-)$	33.2 ± 17.9
$N_{\rm sig}(\Omega_b^- \to pK^-K^-)$	-2.8 ± 2.5
$N_{\rm sig}(\Omega_b^- \to pK^-\pi^-)$	-7.6 ± 9.2
$N_{\rm sig}(\Omega_b^- \to p\pi^-\pi^-)$	20.1 ± 13.8
Nuisance parameters	
$N_{\rm cross-feed}(\Xi_b^-(\Omega_b^-) \to pK^-\pi^- \to pKK)$	$(0.0199 \pm 0.0001) \times N_{\rm sig}(\Xi_b^-(\Omega_b^-) \to pK^-\pi^-)$
$N_{\rm cross-feed}(\Xi_b^-(\Omega_b^-) \to pK^-K^- \to pK\pi)$	$(0.2532 \pm 0.0005) \times N_{\rm sig}(\Xi_b^-(\Omega_b^-) \to pK^-K^-)$
$N_{\rm cross-feed}(\Xi_b^-(\Omega_b^-) \to p\pi^-\pi^- \to pK\pi)$	$(0.0156 \pm 0.0001) \times N_{\rm sig}(\Xi_b^-(\Omega_b^-) \to p\pi^-\pi^-)$
$N_{\rm cross-feed}(\Xi_b^-(\Omega_b^-) \to pK^-K^- \to p\pi\pi)$	$(0.0686 \pm 0.0002) \times N_{\rm sig}(\Xi_b^-(\Omega_b^-) \to pK^-K^-)$
$N_{\rm cross-feed}(\Xi_b^-(\Omega_b^-) \to pK^-\pi^- \to p\pi\pi)$	$(0.1640 \pm 0.0004) \times N_{\rm sig}(\Xi_b^-(\Omega_b^-) \to p\pi^-\pi^-)$
$N_{\rm cross-feed}(B^+ \to K^+K^-K^+ \to pKK)$†	73.36 ± 0.57
$N_{\rm cross-feed}(B^+ \to K^+\pi^-K^+ \to pK\pi)$†	36.15 ± 0.97
$N_{\rm cross-feed}(B^+ \to K^+\pi^-\pi^+ \to pK\pi)$†	372.13 ± 1.14
$N_{\rm cross-feed}(B^+ \to K^+\pi^-\pi^+ \to p\pi\pi)$†	62.92 ± 0.22
$N_{\rm cross-feed}(B^+ \to \pi^+\pi^-\pi^+ \to p\pi\pi)$†	66.46 ± 0.62
$N_{\rm PRB}(\Xi_b^- \to N(p\pi^0)K^-K^- \to pKK)$	87.18 ± 16.76
$N_{\rm PRB}(\Xi_b^- \to N(p\pi^0)K^-\pi^- \to pK\pi)$	77.90 ± 39.49
$N_{\rm PRB}(\Xi_b^- \to N(p\pi^0)\pi^-\pi^- \to p\pi\pi)$	11.87 ± 46.92
$c_{\text{comb. bkgd}}$	$(-24.77 \pm 1.25) \times 10^{-4}\ c^2/\text{MeV}$

taken directly from the result of the fit to data; these are simply scaled to obtain uncertainties on the efficiency-corrected yields.

The signal yields are given in Table 4.19 with their respective significance (statistical only). In Table 4.19 I also quote the total average efficiency (see Eq. 4.19) and the efficiency-corrected yield. For significant modes ($> 3\sigma$) this is calculated using Eq. 4.17. For non-significant modes the total efficiency is simply taken from Table 4.12 (the lifetime-corrected values are used) and the corrected yield is just $\frac{N_{\Xi_b^-(\Omega_b^-)}}{\epsilon^{\rm tot}_{\Xi_b^-(\Omega_b^-)}}$. For completeness I also show here the numbers for the normalisation mode.

Table 4.18: Results of the fit to the $B^+ \to K^+K^-K^+$ normalisation mode. Here N denotes the yield, c denotes the slope of the exponential parameter, $\mu_{\rm shift}$ & $\sigma_{\rm scale}$ represent the common shift and scale factors for the means and widths of the PDFs. The † symbol indicates Gaussian constrained parameters.

Parameter	Value
Parameter of interest	
$N_{\rm sig}(B^+ \to K^+K^-K^+)$	50493 ± 250
Nuisance parameters	
$N_{\rm cross-feed}(B^+ \to K^+\pi^-K^+)$†	295.63 ± 7.75
$N_{\rm cross-feed}(B^+ \to K^+\pi^-\pi^+)$†	56.37 ± 0.15
$N_{\rm PRB}(B_s^0 \to \overline{K}^{*0}(K^-\pi^+)\phi(K^+K^-))$	428.65 ± 196
$N_{\rm PRB}(B^+ \to K^{*+}(K^+\pi^0)\phi(K^-K^+))$	2879.90 ± 96
$N_{\rm comb.\ bkgd}$	1804.50 ± 355
$c_{\rm comb.\ bkgd}$	$(-35.00 \pm 9.53) \times 10^{-4}\ c^2/{\rm MeV}$
$\mu_{\rm shift}$	$3.85 \pm 0.08\ {\rm MeV}/c^2$
$\sigma_{\rm scale}$	$(1167.40 \pm 5.49) \times 10^{-3}$

Table 4.19: Results of the fit to the $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal modes after unblinding. I also show the significance determined from the likelihood functions using $\Delta \ln \mathcal{L}$ (statistical uncertainty only). The average efficiency, $\epsilon^{\rm tot}$, is also quoted. For modes without a significant signal, these values are taken from Table 4.12; for modes with significant signal candidate-by-candidate efficiency corrections is used. The efficiency-corrected yield, $N^{\rm corr}_{\Xi_b^-(\Omega_b^-)}$ (see Eq. 4.17), is given for all the signal modes. For completeness, I also collect here relevant numbers for the normalisation mode.

Parameter	Value	Significance (statistical only)	$\epsilon^{\rm tot}$ (%)	$N^{\rm corr} \pm \delta N^{\rm corr}$
$N_{\rm sig}(\Xi_b^- \to pK^-K^-)$	82.9 ± 10.4	11.4	0.3975	20859 ± 2753
$N_{\rm sig}(\Xi_b^- \to pK^-\pi^-)$	59.6 ± 16.0	4.0	0.2925	20378 ± 5035
$N_{\rm sig}(\Xi_b^- \to p\pi^-\pi^-)$	33.2 ± 17.9	1.9	0.5733	5789 ± 3137
$N_{\rm sig}(\Omega_b^- \to pK^-K^-)$	-2.8 ± 2.5	—	0.3753	-736 ± 668
$N_{\rm sig}(\Omega_b^- \to pK^-\pi^-)$	-7.6 ± 9.2	—	0.4183	-1828 ± 2204
$N_{\rm sig}(\Omega_b^- \to p\pi^-\pi^-)$	20.1 ± 13.8	1.5	0.5359	3752 ± 2586
$N_{\rm sig}(B^+ \to K^+K^-K^+)$	50500 ± 250	—	0.6427	7855571 ± 40574

4.6 Systematic uncertainties

Sources of systematic uncertainty can potentially bias the determination of the relative branching fraction using Eq. (4.2). The choice of normalisation channel ($B^+ \to K^+K^-K^+$), with a similar

topology as my signal modes, reduces the size of the possible systematic uncertainty. Nonetheless, some important factors remain and in this section, I investigate various such systematic effects related to the selection efficiency and its variation over the phase-space, the PID calibration, the fit model and fit biases.

In Table 4.20 I give a summary of the systematic uncertainties on the branching fractions for the various $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal modes with respect to the $B^+ \to K^+K^-K^+$ normalisation mode. In Table 4.21 I give a summary of the systematic uncertainties on branching fractions for $\Xi_b^- \to pK^-\pi^-$ and $\Xi_b^- \to p\pi^-\pi^-$ modes with respect to the $\Xi_b^- \to pK^-K^-$ mode. I discuss the origin of these individual sources of systematic uncertainties in the following sub-sections.

Table 4.20: Absolute systematic uncertainties (10^{-5}) on the relative branching fractions for the various $\Xi_b^-(\Omega_b^-) \to phh'^-$ signal modes with respect to the $B^+ \to K^+K^-K^+$ normalisation mode. The total is the sum in quadrature of all other entries.

Mode	Ξ_b^- (Ω_b^-) decay time	Ξ_b^- (Ω_b^-) $p_{\rm T}$	Trigger	Vetoes	MC Stat.	PhSp	PID	Fit	Total
$\Xi_b^- \to pK^-K^-$	5.5	41.4	15.5	0.6	1.6	7.2	11.5	6.1	47.0
$\Xi_b^- \to pK^-\pi^-$	5.4	40.4	15.2	4.8	1.9	12.3	8.9	15.1	48.8
$\Xi_b^- \to p\pi^-\pi^-$	1.5	11.5	4.3	0.7	0.5	28.7	4.1	17.4	36.0
$\Omega_b^- \to pK^-K^-$	0.8	1.5	0.5	0.1	0.1	3.6	0.3	4.7	6.2
$\Omega_b^- \to pK^-\pi^-$	2.0	3.6	1.4	2.2	0.2	11.8	1.1	19.3	23.1
$\Omega_b^- \to p\pi^-\pi^-$	4.1	7.4	2.8	0.3	< 0.1	23.6	2.8	12.2	28.2

Table 4.21: Absolute systematic uncertainties (10^{-3}) on the relative branching fractions for the various $\Xi_b^- \to pK^-\pi^-$ and $\Xi_b^- \to p\pi^-\pi^-$ signal modes with respect to the $\Xi_b^- \to pK^-K^-$ mode. The total is the sum in quadrature of all other entries.

Mode	Ξ_b^- (Ω_b^-) decay time	Ξ_b^- (Ω_b^-) $p_{\rm T}$	Trigger	Vetoes	MC Stat.	PhSp	PID	Fit	Total
$\Xi_b^- \to pK^-\pi^-$	0.2	0.0	59	20	0.15	20	10	54	85
$\Xi_b^- \to p\pi^-\pi^-$	0.0	0.0	16	2	0.01	111	3	69	132

4.6.1 Selection efficiency

Most selection requirements are applied to the $B^+ \to K^+K^-K^+$ normalisation mode in the same way as for the signal modes. Therefore, the ratio of signal efficiencies in MC can be taken as a reliable estimate of the ratio of signal efficiencies in data, since for most selection criteria

any data/MC differences should cancel to first order. Systematic uncertainties arise where the MC modelling of the efficiency may not be reliable, and also where different requirements are made for the signal and normalisation modes.

4.6.1.1 MC and Data discrepancies

Ξ_b^- (Ω_b^-) lifetime: The Ξ_b^- (Ω_b^-) lifetime values used to generate the MC samples for signal modes are $\tau(\Xi_b^-(\Omega_b^-)) = 1.567\ (1.135)\,\mathrm{ps}$, which are different from the most recent measurements quoted in Table 4.1. Due to the dependence of the selection efficiency on the Ξ_b^- (Ω_b^-) decay time, I correct the average efficiency, and assign a corresponding systematic uncertainty related to the precision on the measured lifetime. The correction factors to the average efficiency for Ξ_b^- and Ω_b^- modes are 1.03 and 1.42, respectively. The uncertainty on the lifetime-corrected average efficiency arising due to this correction procedure is at the level of 2% for Ξ_b^- modes and 8% for Ω_b^- modes, owing to the fact that the Ω_b^- lifetime is known less precisely than that of the Ξ_b^- baryon.

Ξ_b^- (Ω_b^-) production kinematics: Similarly as discussed for the Ξ_b^- (Ω_b^-) lifetime, one might expect differences in p_{T} dependence in the fragmentation fractions for MC and data samples. To evaluate the systematic uncertainty from this effect, I calculate the difference in the average efficiency when the MC samples are reweighted to match the p_{T} distribution to that of $\Xi_b^- \to pK^-K^-$ obtained from sWeighted data. Since this is the only channel with a significant yield, I assume same systematic uncertainty for all other signal modes. The systematic uncertainties arising due to MC and data discrepancies in Ξ_b^- (Ω_b^-) production kinematics on the average efficiency from this procedure for all modes is 15.6 %. This is the dominant systematic uncertainty on the most precisely measured modes ($\Xi_b^- \to pK^-K^-$ and $\Xi_b^- \to pK^-\pi^-$).

In Sec. 4.2.3.2 I noted that the distribution of `Bu_PTASYM_1_5` variable is slightly different for $\Xi_b^- \to pK^-K^-$ and $B^+ \to K^+K^-K^+$ MC samples. Since I have already assigned a systematic uncertainty due to data/MC differences in the p_{T} distribution, adding a further source would artificially inflate the uncertainty due to double counting. Therefore, I do not consider this further.

4.6.1.2 Trigger

A possible bias can arise from a mismatch of the MC and data `L0Hadron` trigger efficiency, caused by mis-calibration of the calorimetric transverse energy measurements. A data-driven procedure is employed where the `L0Hadron` efficiency for pions, kaons, and protons, for both magnet configurations, and for 2011 and 2012 are obtained as a function of p_{T} of the particle.

This information is then used to create corrected total efficiency maps as a function of SDP variables. The systematic uncertainty from this effect is taken as the difference in the average efficiency evaluated with uncorrected and corrected efficiency maps. The size of the systematic uncertainty on the average efficiency for $\Xi_b^- \to pK^-K^-$ mode is 5.9%. The uncertainty on other modes have been evaluated in the similar fashion to that of $\Xi_b^- \to pK^-K^-$.

4.6.1.3 Vetoes

Vetoes are applied to remove possible contributions from $\Xi_b^- \to \Xi_c^0(ph)h'$ decays. The efficiency correction does not account for regions of phase space that are completely removed, and this therefore introduces a systematic uncertainty. To evaluate the effect, I increase the veto window by 50%; the shift in the result is assigned as the associated systematic uncertainty. A similar treatment is applied for the veto of $B^+ \to \overline{D}^0(h^+h'^-)h^+$ decays in the $B^+ \to K^+K^-K^+$ normalisation mode. The relative systematic uncertainties on the branching fraction ratio with respect to $B^+ \to K^+K^-K^+$ from this source is largest for the $\Omega_b^- \to pK^-\pi^-$ mode and is at the level of 9.5%.

4.6.1.4 Monte Carlo statistics

The efficiencies are evaluated from Monte Carlo samples with finite statistics. The effect is evaluated from the distribution of results when the average efficiencies (including the SDP dependent efficiencies) are varied within their uncertainties. The standard deviation of the distribution of the branching fraction ratio, calculated using these average efficiencies is assigned as the systematic uncertainty. The relative systematic uncertainties (%) on the branching fraction ratio with respect to $B^+ \to K^+K^-K^+$ from this source is largest for the $\Omega_b^- \to p\pi^-\pi^-$ mode and is at the level of 0.8%.

4.6.1.5 Phase-space (Phsp)

As highlighted in Sec. 4.5.2, for modes where a significant signal yield is observed, the average efficiency is a weighted harmonic mean of efficiencies evaluated on a candidate-by-candidate basis, where the weights are the *sWeights* obtained from the fit to signal modes. In this case, the uncertainty related to the procedure of obtaining the SDP dependent efficiencies is estimated by changing the SDP binning scheme to an adaptive one (approximately equal number of candidates in all bins of 2D SDP). If no significant signal is observed however, the average efficiency is assumed. The corresponding systematic uncertainty is taken from the scale of the variation of the efficiency across the SDP. The relative systematic uncertainty (%) on the branching fraction ratio

with respect to $B^+ \to K^+K^-K^+$ from this source is largest for the $\Omega_b^- \to p\pi^-\pi^-$ mode and is at the level of 42.9%.

4.6.1.6 Particle Identification (PID)

The uncertainty introduced by the finite size of the calibration and MC control samples are already taken into account in the statistical error quoted in Table 4.11 in Sec. 4.3.3. The assumptions that are made in the calculation of PID efficiencies are as follows:

- The RICH response can be completely parametrised by p, $p_{\rm T}$, and `nTracks`. This is generally accepted to be true to a good approximation.
- The efficiency is smoothly varying within the chosen bins of the p, $p_{\rm T}$, and `nTracks` variables. To evaluate the uncertainty from this, I consider different binning schemes for the `PIDCalib` calibration data and repeat the calculation of the efficiencies.
- An extra $0.3\,\%$ systematic uncertainty ($0.1\,\%$ per track) is added to account for the correlation between p, $p_{\rm T}$, `nTracks` and the variable fitted to extract the sWeights by the `PIDCalib` package [121].
- The kinematic distributions of the signal for MC and data samples are assumed to not only match amongst themselves but also match the corresponding distributions of the `PIDCalib` calibration data sample. The binning in p and $p_{\rm T}$ deals with any disagreement between the kinematics of the signal channel and calibration channels. A source of potential differences in kinematics between data and MC could arise due to mis-modelling of the Ξ_b^- (Ω_b^-) differential production cross-section (with respect to $p_{\rm T}$ and rapidity). This has not yet been measured, and is modelled in the MC using PYTHIA. The PYTHIA distribution is reasonably reliable, since it is tuned to reproduce the known production characteristics of other b hadrons [137].

The relative systematic uncertainties (%) on the branching fraction ratio with respect to $B^+ \to K^+K^-K^+$ from the assumptions made above in the PID efficiency evaluation is largest for the $\Omega_b^- \to p\pi^-\pi^-$ mode and is at the level of 5.8%.

4.6.2 Fit model

The fitting procedure introduces three sources of systematic uncertainty. The systematic uncertainties related to the fit model are presented as absolute uncertainties on the value of $N_{\rm sig}/N_{\rm norm}$. For the reader's convenience, we present in Table 4.22 the central values and statistical uncertainties of these values. We also show here the combined statistical plus systematic significance

for the modes with statistical significance $> 3\sigma$ (see Table 4.19). These three sources are briefly described below.

Table 4.22: Central values of ratio of signal yield to normalisation yield and the statistical uncertainty on them. The total (statistical plus systematic) significance are also given.

Mode	$N_{\rm sig}/N_{\rm norm}$ (10^{-6})	Stat. Error (10^{-6})	Significance (stat + syst)
$\Xi_b^- \to pK^-K^-$	1642	206	8.7σ
$\Xi_b^- \to pK^-\pi^-$	1181	318	3.4σ
$\Xi_b^- \to p\pi^-\pi^-$	657	354	—
$\Omega_b^- \to pK^-K^-$	-55	50	—
$\Omega_b^- \to pK^-\pi^-$	-151	182	—
$\Omega_b^- \to p\pi^-\pi^-$	398	273	—

4.6.2.1 Fit model choice

The choice of the fit model introduces an uncertainty that is estimated by repeating the fit with alternative fit models for each component, by considering the inclusion of components that are excluded in the baseline fit, and by relaxing some of the constraints. The nominal fit consists of DCB PDFs for the signal and baryonic cross-feed distributions, an ARGUS function convolved with a Gaussian resolution function to model the partially-reconstructed background and an exponential function to model mesonic cross-feed distributions and model the combinatorial background. To evaluate systematic uncertainty from these sources I make the following changes to the model:

- The signal and normalisation models are replaced with a double Hypatia function.
- The cross-feed models and partially reconstructed background models are replaced with RooKeysPdf kernel density estimates.
- The combinatorial background model is replaced with a second order Chebychev polynomial.
- In order to simplify the fit, a reduced set of cross-feed and partially reconstructed contributions is included in the baseline model (see Sec. 4.4.1.2). The effect of this is evaluated by considering the effect on the signal of including additional fit components.
- In the normalisation fit model, to account for the data/MC differences in the mean and width of different (signal and cross-feed) components modelled as either single or double

Crystal Ball PDFs, I have introduced the floating $\mu_{\rm shift}$ and $\sigma_{\rm scale}$ parameters respectively.

I expect similar data/MC differences in the fit to the signal modes. In my nominal fit model, the μ of the signal component (which is modelled as DCB) is fixed to the most precise measurement of Ξ_b^- (Ω_b^-) mass. For baryonic cross-feed, to simplify the fit model, I fix it to the one obtained from a fit to MC and do not introduce a floating $\mu_{\rm shift}$ term here. I evaluate a systematic uncertainty arising from this effect by introducing a common $\mu_{\rm shift}$ parameter on the means of baryonic cross-feed components, whose value will be fixed to the one taken from the normalisation fit. The difference in the expected signal yields obtained from this procedure and from my nominal fit model is assigned as a systematic.

In the case of the width of signal, baryonic cross-feed and PRB components, I have assumed that the magnitude of the $\sigma_{\rm scale}$ is the one taken from the normalisation mode fit. To evaluate a systematic uncertainty arising from this assumption, I first do a fit to the normalisation mode using DCB PDFs with non-common widths and two separate scale factors for each width. I then use these separate values of the scale factors obtained from the normalisation mode in the fit to signal modes to evaluate the systematic uncertainty (the biggest of the systematic uncertainty from the two scales is used).

- Another simplification of the fit is made by constraining some of the parameters to be identical to each other (see Sec. 4.4.2). The systematic uncertainty from these choices is evaluated by running the fit with different configurations. One such constraint is the requirement that the slope parameter of the combinatorial background be the same for all signal final states.

Table 4.23 lists all the systematics from the fit model choice. The systematics are reported in terms of the difference (as a percentage) in the ratio of the signal yield to the normalisation yield. We will also quote the total systematic from the fit model – the sum in quadrature of all model systematics for a particular mode.

4.6.2.2 Fixed parameters

To increase the simultaneous fit stability certain shape parameters are fixed in the final fit. To estimate the systematic uncertainty introduced by these fixed parameters, I generate an ensemble of toy pseudo-experiments based on the nominal fit results. Using the covariance matrix from the fits to simulation I also generate new sets of values for the fixed PDF parameters. Every toy experiment is fitted using each of these new sets of values as well as the nominal values. The distribution of the difference between the yield returned by the fit using the nominal parameter values and the yields from each of the fits using the modified parameter values is plotted for

Table 4.23: Absolute systematic uncertainties arising from various choices of fit model on the ratio of signal yield to the normalisation yield. The total is the sum in quadrature of all other entries.

Mode	Model I (10^{-6})	Model II (10^{-6})	Model III (10^{-6})	Model IV (10^{-6})	Model V (10^{-6})	Total (10^{-6})
$\Xi_b^- \to pK^-K^-$	4.96	10.05	11.07	1.69	24.46	29.15
$\Xi_b^- \to pK^-\pi^-$	4.07	25.99	46.57	33.40	18.64	65.76
$\Xi_b^- \to p\pi^-\pi^-$	11.09	51.76	91.03	80.34	79.56	154.52
$\Omega_b^- \to pK^-K^-$	1.98	13.66	13.63	5.67	18.73	27.56
$\Omega_b^- \to pK^-\pi^-$	53.00	65.10	65.24	64.78	14.53	125.35
$\Omega_b^- \to p\pi^-\pi^-$	42.58	49.78	29.74	69.85	16.31	101.60

every toy experiment and the standard deviation of each of these distributions is determined. The systematic uncertainty is assigned to be the average value of the standard deviation over the ensemble of toy experiments. Table 4.24 shows the absolute systematic uncertainties on the ratio of the signal yield to the normalisation yield arising due to this procedure.

Table 4.24: Absolute systematic uncertainties arising from fixing various shape parameters on the ratio of signal yield to the normalisation yield.

Mode	Fixed Parameters (10^{-6})
$\Xi_b^- \to pK^-K^-$	24.02
$\Xi_b^- \to pK^-\pi^-$	17.81
$\Xi_b^- \to p\pi^-\pi^-$	9.98
$\Omega_b^- \to pK^-K^-$	0.87
$\Omega_b^- \to pK^-\pi^-$	2.59
$\Omega_b^- \to p\pi^-\pi^-$	5.91

4.6.2.3 Fit bias

Using a similar procedure as the blind fit stability study in Sec. 4.4.3, multiple toy pseudo-experiments are generated with the yield and fit parameters extracted from the nominal unblinded fit to signal modes. The results of the study show that all fitted yields are consistent with being unbiased; since the precision of this statement is limited by the size of the ensemble generated, the difference in the mean yield value of the toys and the nominal yield will be assigned as a systematic uncertainty. Table 4.25 shows the absolute systematic uncertainties on the ratio of the signal yield to the normalisation yield arising due to this procedure. These uncertainties are negligible compared to other sources.

Table 4.25: Absolute systematic uncertainties arising from the fit bias on the ratio of signal yield to the normalisation yield.

Mode	Fit bias (10^{-6})
$\Xi_b^- \to pK^-K^-$	0.24
$\Xi_b^- \to pK^-\pi^-$	7.51
$\Xi_b^- \to p\pi^-\pi^-$	4.50
$\Omega_b^- \to pK^-K^-$	1.69
$\Omega_b^- \to pK^-\pi^-$	6.59
$\Omega_b^- \to p\pi^-\pi^-$	6.94
$\Xi_b^- \to pK^-K^-$	10.63
$\Xi_b^- \to pK^-\pi^-$	4.65
$\Xi_b^- \to p\pi^-\pi^-$	14.62
$\Omega_b^- \to pK^-K^-$	51.31
$\Omega_b^- \to pK^-\pi^-$	159.1
$\Omega_b^- \to p\pi^-\pi^-$	10.11

4.7 Results of the analysis

In this analysis, I have searched for the decays of the Ξ_b^- and Ω_b^- baryons to the charmless final states $ph^-h'^-$ using the proton-proton collision data collected by LHCb at centre-of-mass energies $\sqrt{s} = 7$ and $8\,\text{TeV}$, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. Taking into account both statistical and systematic uncertainties arising from the fitting procedure, the $\Xi_b^- \to pK^-K^-$ decay is observed with a significance of 8.7 standard deviations, and evidence at the level of 3.4 standard deviations is found for the $\Xi_b^- \to pK^-\pi^-$ decay. To obtain these significances one first calculates the profile likelihood ratio ($-2\ln\frac{\mathcal{L}(N_{\text{sig}})}{\mathcal{L}(N_{\text{best}})}$) as function of signal yield (N_{sig}), where N_{best} denotes the best estimate of the signal yield from the likelihood minimisation. The profile likelihood ratio is then convolved with a Gaussian of mean zero and width given by the systematic uncertainty arising from the fitting procedure alone. By invoking Wilks theorem, the p-value is obtained from $\sqrt{-2\Delta\ln\mathcal{L}}$ assuming a χ^2 distribution with degrees of freedom equal to the difference in dimensionality of $\mathcal{L}(N_{sig})$ and $\mathcal{L}(N_{\text{best}})$. The p-value is then converted into equivalent "Gaussian significance" using the relation $\sqrt{2}\,\text{erfc}^{-1}(p)$.

In this analysis, a charmless decay of a baryon containing both beauty and strange quarks (Ξ_b) is observed for the first time. The measurement of the branching fractions and investigation of the Dalitz plot distribution of the observed $\Xi_b^- \to pK^-K^-$ mode is presented in the following subsections.

4.7.1 Branching fractions

In this subsection, I report the products of branching fractions and b-hadron production fractions relative to the $B^- \to K^+K^-K^-$ normalisation channel. These are determined using Eq. 4.2 for all channels,

$$
\begin{aligned}
\frac{\mathcal{B}(\Xi_b^- \to pK^-K^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Xi_b^-}}{f_u} &= (\ 265 \pm 35\,(\text{stat}) \pm 47\,(\text{syst})\ ) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Xi_b^- \to pK^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Xi_b^-}}{f_u} &= (\ 259 \pm 64\,(\text{stat}) \pm 49\,(\text{syst})\ ) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Xi_b^- \to p\pi^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Xi_b^-}}{f_u} &= (\ 74 \pm 40\,(\text{stat}) \pm 36\,(\text{syst})\ ) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Omega_b^- \to pK^-K^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Omega_b^-}}{f_u} &= (\ -9 \pm\ \ 9\,(\text{stat}) \pm\ \ 6\,(\text{syst})\ ) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Omega_b^- \to pK^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Omega_b^-}}{f_u} &= (\ -23 \pm 28\,(\text{stat}) \pm 23\,(\text{syst})\ ) \times 10^{-5}\,,
\end{aligned}
$$

$$
\frac{\mathcal{B}(\Omega_b^- \to p\pi^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Omega_b^-}}{f_u} = (\ 48 \pm 33\,(\text{stat}) \pm 28\,(\text{syst})\ ) \times 10^{-5}\,.
$$

The branching fractions of $\Xi_b^- \to pK^-\pi^-$ and $\Xi_b^- \to p\pi^-\pi^-$ relative to $\Xi_b^- \to pK^-K^-$ decays are also measured

$$
\begin{aligned}
\frac{\mathcal{B}(\Xi_b^- \to pK^-\pi^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} &= (\ 977 \pm 274\,(\text{stat}) \pm 85\,(\text{syst})\ ) \times 10^{-3}\,, \\
\frac{\mathcal{B}(\Xi_b^- \to p\pi^-\pi^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} &= (\ 279 \pm 155\,(\text{stat}) \pm 132\,(\text{syst})\ ) \times 10^{-3}\,.
\end{aligned}
$$

Since the signals for $\Xi_b^- \to p\pi^-\pi^-$, $\Omega_b^- \to pK^-K^-$, $\Omega_b^- \to pK^-\pi^-$ and $\Omega_b^- \to p\pi^-\pi^-$ are not significant, I set upper limits on their relative branching and fragmentation fractions. Specifically, the likelihood obtained from the fit (as a function of signal yield) is converted into a likelihood as a function of the relative branching and fragmentation fractions. This is then convolved with a Gaussian of width given by the total systematic uncertainty. The upper limits are set by integrating the likelihood after multiplying by a prior probability distribution that is uniform in the region of positive branching fraction. The obtained upper limits

are

$$\begin{aligned}
\frac{\mathcal{B}(\Xi_b^- \to p\pi^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Xi_b^-}}{f_u} &< 147\ (166) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Omega_b^- \to pK^-K^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Omega_b^-}}{f_u} &< 18\ (22) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Omega_b^- \to pK^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Omega_b^-}}{f_u} &< 51\ (62) \times 10^{-5}\,, \\
\frac{\mathcal{B}(\Omega_b^- \to p\pi^-\pi^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Omega_b^-}}{f_u} &< 109\ (124) \times 10^{-5}\,,
\end{aligned}$$

where the limits are at 90 (95) % confidence level.

I also quote the upper limits at 90 (95) % confidence level for the relative branching fraction of $\Xi_b^- \to p\pi^-\pi^-$ with respect to $\Xi_b^- \to pK^-K^-$.

$$\frac{\mathcal{B}(\Xi_b^- \to p\pi^-\pi^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} < 554\ (626) \times 10^{-3}\,,$$

Figure 4.29, shows $\Delta \ln \mathcal{L}$ scans as functions of the relative branching fraction and fragmentation fractions. For non-significant modes, the upper limits are also shown. I also show $\Delta \ln \mathcal{L}$ scans as functions of the relative branching fraction of $\Xi_b^- \to pK^-\pi^-$ and $\Xi_b^- \to p\pi^-\pi^-$ with respect to $\Xi_b^- \to pK^-K^-$ in Figure 4.30.

4.7.2 Dalitz plot of $\Xi_b^- \to pK^-K^-$ mode

In this subsection, I inspect any quasi-two-body contributions to the observed $\Xi_b^- \to pK^-K^-$ mode. The signal *sWeighted* distributions of the $\Xi_b^- \to pK^-K^-$ data in the Dalitz plot is shown in Figure 4.31. In the $m(Kp)_{\min}$ spectrum, one can notice a clear peak from $\Lambda(1520)$ ($J^P = \frac{3}{2}^-$) and another that might arise from the combination of $\Lambda(1670)$ ($\frac{1}{2}^-$) and $\Lambda(1690)$ ($\frac{3}{2}^-$) resonances, together with broader contributions that extend above 2000 MeV. Compared to the pK^- structures seen in the amplitude analysis of $\Lambda_b^0 \to J/\psi pK^-$ (see Figure 4.32) [21], there appears to be less contribution from $\Lambda(1600)$ and $\Lambda(1810)$, both of which have $J^P = \frac{1}{2}^+$, in my data. There seems to be no clear resonant contribution in $m_{\max}(pK)$ and no exotic contribution in the $m_{\min}(KK)$ distribution.

The plan of the next chapter is to increase the yield of $\Xi_b^- \to pK^-K^-$, investigate the quasi-two-body contributions and search for *CP* violating effects through an amplitude analysis

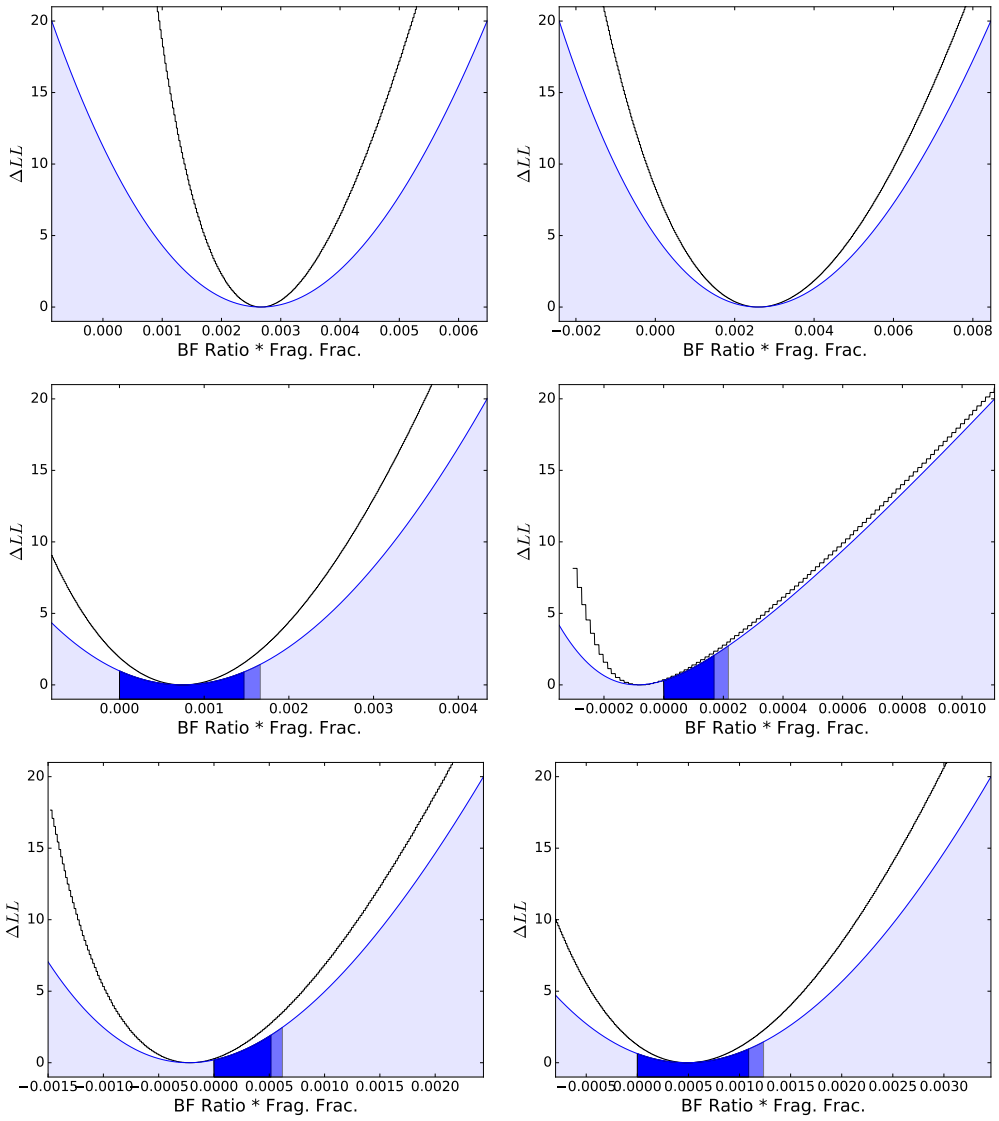

Figure 4.29: $\Delta \ln \mathcal{L}$ scans as functions of relative branching fraction and fragmentation fractions for the signal modes. For non-significant modes, blue shaded areas which represent the 90 (95) % confidence intervals. The black line is with statistical uncertainty only and the blue line includes systematic uncertainties. From left to right, top to bottom: $\Xi_b^- \to pK^-K^-$, $\Xi_b^- \to pK^-\pi^-$, $\Xi_b^- \to p\pi^-\pi^-$, $\Omega_b^- \to pK^-K^-$, $\Omega_b^- \to pK^-\pi^-$ and $\Omega_b^- \to p\pi^-\pi^-$.

of this decay mode.

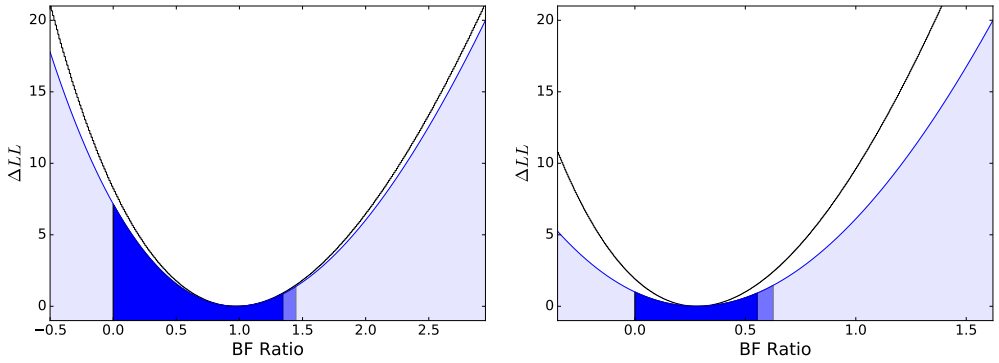

Figure 4.30: $\Delta \ln \mathcal{L}$ scans as functions of relative branching fraction of (left) $\Xi_b^- \to pK^-\pi^-$ and (right) $\Xi_b^- \to p\pi^-\pi^-$ with respect to the $\Xi_b^- \to pK^-K^-$ mode. The blue shaded areas represent the 90 (95) % confidence interval. The black line is with statistical uncertainty only and the blue line includes systematic uncertainties.

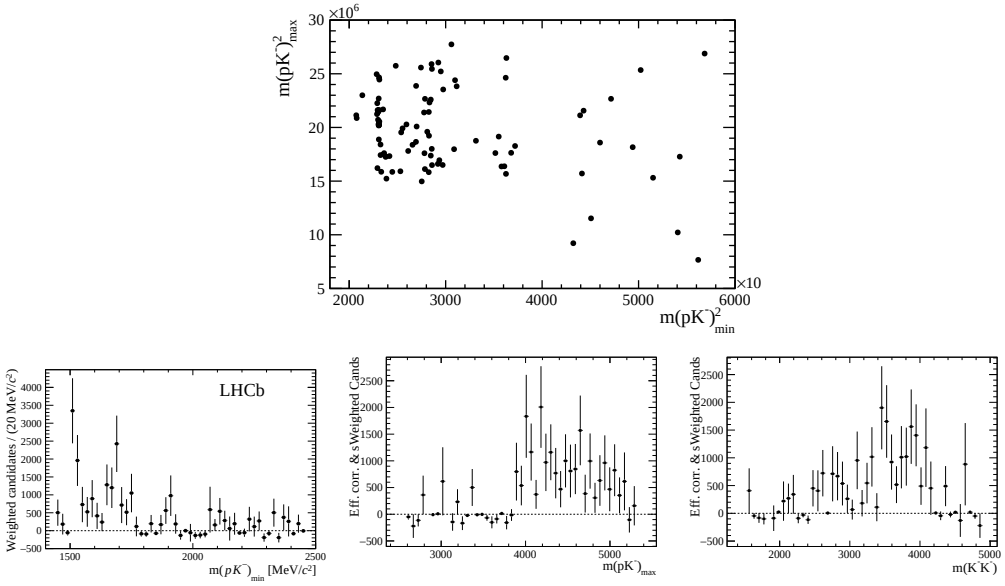

Figure 4.31: Scatter plot of the phase space (top) with background subtraction applied for $\Xi_b^- \to pK^-K^-$ data sample. (Bottom) Background-subtracted (through signal *sWeights*) and efficiency-corrected 1D projections of $m_{\min}(pK)$, $m_{\max}(pK)$ and $m_{\min}(KK)$. Here I have folded the DP due to the Bose symmetry of the indistinguishable pseudo-scalar particles (K_1, K_2) in the final state.

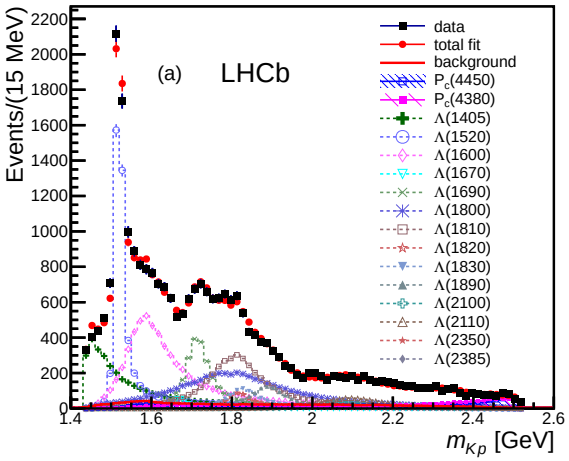

Figure 4.32: Results for $m(K^-p)$ for the best model obtained for $\Lambda_b^0 \to J/\psi p K^-$ decays that contains two charmonium-pentaquark (P_c^+) states [21]. The data are shown as (black) squares with error bars, while the (red) circles show the results of the fit. Each Λ^* component is also shown. The (blue) open squares and (purple) solid squares show the two P_c^+ states.

CHAPTER 5

Search for CP violation in $\Xi_b^- \to pK^-K^-$ decays

“ Sometimes, if you pay real close attention to the pebbles you will find out about the Ocean. ”

Sir Terry Prachett, *Lords and Ladies*

5.1 Introduction

The phenomenon of CP violation has been well studied in B mesons decays, however no significant manifestation of CP violation has yet been observed in the decay of any baryon. It is therefore of great interest to search for CP violation in the baryon sector and thereby test the Standard Model prediction. It was mentioned in the previous chapter that, limits on (phase-space integrated) CP violation parameters have been set for $\Lambda_b^0 \to pK^-$ and $p\pi^-$ [106, 138], $K^0_S p\pi^-$ [102], ΛK^+K^- and $\Lambda K^+\pi^-$ channels [104]. Searches for asymmetries in regions of the phase-space of the four-body charmless hadronic $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$, $\Lambda_b^0 \to p\pi^-K^+K^-$, $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$ and $\Xi_b^0 \to pK^-\pi^+K^-$ decays have also been carried out, with $3.3\,\sigma$ evidence of a CP violation effect in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays [107, 108].

Since large CP asymmetries have been seen in B meson decays to three charged particles [63, 99, 100], it is natural to look for such effects in similar decays of b baryons. In the previous chapter, I searched for Ξ_b^- (bsd) and Ω_b^- (bss) decays to $ph^-h'^-$ final states with LHCb

Run I (2011–2012) data, which revealed the first observation of $\Xi_b^- \to pK^-K^-$ decays [135]. This channel is therefore the most attractive of the $ph^-h'^-$ channels with which to search first for CP violation effects.

The analysis performed in the previous chapter resulted in a signal yield of around 80 $\Xi_b^- \to pK^-K^-$ decays, which is not sufficient for a detailed study of CP violation effects varying across phase space. In order to have sensitivity to such effects, it is necessary to increase the yield as much as possible, this will be achieved in two ways.

The first is to include Run II (2015–2016) data. The analysis described in this chapter is based on the data sample accumulated by LHCb in 2011-2016, comprising $1\,\text{fb}^{-1}$ collected at pp collision centre-of-mass energy $\sqrt{s} = 7\,\text{TeV}$ (2011), $2\,\text{fb}^{-1}$ collected at $\sqrt{s} = 8\,\text{TeV}$ (2012) and $2\,\text{fb}^{-1}$ collected at $\sqrt{s} = 13\,\text{TeV}$ (2015–16). Since the b production cross-section varies with $\sqrt{s}$, and due to changes in operating conditions (in particular, trigger efficiency) between run periods, we can estimate as increase in yield of more than a factor of two with respect to the previous chapter.

The second improvement in yield comes from a better signal efficiency, achieved in this chapter by improving the selection criteria. In particular, a dedicated stripping line (as compared to the line designed for $B^+ \to h^+h'^-h''^+$ decays used in the analysis presented in the previous chapter) will be exploited. Further improvement comes from a more sophisticated use of the particle identification variables and the multivariate classifier: the key tools to discriminate signal from background. An increase of efficiency of around a factor of 2 has been achieved, for comparable background levels.

Phase space dependent CP violation effects in multibody b-hadron decay can be searched using either model-independent or a model-dependent approach. The former can be used with smaller data samples, and relies less on understanding the resonant structures that are present in the decay. It is therefore of limited use in associating any observed effects to the underlying physical decay processes. The model-dependent approach aims at a complete description of the underlying physical processes and therefore has, in principle, the best possible sensitivity. In practice, uncertainties associated with the modelling of the hadronic resonances and other issues associated with the complexity of the fit (*e.g.* instabilities, correlations between parameters, multiple solutions) can be significant and in some cases prevent a clear interpretation of the results.

Until recently, very few amplitude analyses of multibody baryon decays had been carried out. However, LHCb has shown that the challenges can be met, at least for high yield channels such as $\Lambda_b^0 \to J/\psi\, pK^-$ [21], $J/\psi\, p\pi^-$ [139] and $D^0 p\pi^-$ [126]. It therefore seems timely to attempt to address the challenge of performing an amplitude analysis of a charmless three-body b baryon decay. This challenge includes the fact that CP violation effects should be included for

the first time in any amplitude analysis of a baryon decay.

In Sec. 5.2 of this chapter, I discuss how the signal selection efficiency is improved. In Sec. 5.3, I determine the signal and background fit fractions that form inputs to the amplitude analysis of $\Xi_b^- \to pK^-K^-$ decays. An update on the result for the ratio of production fractions and branching fractions $\left(f_{\Omega_b^-}/f_{\Xi_b^-}\right) \times \left(\mathcal{B}\left(\Omega_b^- \to pK^-K^-\right)/\mathcal{B}\left(\Xi_b^- \to pK^-K^-\right)\right)$ is included. The description of the amplitude analysis, and the method used to establish the baseline model, is given in Sec. 5.4. Since the analysis is currently blind to the CP violation parameters, I present the sensitivity that can be achieved on the CP violation parameters with the present data using a toy study in Sec. 5.5. Finally, the future work involved and the summary of the chapter is presented in Sec. 5.6.

5.2 Gain in the signal selection efficiency

In this section, I discuss in detail the signal selection procedure employed to improve the selection efficiency by a factor > 2 for Run I (2011–2012) data when compared to previous chapter. The changes to the signal selection procedure are discussed in Sec. 5.2.1 and the variation of the efficiency of the selection procedure over the phase space is determined in Sec. 5.2.2.

5.2.1 Changes to the signal selection

The improvement in the signal selection has been achieved by use of a dedicated stripping line, an improved use of particle identification variables and use of a better performing multivariate classifier algorithm. Details are discussed in the following sub-sections.

5.2.1.1 Stripping

This analysis, is based on a data sample that consists of candidates that are reconstructed from three charged particles. This sample undergoes a set of pre-selection requirements that is different to the analysis presented in the previous chapter. In this analysis, the reconstructed candidates are selected using the `StrippingXb2phhLine` stripping line designed to preserve the whole phase space of the decays, and in particular the low two-body invariant mass region, for the ph system, where $\Lambda^{*0}(1520)$ and (for $p\pi^-$) the $N^{*0}(1520)$ resonances are expected to be present. Only a loose requirement on the kinematic variables is applied on the daughter tracks and none on the two-body pairs is applied. The two like-charged particles have no particle identification (PID) requirement, but the particle that has been assigned a proton mass has a loose ProbNNp requirement applied in order to reduce the retention rate to an acceptable level. Moreover, in order to avoid border effects in the invariant mass of Ξ_b^- (Ω_b^-) candidates, the minimum invariant

mass is calculated using the pKK mass hypothesis, while the maximum is calculated using the $p\pi\pi$ mass hypothesis.

A summary of the stripping line selection requirements is presented in Table 5.1. All topological selections used in this analysis are in terms of variables significance, in order retain the best reconstructed candidates irrespective of the absolute variable value *e.g.* flight distance χ^2 is used instead of flight distance. The requirement on the minimal $p_{\rm T}$ of the daughters has been tuned in order to cope with the allowed retention rate and to preserve the signal efficiency of the quasi 2-body decays at threshold for the intermediate resonances.

Table 5.1: Requirements imposed in the stripping line `StrippingXb2phhLine`.

Variable definition	Requirement
Requirements on daughter tracks (p,π)	
Daughter tracks' momentum	$p > 1500$ MeV/c
Daughter tracks' transverse momentum	$p_{\rm T} > 250$ MeV/c
Daughter tracks' χ^2 over degrees of freedom	Trk. $\chi^2/{\rm ndf} < 3.0$
Daughter tracks' minimum impact parameter χ^2 to any PV	Min. $\chi^2_{\rm IP} > 16.0$
Daughter tracks' probability of being a ghost track	Prob$ghost < 0.40$
Proton track's probability of being a proton	ProbNN$p > 0.05$
Combination requirements (before vertex fit)	
Mass of the X_b^- candidate with pKK tracks hypothesis	$m_{pKK} > 5195$ MeV/c^2
Mass of the X_b^- candidate with $p\pi\pi$ tracks hypothesis	$m_{p\pi\pi} < 6405$ MeV/c^2
X_b^- candidate's transverse momentum	$p_{\rm T} > 1500$ MeV/c
Vector sum of the daughter tracks' transverse momentum	$\Sigma_{\rm daug.} p_{\rm T} > 3500$ MeV/c
Distance of closest approach χ^2 of any two daughters	$\chi^2_{\rm DOCA} < 20$
Combination requirements (after vertex fit)	
X_b^- candidate's vertex χ^2	$\chi^2_{\rm vtx} < 20$
X_b^- candidate's flight distance χ^2 w.r.t. best PV	$\chi^2_{\rm FD} > 50$
X_b^- candidate's impact parameter χ^2 w.r.t. best PV	$\chi^2_{\rm IP} < 16$
Cosine of the X_b^- candidate pointing angle	$\cos(\theta_{\rm DIRA}) > 0.9999$

5.2.1.2 Trigger

The ability to associate the offline particles with the trigger objects allows us to form TOS and TIS categories (See Sec. 3.5 for more details). Hadronic final states are selected, at the L0 hardware trigger stage, requiring either the L0 hadron trigger firing as `TOS` or the global L0 trigger firing as `TIS`. At the first stage of the high level trigger (HLT1) candidates are re-

quired to have high p_{T} and include tracks with significant impact parameter (IP) (triggering the `Hlt1TrackAllL0Decision_TOS` line and either `Hlt1TrackMVADecision_TOS` or `Hlt1TwoTrackMVADecision_TOS` for Run I (2011–2012) and Run II (2015–2016), respectively). At the second stage of the high level trigger (HLT2) one of the topological triggers is required to fire.

These are implemented for Run I (2011–2012) as requiring

```
(L0Global_TIS || L0HadronDecision_TOS) &&
     Hlt1TrackAllL0Decision_TOS &&
   (Hlt2Topo{2,3}BodyBBDTDecision_TOS)
```

and Run II (2015–2016) as requiring

```
          (L0Global_TIS || L0HadronDecision_TOS) &&
(Hlt1TrackMVADecision_TOS || Hlt1TwoTrackMVADecision_TOS)&&
              (Hlt2Topo{2,3}BodyDecision_TOS)
```

5.2.1.3 Offline selection

Candidates with Muons in the final state are vetoed by requiring `!isMuon` in data and MC samples prior to MVA training and selection. The invariant mass of pKK ($m(pKK)$) is required to lie within the fit range $[5545, 6470]\,\mathrm{MeV}/c^2$. The lower limit has been chosen to minimise both loss in the signal efficiency and in background contribution. The upper limit is chosen to avoid any effect on $m(pKK)$ from the stripping requirement on $m(p\pi\pi)$.

Use of particle identification (PID) information: From the previous chapter, we learned that the major background contribution originated from combinatorial background, consisting mainly of random pion tracks from the event reconstructed with the mass hypothesis of the signal final state particles. Significant backgrounds can also occur from cross-feed, in which a b-hadron decay has a final state particle misidentified. Since the global neural network based (ProbNN) PID variables were designed to reduce mis-identification of final state particles, developing a selection requirement that takes into account the correlation of the PID variables with other topological and kinematic variables could be beneficial in discriminating signal from background. Hence, these PID variables are included in the MVA training along with other discriminating variables.

However, we have already discussed how the response of the PID is not very well modelled in simulation. Therefore, one needs to either resample or correct these PID variables using calibration samples. I use the `PIDGen` package that facilitates resampling, with the limitation here that only one ProbNN variable per track can be used since any correlation with

other ProbNN variables of the same track is lost (therefore as a future improvement one can use the `PIDCorr` package which is under development). The corrections are implemented using PDFs that are functions of the PID variable being resampled, the track p, η and event multiplicity (`nTracks`) variables. Therefore I choose to resample the following PID variables using `PIDGen`:

1st final state particle (K mass hypo.) : `h1_PROBNNK` ,
2nd final state particle (K mass hypo.) : `h2_PROBNNK` ,
3rd final state particle (p mass hypo.) : `p_PROBNNp` .

Note that the PDF templates of the calibration samples can in principle be different for particle and their charge conjugate as they can interact differently with the sub-detectors that provide the PID information. To account for such asymmetry in PID efficiencies the charge of the particle is taken into account and separate templates are used.

MVA training samples and input variables: The following samples are used as signal and background samples to train the MVA algorithm:

- Signal sample: MC sample with all the above selection criteria applied. This sample consists of 135055 and 324958 signal MC candidates for Run I (2011–2012) and Run II (2015–2016), respectively.
- Background sample: Data sample with all the above selection criteria applied, excluding the $m(pKK)$ selection, which is modified to select candidates that lie in the signal region defined as $5600 < m(pKK) < 6250\,\mathrm{MeV}/c^2$, *i.e.* approximately between the Ξ_b^- mass $(5794.4\,\mathrm{MeV}/c^2) - 10\sigma$ and the Ω_b^- mass $(6049\,\mathrm{MeV}/c^2) + 10\sigma$ where σ is the signal resolution and is approximately $16\,\mathrm{MeV}/c^2$ (obtained from the fit to $m(pKK)$ in the $\Xi_b^- \to pK^-K^-$ MC sample). This sample consists of 563497 and 1367519 background data candidates for Run I (2011–2012) and Run II (2015–2016), respectively.

It should be noted that the background sample contains sidebands both above and below the wide signal region. In many analyses, only the upper sideband is used, since this region is usually dominated by combinatorial background with little contribution from the more signal-like partially reconstructed or cross-feed backgrounds. In such a case, the MVA is used essentially to discriminate against combinatorial background only, and other selections are typically used to reject other sources of background (for example, particle identification requirements in the case that the associated variables are not included in the MVA). In the current case, however, the variables that are most powerful to reject cross-feed (namely the particle identification variables)

and partially reconstructed backgrounds[1] are included in the MVA. Therefore, my MVA classifier is used to reject all types of background and hence trained it accordingly, with a background sample that contains all backgrounds.

As before, it was checked that the input variables for MVA training:

- Exhibit good discriminating power between signal and background samples. A list of the input variables used to train the MVA with their definitions is shown in Table 5.2. The discriminating power of these variables can be seen in Figures 5.1–5.2.

- Exhibit good agreement between MC and background-subtracted data. This issue was studied by using a control mode ($B^- \to p\overline{p}K^-$) that has high signal statistics and a similar topology to my signal mode (see Appendix B for more detail). It was concluded from this study that the variables used in the classifier training do not show any discrepancy between simulated and background-subtracted data samples.

- Show no significant correlation between the phase space variables (Dalitz plot and angular variables), $m(pKK)$ and between each other. The correlation matrices of the input variables for the background and signal samples for Run I (2011–2012) are shown in Figure 5.3. A similar correlation matrix is obtained for Run II (2015–2016) samples. It was also checked that the MVA output requirement does not introduce any bias in the $m(pKK)$ distribution in a purely combinatorial background sample *i.e.* a sample that consists of like-charged final state particles, referred to here as the "same sign sample". In other words, variables that introduce fake structures in the $m(pKK)$ distribution in the same sign sample after applying the MVA output requirement were removed from the training.

MVA training performance: In this section I investigate the performance of different types of MVA classifier algorithms and select that which gives the best performance.

I investigate decision tree based algorithms such as simple decision tree [140], random forest [141], gradient boosted decision tree (`xGBoost`) [142] and uniform boosted decision tree (`uBoost`) [143]. The aim of `uBoost` is to maintain an uniform efficiency of MVA output requirement over a certain set of chosen variables such as phase space variables. This algorithm took a lot of time to train and gave a lower accuracy of prediction in the test sample compared to the random forest and `xGBoost` methods. Since my input variables do not show any significant correlation with the phase-space variables, small gain was achieved from employing `uBoost` and it was not considered in further investigations.

[1]The isolation requirements such as $p_{\rm T}$ asymmetry for a cone in the η–ϕ plane and $\Delta\chi^2_{\rm vtx}$, calculated by adding and removing the most compatible extra track from B cand vertex fit

Table 5.2: List of the input variables that enter the training of the MVA algorithm. Note that the transformations on the variables are made to smooth out any sharp structures in the distributions.

Variable definition	Short form
X_b^- candidate pointing angle ($\theta_{\rm DIRA}$).	COSINV_Xib_DIRA_OWNPV
$p_{\rm T}$ asymmetry for a cone in η–ϕ plane with radius 1.7	B_STRIP_PTASYM_1_7
$ln(\Delta\chi^2_{\rm vtx}+2)$, calculated by adding and removing the most compatible extra track from B cand vertex fit	Transf_B_STRIP_VTXISODCHI2ONETRACK
ln($\chi^2_{\rm IP}$ of the leading $p_{\rm T}$ track)	Log_IPCHI2_LEADING_DAUG
ln(B cand vertex χ^2/ndof)	Log_Xib_ENDVERTEX_CHI2NDOF
ln($\chi^2_{\rm IP}$ of B candidate wrt PV)	Log_Xib_IPCHI2_OWNPV
ln($\chi^2_{\rm FD}$ of B candidate wrt PV)	Log_Xib_VDCHI2_OWNPV
$\sum p_{\rm T}$ of all tracks	sum_h_PT
PID var. for 1st track with K mass hypo.	h1_PROBNNK
PID var. for 2nd track with K mass hypo.	h2_PROBNNK
PID var. for 3rd track with p mass hypo.	p_PROBNNK

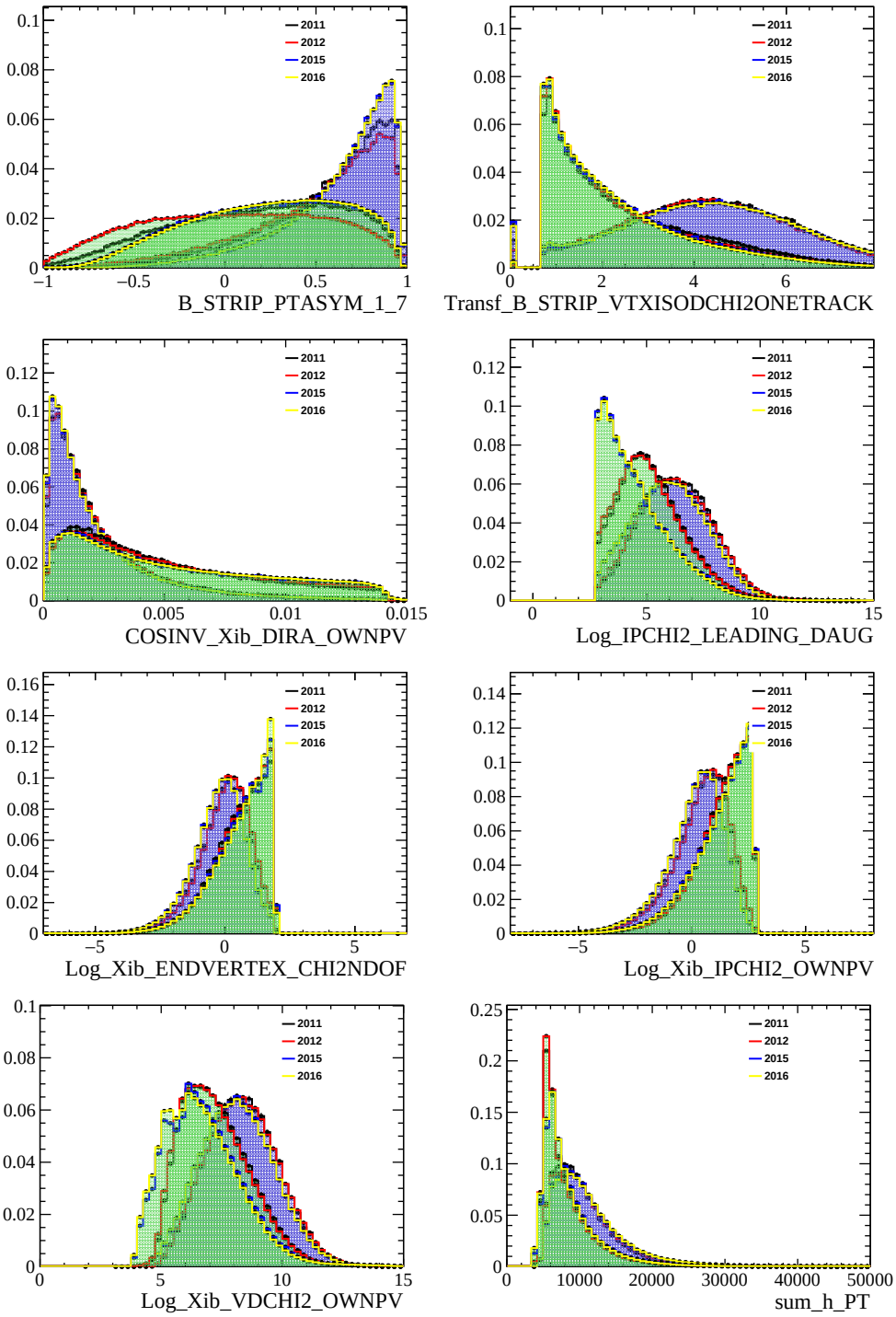

Figure 5.1: MVA input variable distributions for signal (blue) and background (green) samples for different data taking periods.

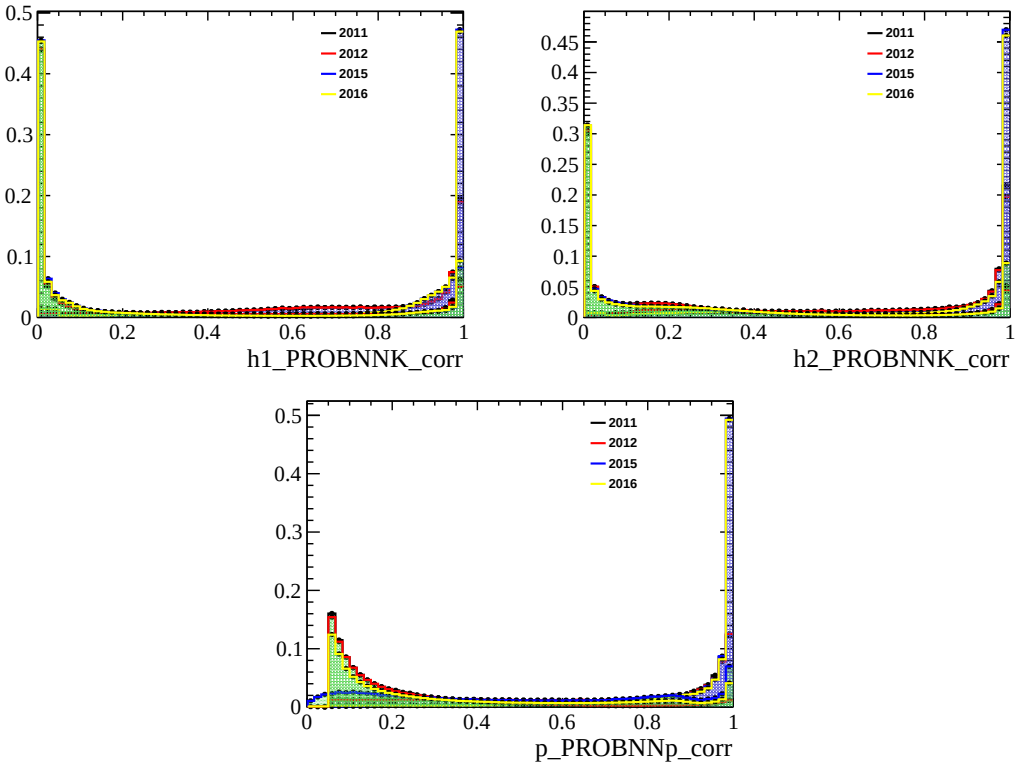

Figure 5.2: Global PID variables resampled using the `PIDGen` package for signal (blue) and background (green) samples for different data taking periods. Here 'h1', 'h2' and 'p' are K, K and p particles respectively. For the 2015 data sample, the stripping requirement `p_PROBNNp`>0.05 was applied with the Run I (2011–2012) tune (MC12TuneV2) instead of the Run II (2015–2016) tune (MC15TuneV1) which I use in the MVA training. Therefore, for the 2015 MC sample the resampling of `p_PROBNNp` was carried after the same requirement was imposed on the calibration samples.

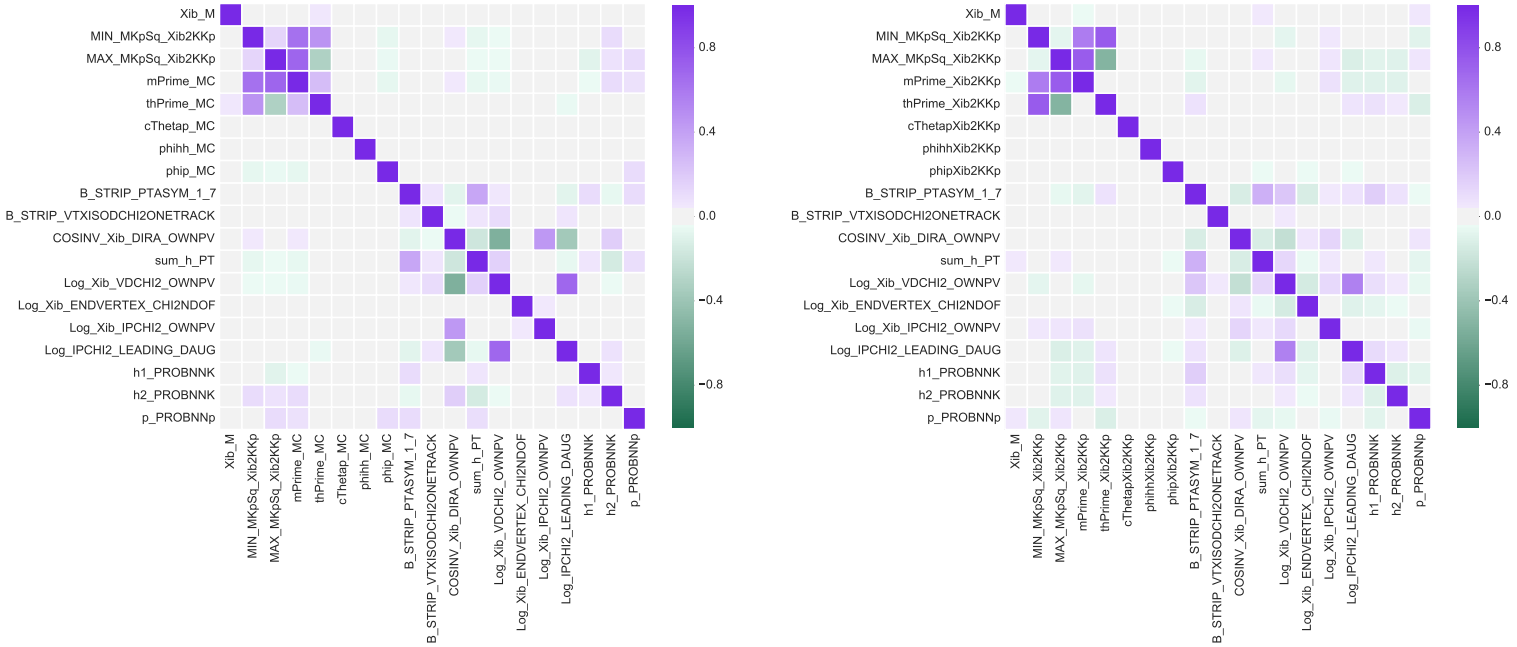

Figure 5.3: For Run I (2011–2012), correlation of the input variables with themselves, phase space variables (Conventional DP vars:`MIN_MKpSq, MAX_MKpSq`; Square DP vars:`mPrime,thPrime`; angular variables:`cThetap, phihh,phip`) and $m(pKK)$ (denoted here `Xib_M`) for (left) signal and (right) background samples. Note that the phase space variables have suffix `_MC` for signal and `Xib2KKp` for bkg samples.

I also investigate neural network based algorithms such as NEUROBAYES [144], `PyBrain` [145] and `Neurolab` [146]. The accuracy of the prediction of `Neurolab` was low compared to `PyBrain` and it was also time consuming to train, hence `Neurolab` was excluded from further consideration. All the neural network based algorithms were trained with one hidden layer with the same number of nodes as the number of input variables. Optimisation of the number of hidden layers and the number of nodes inside each hidden layer was carried out, however it was found that the gain in accuracy and ROC curve performance was very small at the expense of the time spent in training.

I use the implementation of some of the above algorithms in the `scikit-learn` package [147] and also use the Reproducible Experiment Platform (`REP`) python wrappers [148] when investigating the performance of the above algorithms.

To enhance the performance of the classifiers and also not over-train them, I investigated a few ensemble methods such as bagging, boosting and folding techniques [147]. In the end, I decided to apply manual folding to avoid over-training, *i.e.* I trained 2 classifiers separately for each run period on odd and even event number samples and evaluated them on the opposite event number samples.

The performance of the classifiers (ROC curves) on test samples for Run I (2011–2012) and Run II (2015–2016) is shown in Figure 5.4. The samples that contain even event number candidates show a similar trend. It can be seen that for both runs the performance of `xGBoost` is better than any of the other investigated classifiers. Therefore, I choose to employ this classifier to reduce background contamination.

Optimisation of MVA output requirement to reject combinatorial bkg: To optimise a requirement on the MVA output to reduce the combinatorial background contribution, I use $\frac{S}{\sqrt{S+B}}$ as a figure of merit (FoM). Here, 'S' and 'B' are respectively the number of signal and background candidates in the signal region, defined as $5700 < m(pKK) < 5900\,\text{MeV}/c^2$. The lower limit $\sim m(\Xi_b^-) - 2.5\sigma$ and upper limit $\sim m(\Xi_b^-) + 2.5\sigma$, (here $\sigma \approx 16\,\text{MeV}$ and the choice of 2.5 is chosen as it gives a high purity sample with minimal $\Xi_b^- \to pK^-\pi^-$ cross-feed contribution), are taken from the fit to the MC sample. An alternative FoM, $\frac{S^2}{(S+B)^{\frac{3}{2}}}$ is often used for amplitude analysis since it gives a cleaner sample, which can be better for determining relative phases; however, my main concern is to obtain a sufficiently large sample to make the amplitude analysis viable and therefore an MVA selection that results in a larger signal efficiency is preferred, while still controlling the background contribution. Hence, $\frac{S}{\sqrt{S+B}}$ is used as the FoM in my case.

The number of combinatorial background candidates (B) is estimated by fitting a first order Chebychev polynomial (*i.e.* a straight line) to the higher mass sideband (defined as

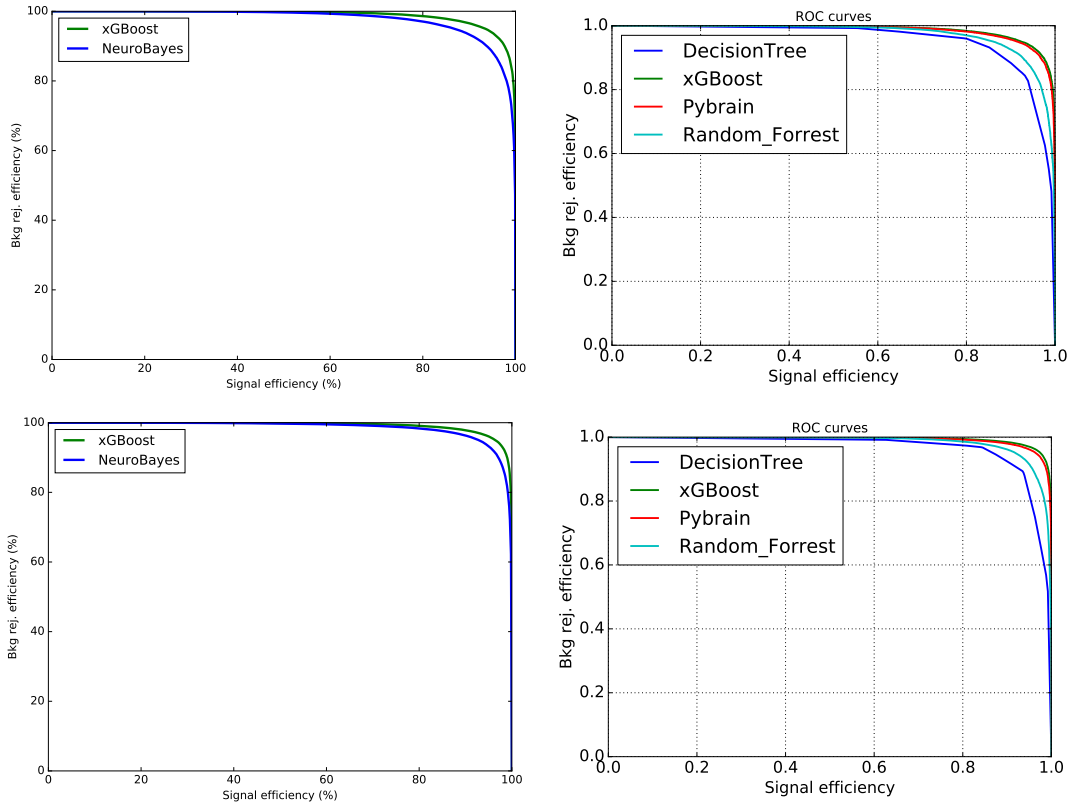

Figure 5.4: Performance of (left) NEUROBAYES algorithm and (right) the python based MVA algorithms on the test sample for (top) Run I (2011–2012) and (bottom) Run II (2015–2016) odd event number sample. Similar plots are seen for even event number sample.

$6125 < m(pKK) < 6470\,\mathrm{MeV}/c^2$, where the lower limit $\sim m(\Omega_b^-) + 5\sigma$ and the upper limit is the limit above which the stripping requirement on $m(p\pi\pi)$ is visible) and extrapolating the background yield into the signal region defined earlier. It may be noted that for the background estimate in the optimisation of the MVA output requirement I consider only combinatorial background (by extrapolating from an upper sideband), while in the MVA training I consider all types of background (by including both lower and upper sidebands). This does not introduce any inconsistency into the analysis as the dominant background in the signal region after optimisation of the MVA requirement is anyway expected to be combinatorial in nature.

The number of signal candidates (S) expected in the signal region is obtained using the following expression:

$$S = 2 \times \sigma(pp \to b\bar{b}X) \times \mathcal{L} \times f_{\Xi_b^-} \times \mathcal{B}(\Xi_b^- \to pK^-K^-) \times \epsilon^{\mathrm{tot}}\,. \tag{5.1}$$

It is know from previous chapter:

$$R_{\Xi_b^- \to pK^-K^-} = \frac{\mathcal{B}(\Xi_b^- \to pK^-K^-)}{\mathcal{B}(B^+ \to K^+K^-K^+)} \times \frac{f_{\Xi_b^-}}{f_u} = (265 \pm 35\,(\text{stat}) \pm 47\,(\text{syst})) \times 10^{-5} \quad (5.2)$$

so plugging Eq. 5.2 into Eq. 5.1, I obtain

$$\begin{aligned} S = 2 \times \sigma(pp \to b\bar{b}X) \times \mathcal{L} \\ \times (f_u \times \mathcal{B}(B^+ \to K^+K^-K^+) \times R_{\Xi_b^- \to pK^-K^-}) \times \epsilon^{\text{tot}}\,. \end{aligned} \quad (5.3)$$

In the above equations:

- The factor of '2' arises since both b and $\bar{b}$ can hadronise into Ξ_b^-.
- $\sigma(pp \to b\bar{b}X)$ is the $b\bar{b}$ production cross-section in the acceptance of the LHCb experiment which is measured in Ref. [149] to be $(72.0 \pm 0.3 \pm 6.8)$ µb and $(154.3 \pm 1.5 \pm 14.3)$ µb for collisions with centre-of-mass energies of 7 and 13 TeV, respectively. For 8 TeV collisions (data collected during 2012), I simply scale the $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s}$ = 7 TeV by $\frac{8}{7}$.
- $\mathcal{L}$ is the integrated luminosity.
- $f_{\Xi_b^-}$ and f_u are the fragmentation fractions for Ξ_b^- and B^- respectively. The value of $f_u = (33.7 \pm 2.2)\%$ is stated in Ref. [150] and is assumed to be the same for different data taking periods (this assumption does not have a significant impact on the optimal MVA output requirement).
- $\mathcal{B}$ stands for branching fraction: $\mathcal{B}(B^+ \to K^+K^-K^+) = (3.40 \pm 0.14) \times 10^{-5}$ as stated in the PDG [151].
- $R_{\Xi_b^- \to pK^-K^-}$ is given in Eq. 5.2.
- $\epsilon^{\text{tot}} = \epsilon^{\text{geom}} \times \epsilon^{\text{sel|geom}}$, where ϵ^{geom} is the geometrical efficiency (defined as the fraction of accepted b-hadron candidates whose charged daughters are all in the acceptance region of $10 < \theta < 400\,\text{mrad}$) taken from generated MC samples from the analysis presented in previous chapter to be $(53.16 \pm 0.16)\%$ and $\epsilon^{\text{sel|geom}}$ is the combined reconstruction, stripping, trigger and offline (includes MVA output requirement) selection efficiencies taken from MC samples.

It is important to note here that some of the assumptions that I have made in the calculation of 'S' may not hold for some data taking periods and since the FoM depends on 'S', I investigate

the dependence of the optimal MVA output requirement by varying 'S' by $\pm 20\%$ of its expected value. Figures 5.5–5.6 show the values of 'S', 'B', $\epsilon^{\text{sel|geom}}$ and the FoM as a function of the MVA output requirement for Run I (2011–2012) and Run II (2015–2016), respectively. It is evident from these figures that the optimal MVA output requirement does not significantly change when 'S' gets varied by $\pm 20\%$ of its original expected value. The optimal MVA output requirements are MVA> 0.945 and MVA> 0.965 for Run I (2011–2012) and Run II (2015–2016), respectively. It is interesting to note that this procedure predicts signal yields of around 170 in Run I (2011–2012) and around 300 in Run II (2015–2016).

It can be seen from Fig. 5.1–5.2 that the input variables that go into the training of MVA classifier show significant differences between Run I (2011–2012) and Run II (2015–2016) samples. Due to this reason we have trained the MVA classifier on these two samples separately, as a result we also choose to optimise a cut on these samples separately (one could also choose to optimise a MVA cut using a combined Run I (2011–2012) and Run II (2015–2016) FoM. Whether this choice would lead to a gain in signal efficiency and less background contribution would require a separate study).

Cross-feed background rejection: The major potential cross-feed background contribution to the $m(pKK)$ spectrum comes from $B^+ \to K^+K^-K^+$. I remove this contribution with a veto region in the $m(KKK)$ spectrum defined as $m(B^+) - 3\sigma < m(KKK) < m(B^+) + 3\sigma$. Figure 5.7 shows the veto of $B^+ \to K^+K^-K^+$ in $m(KKK)$ spectrum and the effect of the $B^+ \to K^+K^-K^+$ veto on $m(pKK)$ for different veto regions for Run I (2011–2012) and Run II (2015–2016) samples. It can be seen that this background would be significant without a veto, but is suppressed to negligible levels with the $\pm 3\sigma$ requirement. The final veto region of 3σ around the B^+ mass is selected since increasing the veto region leads to reduction in signal efficiency and no significant decrease in background.

There could potentially be a contribution from the $\Xi_b^- \to pK^-\pi^-$ decay for which evidence was found in the Run I (2011–2012) analysis [135]. The expected level of $\Xi_b^- \to pK^-\pi^-$ crossfeed background can be calculated using the following formula

$$\begin{aligned} N(\Xi_b^- \to pK^-\pi^-) = N(\Xi_b^- \to pK^-K^-) \times \frac{\mathcal{B}(\Xi_b^- \to pK^-\pi^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} \\ \times \frac{\epsilon^{\text{sel|geom}}(\Xi_b^- \to pK^-\pi^-)}{\epsilon^{\text{sel|geom}}(\Xi_b^- \to pK^-K^-)}, \end{aligned} \tag{5.4}$$

where N is the yield and $\epsilon^{\text{sel|geom}}$ includes all efficiency effects except those due to the geometry (which are assumed to be the same between $\Xi_b^- \to pK^-K^-$ and $\Xi_b^- \to pK^-\pi^-$ channels *i.e.* see Table 4.9). To obtain $N(\Xi_b^- \to pK^-\pi^-)$ from Eq. 5.4, I use:

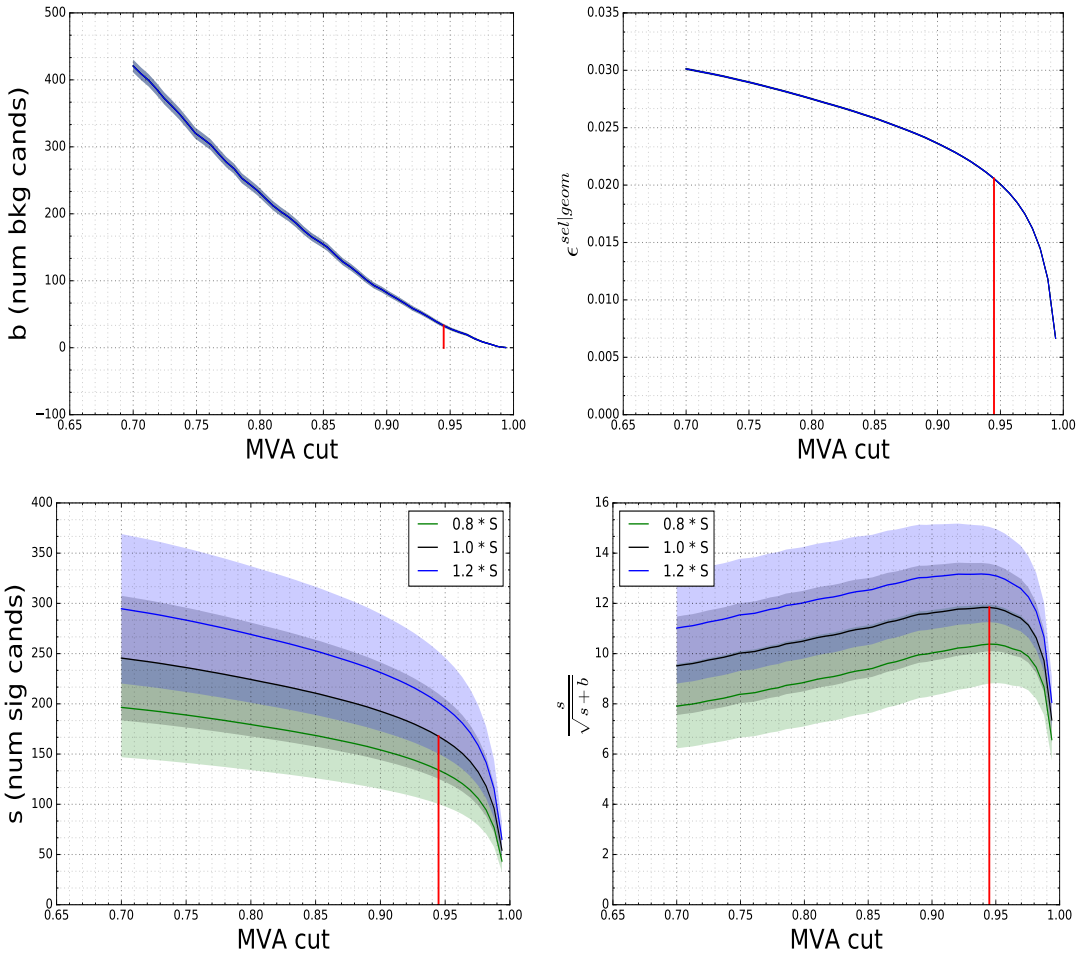

Figure 5.5: Optimisation of `xGBoost` classifier output requirement for Run I (2011–2012). (Top left) number of background candidates (B), (top right) full offline selection efficiency for signal ($\epsilon^{\text{sel}|\text{geom}}$), (bottom left) number of signal candidates (S) and (bottom right) FoM ($\frac{S}{\sqrt{S+B}}$) at different scaling factors for 'S'. The red line shows the optimal MVA output requirement.

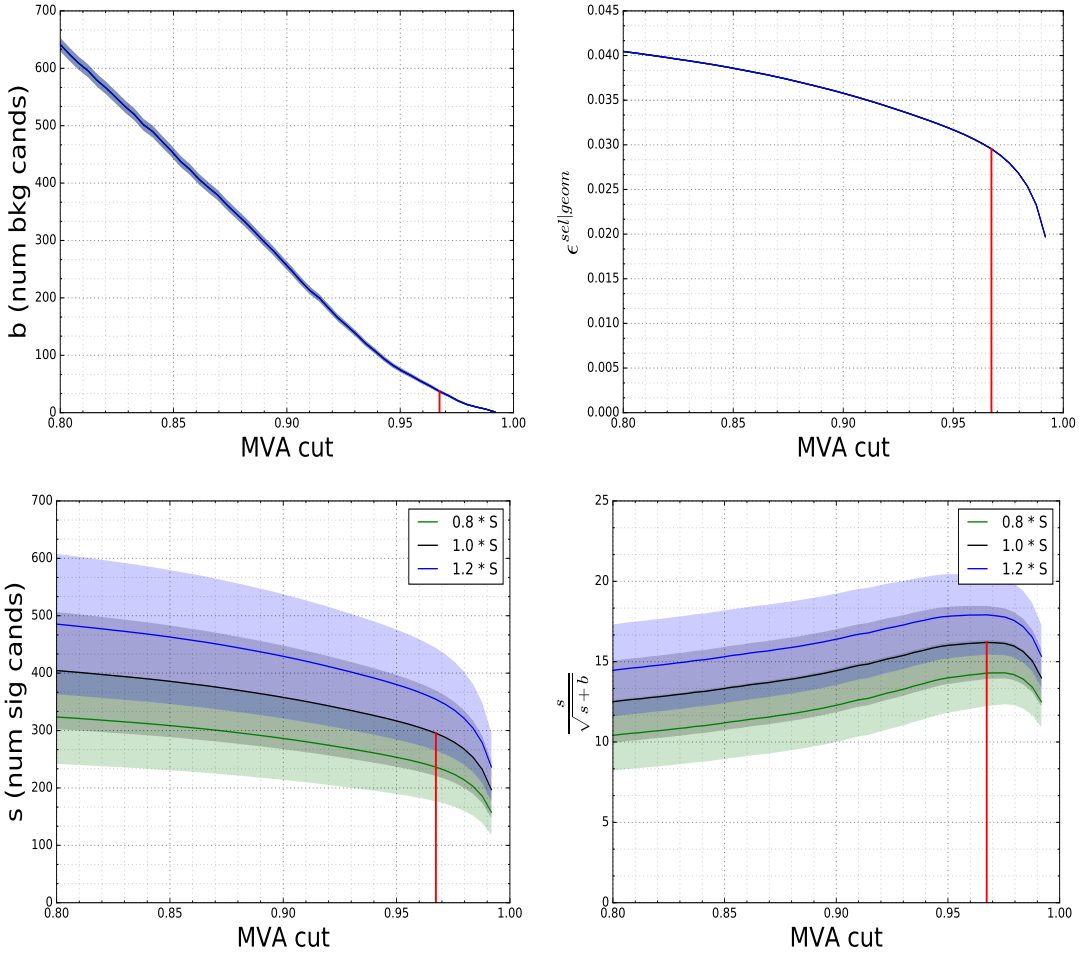

Figure 5.6: Optimisation of `xGBoost` classifier output requirement for Run II (2015–2016). (Top left) number of background candidates (B), (top right) full offline selection efficiency for signal ($\epsilon^{\text{sel|geom}}$), (bottom left) number of signal candidates (S) and (bottom right) FoM ($\frac{S}{\sqrt{S+B}}$) at different scaling factors for 'S'. The red line shows the optimal MVA output requirement.

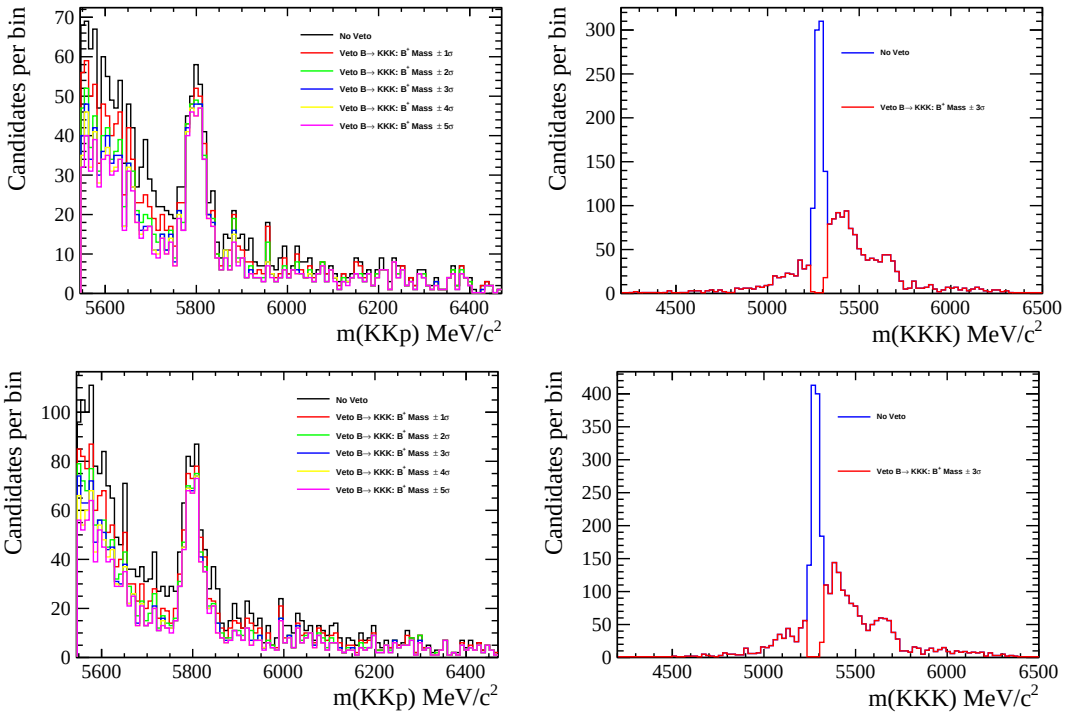

Figure 5.7: For Run I (2011–2012) data, (top left) $m(pKK)$ invariant mass, after optimal requirement on MVA output, at different veto regions defined in $m(KKK)$ spectrum, constructed when the p candidate is given K mass hypothesis, to reduce contribution from $B^+ \to KKK \to pKK$ cross-feed background. (Top right) The corresponding $m(pKK)$ distribution, with the veto region of $\pm 3\sigma$ indicated. (Bottom) Similar plots for Run II (2015–2016) data. The value of $\sigma = 15\,\mathrm{MeV}$ is taken from Ref. [135].

- the value of $\frac{\mathcal{B}(\Xi_b^-\to pK^-\pi^-)}{\mathcal{B}(\Xi_b^-\to pK^-K^-)} = 0.98 \pm 0.29$ taken from the result presented in previous chapter. Note that systematic and statistical uncertainties are added in quadrature to obtain the total uncertainty shown here;
- the values $\frac{\epsilon^{\text{sel|geom}}(\Xi_b^-\to pK^-\pi^-)}{\epsilon^{\text{sel|geom}}(\Xi_b^-\to pK^-K^-)} = 0.28 \pm 0.01$ for Run I (2011–2012) and $\frac{\epsilon^{\text{sel|geom}}(\Xi_b^-\to pK^-\pi^-)}{\epsilon^{\text{sel|geom}}(\Xi_b^-\to pK^-K^-)} = 0.14 \pm 0.01$ for Run II (2015–2016), taken from the corresponding MC samples. Better rejection of $pK\pi$ background is achieved in Run II (2015–2016) due to better PID performance.

In the fit to $m(pKK)$, I apply Gaussian constraints to both $\frac{\mathcal{B}(\Xi_b^-\to pK^-\pi^-)}{\mathcal{B}(\Xi_b^-\to pK^-K^-)}$ and $\frac{\epsilon^{\text{sel|geom}}(\Xi_b^-\to pK^-\pi^-)}{\epsilon^{\text{sel|geom}}(\Xi_b^-\to pK^-K^-)}$. Note that in the fits to $m(pKK)$ and also in the amplitude analysis, I weight the $\Xi_b^- \to pK^-\pi^-$ MC samples such that the Dalitz Plot (DP) distribution matches a cocktail DP model (see Sub-sec. 5.4.3 for more details on the model). Systematic uncertainties will be evaluated due to the model dependence.

A summary of all the offline requirements can be seen in Table 5.3.

Table 5.3: Summary of the offline requirements imposed on data sample.

Requirements
`!isMuon` on all final state tracks
$5545\,\text{MeV} < m(pKK) < 6470\,\text{MeV}$
MVA > 0.945 (Run I (2011–2012))
MVA > 0.965 (Run II (2015–2016))
$m(B^+) - 3\times 15\,\text{MeV} < m(KKK) < m(B^+) + 3\times 15\,\text{MeV}$

5.2.2 Efficiency of the signal selection

As mentioned in Sec 4.3 of the previous chapter, the detector geometry and reconstruction, in addition to small correlations between the variables used in the selection and the phase space variables, introduce some biases that favour certain parts of the phase space over others. Hence accounting for the variation of the efficiency across the phase space in any branching fraction calculation or amplitude analysis is necessary.

The efficiency can be factorised into components as

$$\epsilon^{\text{tot}} = \epsilon^{\text{geom}} \times \epsilon^{\text{sel|geom}}, \tag{5.5}$$

where

- ϵ^{geom} is the geometrical efficiency, and is determined from MC samples generated uniform in square Dalitz Plot variables. The phase space integrated ϵ^{geom} values for all Run I (2011–2012) and Run II (2015–2016) years and for both magnet polarity configurations are reported in Table 5.4. It also shows the weighted average for each run period, *i.e.* weighted according to the number of candidates in each sub-sample.

- $\epsilon^{\text{sel|geom}}$ is the selection efficiency and is calculated from the standard MC samples generated uniform in square Dalitz Plot variables. It includes contributions from reconstruction, stripping, trigger and offline (includes MVA output requirement) selections. I present, in Table 5.5, the phase space integrated selection efficiency at each of the selection stages in order to provide more precise information on where inefficiencies arise.

It was highlighted in Sec. 4.3 that the phase space of $\Xi_b^- \to pK^-K^-$ decays is characterised by five independent variables that describe the decay kinematics. In this section I present the variation of the efficiency over the two DP variables only and not the three angular variables, since in the case of unpolarised Ξ_b^- (Ω_b^-) decays the probability density should only depend on the DP variables. In the case that the Ξ_b^- (Ω_b^-) is polarised, the effect of the polarisation cancels if the efficiency is symmetric over $\cos(\theta_{Res}^{[\Xi_b^-(\Omega_b^-)]})$ where $\theta_{Res}^{[\Xi_b^-(\Omega_b^-)]}$ is the polar angle of momentum of the resonance (Res) in the rest frame of Ξ_b. I show in Figure 5.8, the efficiency as a function of $\cos(\theta_{Res}^{[\Xi_b]})$ folded at $\cos(\theta_{Res}^{[\Xi_b]}) = 0$. The relative difference in the efficiencies between $\cos(\theta_{Res}^{[\Xi_b]}) > 0$ and $\cos(\theta_{Res}^{[\Xi_b]}) < 0$ is also shown. The efficiency over $\cos(\theta_{Res}^{[\Xi_b]})$ is not fully symmetric. However, the bias introduced by the assumption of symmetric efficiency over $\cos(\theta_{Res}^{[\Xi_b]})$ is small and would be of second order since the difference in efficiency is multiplied by the Ξ_b polarisation which is expected to be small [126]. Therefore, for the nominal fit I only consider the efficiency variation over the DP variables. I discuss below the total efficiency of the signal selection procedure.

Total efficiency: The phase-space integrated total efficiencies (ϵ^{tot}), evaluated using Eq. 5.5, for signal modes are shown in Table 5.6. The spline interpolated total efficiency variation over SDP position (with associated statistical errors) for the signal modes for Run I (2011–2012) and Run II (2015–2016) are shown in Figure 5.9. I only show here efficiency variation over the phase space for the $\Xi_b^- \to pK^-K^-$ signal mode, the variation of efficiency for the $\Omega_b^- \to pK^-K^-$ mode is very similar.

Comparing the total efficiency of signal selection presented in this section for Run I (2011–2012) (Table 5.6) to that of the total efficiency presented in the previous chapter (see Table 4.12), we see that a factor 3.4 improvement in the signal efficiency is achieved.

From Figure 5.9, it can be seen that the efficiency varies across the SDP between about

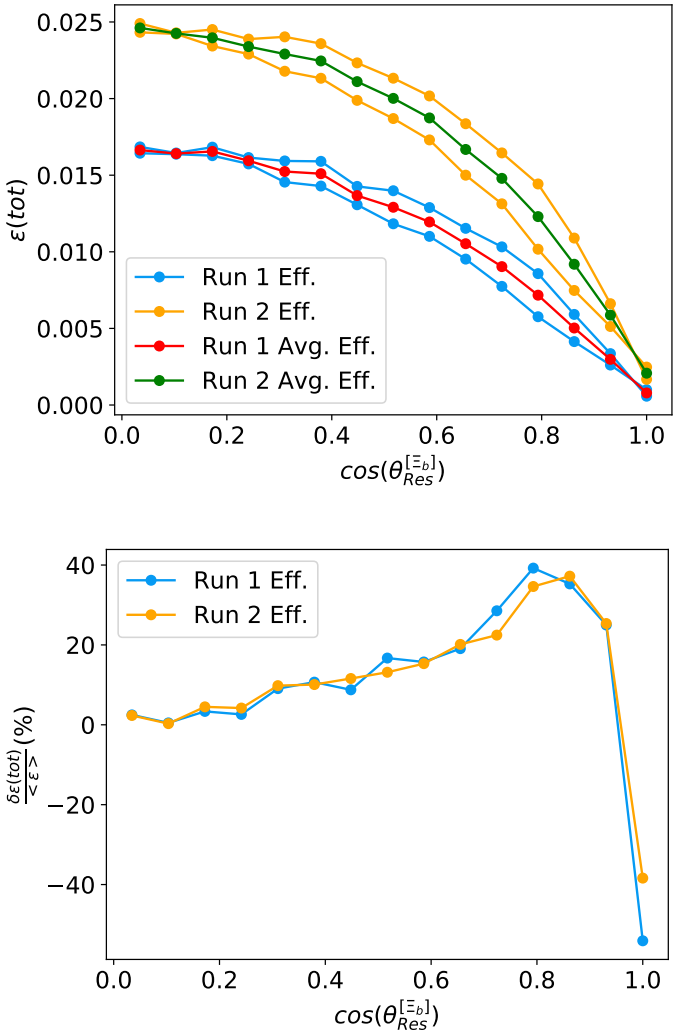

Figure 5.8: (Top) Efficiency as a function of $\cos(\theta^{[\Xi_b]}_{Res})$ folded at $\cos(\theta^{[\Xi_b]}_{Res}) = 0$. (Bottom) Relative difference in efficiency between $\cos(\theta^{[\Xi_b]}_{Res}) > 0$ and $\cos(\theta^{[\Xi_b]}_{Res}) < 0$. Note here that the large relative differences in efficiency are in the areas where the statistics of the MC sample are low and therefore these large differences are accompanied by large statistical errors.

0.08%–0.60% for Run I (2011–2012), and between about 0.2%–0.9% for Run II (2015–2016). However, it should be noted that the regions of lowest efficiency correspond to tiny parts of the conventional DP, which are magnified by the transformation. Neglecting these edges of the DP phase space, the variation of the efficiency is relatively modest. In all the efficiency maps it is evident that the b-hadron candidates with higher m' values (corresponding to higher proton or

Table 5.4: The fraction of generated b-hadron candidates that lie in the LHCb acceptance region of $0 < \theta < 400\,\mathrm{mrad}$ and whose charged daughters are all in the acceptance region of $10 < \theta < 400\,\mathrm{mrad}$ for signal (ϵ^{geom}).

Mode	ϵ^{geom} (%)									
	2011		2012		Wght. avg. Run I (2011–2012)	2015		2016		Wght. avg. Run II (2015–2016)
	MagDown	MagUp	MagDown	MagUp		MagDown	MagUp	MagDown	MagUp	
$\Xi_b^- \to pK^-K^-$	53.32±0.06	53.32±0.06	53.79±0.06	53.79±0.06	53.67 ± 0.04	55.28±0.06	55.28±0.06	55.31±0.06	55.31± 0.06	55.31 ± 0.04
$\Omega_b^- \to pK^-K^-$	52.68±0.06	52.68±0.06	53.13±0.06	53.13±0.06	53.02 ± 0.03	54.75±0.06	54.75±0.06	54.71±0.06	54.71±0.06	54.72 ± 0.04

Table 5.5: Phase-space integrated offline selection efficiencies for signal ($\Xi_b^-(\Omega_b^-) \to pK^-K^-$) modes. Statistical uncertainties are also given. Note 'R' here refers to Reconstruction stage, 'S' refers to stage after stripping requirements 'Trig' here refers to trigger requirements and 'sel' refers to full offline selection (*i.e.* R+S+Trig+Veto+MVA). Truth matched MC samples have been used to evaluate these efficiencies.

Mode	$\epsilon^{\mathrm{R+S\|geom}}$ (%)									
	2011		2012		Wght. avg. Run I (2011–2012)	2015		2016		Wght. avg. Run II (2015–2016)
	MagDown	MagUp	MagDown	MagUp		MagDown	MagUp	MagDown	MagUp	
$\Xi_b^- \to pK^-K^-$	13.47 ± 0.05	13.52 ± 0.05	12.84 ± 0.03	12.78 ± 0.03	12.98 ± 0.02	12.87 ± 0.03	12.95 ± 0.03	13.79 ± 0.02	13.87 ± 0.02	13.76 ± 0.02
$\Omega_b^- \to pK^-K^-$	14.16 ± 0.05	14.12 ± 0.05	13.41 ± 0.03	13.45 ± 0.03	13.6 ± 0.02	13.65 ± 0.03	13.8 ± 0.03	14.64 ± 0.02	14.61 ± 0.02	14.55 ± 0.02
	$\frac{\epsilon^{\mathrm{R+S+Trig\|geom}}}{\epsilon^{\mathrm{R+S\|geom}}}$ (%)									
$\Xi_b^- \to pK^-K^-$	39.23 ± 0.27	39.36 ± 0.27	38.34 ± 0.19	38.44 ± 0.19	38.61 ± 0.11	47.22 ± 0.22	46.9 ± 0.22	45.79 ± 0.15	45.81 ± 0.15	45.9 ± 0.1
$\Omega_b^- \to pK^-K^-$	40.19 ± 0.27	40.07 ± 0.27	39.22 ± 0.19	39.37 ± 0.19	39.49 ± 0.11	47.92 ± 0.22	48.19 ± 0.22	46.94 ± 0.15	46.9 ± 0.15	47.02 ± 0.1
	$\frac{\epsilon^{\mathrm{sel\|geom}}}{\epsilon^{\mathrm{R+S+Trig\|geom}}}$ (%)									
$\Xi_b^- \to pK^-K^-$	43.45 ± 0.47	43.45 ± 0.47	42.39 ± 0.33	43.59 ± 0.34	43.1 ± 0.2	51.2 ± 0.35	52.02 ± 0.35	50.15 ± 0.24	49.77 ± 0.24	50.09 ± 0.16
$\Omega_b^- \to pK^-K^-$	44.65 ± 0.47	45.09 ± 0.47	43.04 ± 0.34	44.7 ± 0.34	44.12 ± 0.2	52.48 ± 0.34	53.35 ± 0.34	51.26 ± 0.23	51.2 ± 0.23	51.37 ± 0.15
	$\epsilon^{\mathrm{sel\|geom}}$ (%)									
$\Xi_b^- \to pK^-K^-$	2.3 ± 0.02	2.31 ± 0.02	2.09 ± 0.01	2.14 ± 0.01	2.16 ± 0.01	3.11 ± 0.02	3.16 ± 0.02	3.17 ± 0.01	3.16 ± 0.01	3.16 ± 0.01
$\Omega_b^- \to pK^-K^-$	2.54 ± 0.02	2.55 ± 0.02	2.26 ± 0.01	2.37 ± 0.02	2.37 ± 0.01	3.43 ± 0.02	3.55 ± 0.02	3.52 ± 0.01	3.51 ± 0.01	3.51 ± 0.01

kaon momentum in the rest frame of the b-hadron) have a larger efficiency, while those with low values of m' (corresponding to lower proton or kaon momentum) have lower efficiency.

Table 5.6: Phase-space integrated total efficiencies ($\epsilon^{\rm tot}$) for signal modes, with statistical uncertainties.

Modes	Run I (2011–2012)	Run II (2015–2016)
	$\epsilon^{\rm tot}$ (%)	
$\Xi_b^- \to pK^-K^-$	1.159 ± 0.005	1.748 ± 0.006
$\Omega_b^- \to pK^-K^-$	1.257 ± 0.005	1.921 ± 0.006

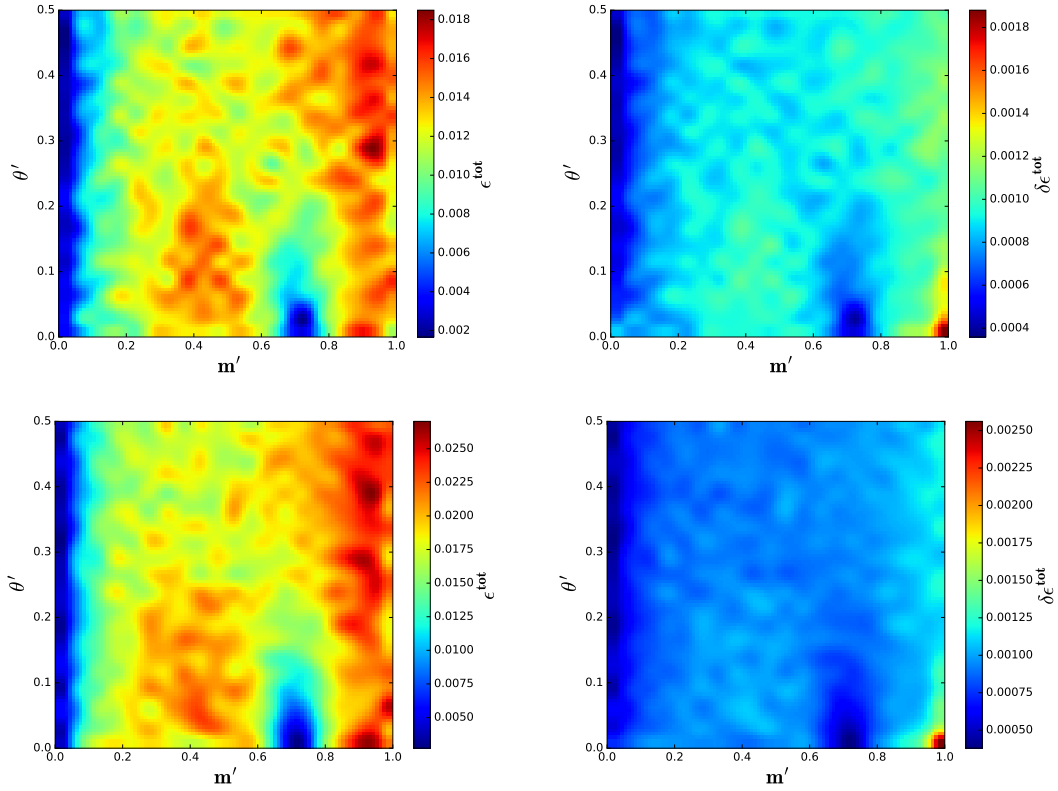

Figure 5.9: For Run I (2011–2012), (Top, left to right) spline interpolation of total efficiency ($\epsilon^{\rm tot}$) map as function of SDP position and associated statistical error ($\delta\epsilon^{\rm tot}$). (Bottom) Same plots for Run II (2015–2016) samples. Similar variation of efficiency is observed for $\Omega_b^- \to pK^-K^-$ mode.

To establish the baseline amplitude model, I use the efficiency maps averaged over initial

state charge (see Figure 5.9) since there are no significant differences between the efficiency maps split by charge of the initial state (see Figure 5.10). However when conducting fits to data sample split by initial state charge to investigate CPV effects, I use separate efficiency maps.

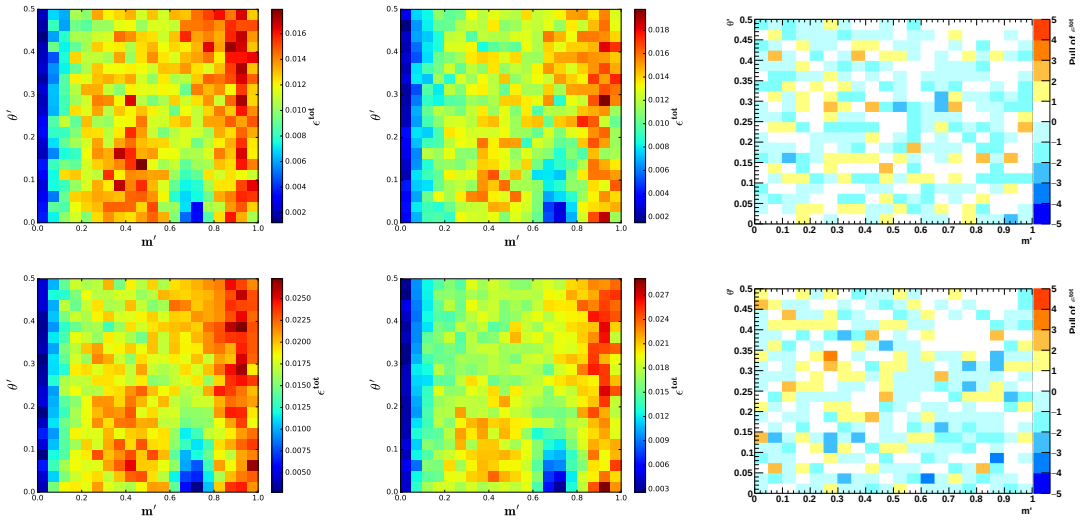

Figure 5.10: (Top) Run I (2011–2012): Total efficiency (ϵ^{tot}) map as function of SDP position for (left) Ξ_b^-, (middle) Ξ_b^+ and (right) pull distribution between them. (Bottom) Same for Run II (2015–2016). Although no significant differences is visible from the pull distributions, I use these efficiency maps when conducting fits to the individual Ξ_b^- and Ξ_b^+ samples.

5.3 Determination of signal and background fit fractions

The amplitude analysis is conducted using a data sample with high signal purity obtained requiring the candidates to lie in the signal region of the $m(pKK)$ spectra. This signal region is chosen to be $m(\Xi_b^-) \pm 2.5\sigma$ where the value of $\sigma = 17\,\text{MeV}$ is taken from the fit to the $\Xi_b^- \to pK^-K^-$ MC sample and $m(\Xi_b^-) = 5797.72\,\text{MeV}$ is taken from Ref. [113]. The signal and background fractions determined in this signal region form an input to the PDF to fit the DP distribution.

In this section I present the results of the unbinned extended maximum likelihood fit to the pKK invariant mass spectrum and determine the signal and background fractions in the defined signal region. The total PDF as a function of $m(pKK)$ is composed of signal and background components whose shape parameters have been taken from fits to the corresponding MC samples with the same selection requirements imposed on them as the data sample. The fit is similar to that used in the previous chapter, except that only the $m(pKK)$ spectrum is studied, in contrast

to the simultaneous fit to pKK, $pK\pi$ and $p\pi\pi$ spectra. There are also differences with regard to the cross-feed background components. The $\Xi_b^- \to pK^-\pi^-$ MC sample has been generated flat in its square DP and its phase-space distribution is not known a priori. To obtain the shape parameters, I assume a DP model which is composed of $\Lambda(1405)$, $\Lambda(1520)$, $\Lambda(1690)$, $N(1440)$, $N(1520)$, $N(1535)$ and $N(1650)$ resonances and weight my MC sample according to this model. A systematic uncertainty will be assigned by varying the DP model. The yield of this component is Gaussian constrained to the expected value(see Sec. 5.2.1.3 for the exact implementation.). Negligible cross-feed from the $\Omega_b^- \to pK^-\pi^-$ mode is expected as it is highly suppressed. The $B^+ \to K^+K^-K^+$ cross-feed contribution has been vetoed from the data sample and any residual contribution from this can be accommodated in the combinatorial background.

The fit results to the data sample for Run I (2011–2012) and Run II (2015–2016) are shown in Table 5.7 and Figure 5.11. Table 5.8 shows the signal and background contributions in the signal region $m(\Xi_b^-) \pm 2.5\sigma$. These are the yields that will be used in the amplitude analysis.

Table 5.7: Results of the fit to the $m(pKK)$ spectrum.

Parameter	Value
Run I (2011–2012) sample	
$N_{\rm sig}(\Xi_b^- \to pK^-K^-)$	193 ± 21
$N_{\rm sig}(\Omega_b^- \to pK^-K^-)$	$-\,4 \pm 6$
$N_{\rm crsfd}(\Xi_b^- \to pK^-\pi^- \to pKK)$	$(0.15 \pm 0.06) \times N_{\rm sig}(\Xi_b^- \to pK^-K^-)$
$N_{\rm PRB}(\Xi_b^- \to K^{*-}(K^-\pi^0)K^-p \to pKK)$	231 ± 34
$N_{\rm comb}$	721 ± 50
$c_{\rm comb}$	$(-26.9 \pm 2.5) \times 10^{-4}\,{\rm GeV}^{-1}$
$\sigma_{\rm scale}$	0.97 ± 0.11
Run II (2015–2016) sample	
$N_{\rm sig}(\Xi_b^- \to pK^-K^-)$	297 ± 23
$N_{\rm sig}(\Omega_b^- \to pK^-K^-)$	15 ± 9
$N_{\rm crsfd}(\Xi_b^- \to pK^-\pi^- \to pKK)$	$(0.14 \pm 0.03) \times N_{\rm sig}(\Xi_b^- \to pK^-K^-)$
$N_{\rm PRB}(\Xi_b^- \to K^{*-}(K^-\pi^0)K^-p \to pKK)$	442 ± 36
$N_{\rm comb}$	775 ± 51
$c_{\rm comb}$	$(-21.9 \pm 2.3) \times 10^{-4}\,{\rm GeV}^{-1}$
$\sigma_{\rm scale}$	0.99 ± 0.08

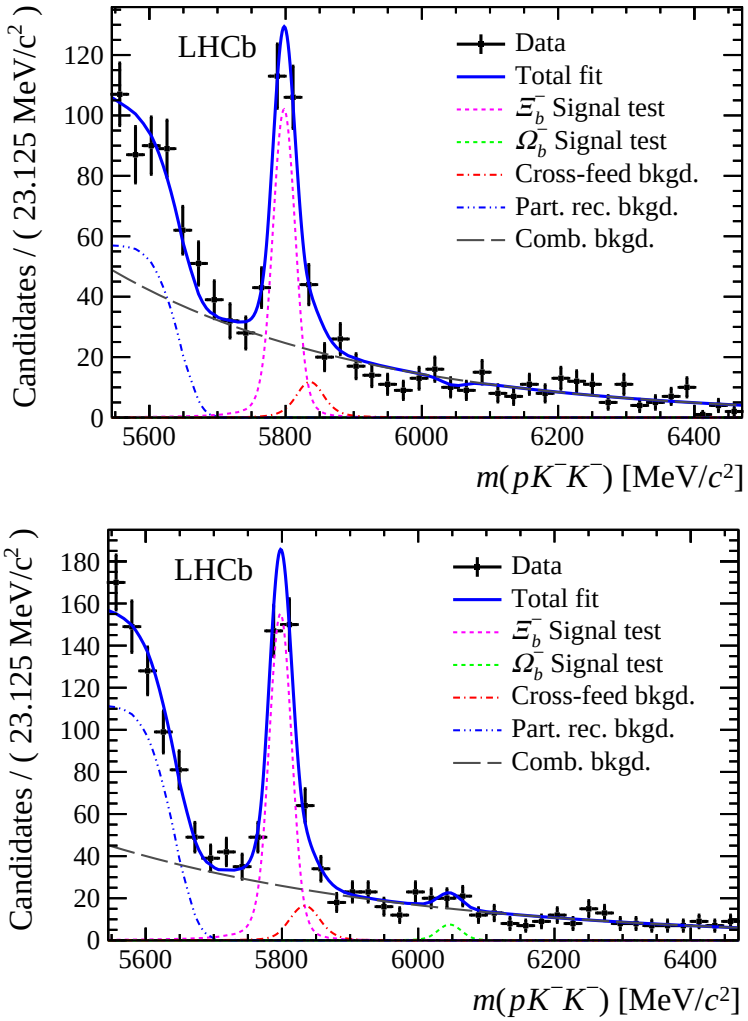

Figure 5.11: The results of the unbinned extended maximum likelihood fit to the $m(pKK)$ invariant mass spectrum for (top) Run I (2011–2012) and (bottom) Run II (2015–2016) data samples.

5.3.1 Update of the ratio of production and branching fraction

With the analysis presented so far, the product of the ratios of fragmentation fractions and branching fractions can be updated:

$$\mathcal{R} = \frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \times \frac{\mathcal{B}(\Omega_b^- \to pK^-K^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} = \frac{N(\Omega_b^- \to pK^-K^-)}{\epsilon^{\rm tot}(\Omega_b^- \to pK^-K^-) \times N_{\rm corr}(\Xi_b^- \to pK^-K^-)} \,,$$

Table 5.8: $\Xi_b^- \to pK^-K^-$ signal and background contributions in the signal region $m(\Xi_b^-) \pm 2.5\sigma$ for Run I (2011–2012) and Run II (2015–2016). The value of $\sigma = 17\,\mathrm{MeV}$ is taken from the fit to the $\Xi_b^- \to pK^-K^-$ MC sample and $m(\Xi_b^-) = 5797.72\,\mathrm{MeV}$ is taken from Ref. [113].

Variables	Run I: $m(\Xi_b^-) \pm 2.5\sigma$	Run II: $m(\Xi_b^-) \pm 2.5\sigma$
S	181 ± 20	278 ± 21
N_{comb}	90 ± 6	95 ± 6
N_{crsfd}	16 ± 7	25 ± 6
$B = N_{\mathrm{comb}} + N_{\mathrm{crsfd}}$	107 ± 9	120 ± 9
$\frac{S}{B}$	1.7 ± 0.2	2.3 ± 0.2
$\frac{S}{S+B}$	$63\% \pm 3\%$	$70\% \pm 2\%$

where $N(\Omega_b^- \to pK^-K^-)$ and $\epsilon(\Omega_b^- \to pK^-K^-)$ are the yield and efficiency of the $\Omega_b^- \to pK^-K^-$ mode respectively. The quantity $N_{\mathrm{corr}}(\Xi_b^- \to pK^-K^-) = \sum_i \frac{w_i}{\epsilon_i^{\mathrm{tot}}(\Xi_b^- \to pK^-K^-)}$ is the efficiency-corrected yield which is obtained from the sum over all candidates in the mass fit of the ratio of the signal sWeight and the total selection efficiency for that candidate calculated from the efficiency maps depending on the candidates phase space location. In the absence of a significant signal for the $\Omega_b^- \to pK^-K^-$ decay, it is not possible to calculate a similar corrected yield for it, and therefore a global efficiency, rather than event-by-event (DP position dependent) efficiencies, is used instead. Note that in contrast to the Run I (2011–2012) analysis [135], this result was not obtained from a blind analysis. However the selection algorithms were designed to be optimised for the $\Xi_b^- \to pK^-K^-$ amplitude analysis, not for the $\Omega_b^- \to pK^-K^-$ search, and therefore no large bias in the result is expected due to this effect.

The results for Run I (2011–2012) and Run II (2015–2016), with statistical uncertainties, are

$$\text{Run I (2011–2012): } \frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \times \frac{\mathcal{B}(\Omega_b^- \to pK^-K^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} = (-20 \pm 30\,(\mathrm{stat})) \times 10^{-3}\,,$$

$$\text{Run II (2015–2016): } \frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \times \frac{\mathcal{B}(\Omega_b^- \to pK^-K^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} = (\ \ 51 \pm 29\,(\mathrm{stat})) \times 10^{-3}\,.$$

The statistical error on the corrected yield is evaluated as outlined in Ref. [152]. Sources of systematic uncertainty are from the fit model, the uncertainty in the efficiency, the variation of the efficiency for the unobserved $\Omega_b^- \to pK^-K^-$ channel across the phase space; possible mis-modelling of the Ω_b^- and Ξ_b^- production in simulation (in particular their p_{T} and decay time distributions). These will all be evaluated in the same way as in the previous analysis [135].

The results from Run I (2011–2012) and Run II (2015–2016) are consistent, and no significant $\Omega_b^- \to pK^-K^-$ signal is seen in either. Therefore the results are combined, and an upper limit set on the ratio of production fractions and branching fractions. The profile likelihood scans as a function of $\mathcal{R}$ are shown in Figure 5.12. The combined results are

$$\text{Combined: } \frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \times \frac{\mathcal{B}(\Omega_b^- \to pK^-K^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} \quad = \quad (\ 24 \pm 21\,(\text{stat})) \times 10^{-3}\,.$$

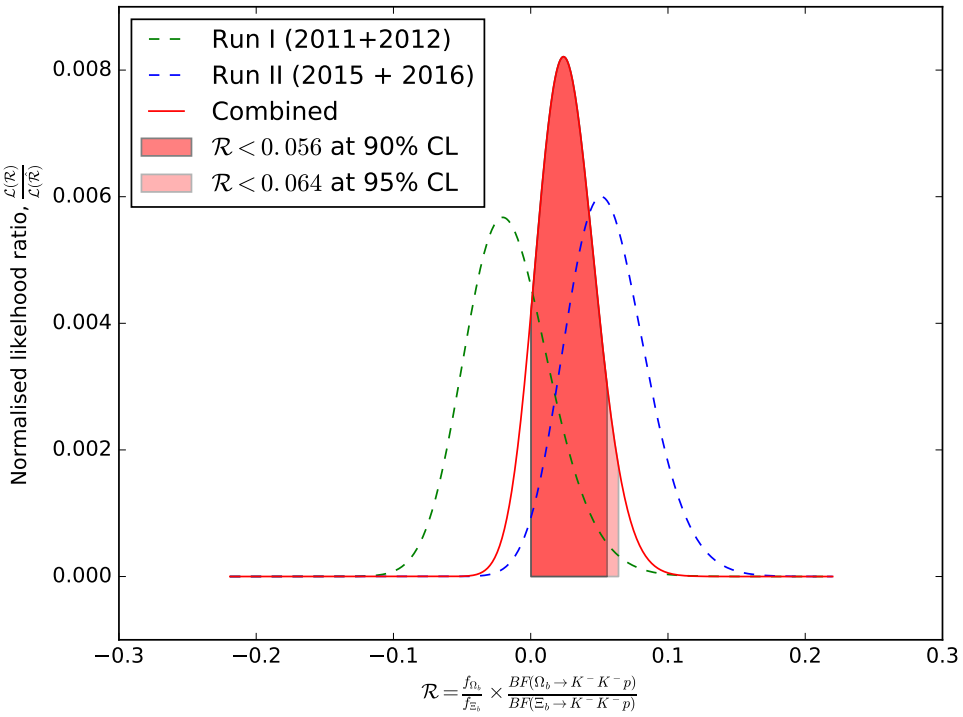

Figure 5.12: Run I (2011–2012), Run II (2015–2016) and combined normalised profile likelihood scans as a function of $\mathcal{R}$.

From the combined value, I calculate the upper limit on $\mathcal{R}$ at 90 (95) % confidence level, $\mathcal{R}_{90\,(95)}$, by integrating the likelihood, $\mathcal{L}(\mathcal{R})$, in the physical region of non-negative branching fraction, such that

$$\int_0^{\mathcal{R}_{90\,(95)}} \mathcal{L}(\mathcal{R})d\mathcal{R} = 0.90\,(0.95) \times \int_0^{\infty} \mathcal{L}(\mathcal{R})d\mathcal{R}\,.$$

Systematic uncertainties are included by convoluting the likelihood with a Gaussian of width corresponding to the systematic uncertainty on $\mathcal{R}$.

The obtained upper limits at 90 (95) % confidence level, combining Run I (2011–2012) and Run II (2015–2016) data (including only statistical uncertainties, systematic uncertainties need to be evaluated), are

$$\mathcal{R}_{90\,(95)} = 56\ (64) \times 10^{-3}\,.$$

In the previous analysis of charmless three-body $\Xi_b^-(\Omega_b^-)$ decays [135], results were not obtained in terms of the quantity $\mathcal{R}$ reported here, but rather

$$\begin{aligned} R_{\Xi_b^-\to pK^-K^-} &\equiv \frac{f_{\Xi_b^-}}{f_u}\frac{\mathcal{B}(\Xi_b^-\to pK^-K^-)}{\mathcal{B}(B^-\to K^+K^-K^-)} = (\ \ 265 \pm 35 \pm 47) \times 10^{-5}\,, \\ R_{\Omega_b^-\to pK^-K^-} &\equiv \frac{f_{\Omega_b^-}}{f_u}\frac{\mathcal{B}(\Omega_b^-\to pK^-K^-)}{\mathcal{B}(B^-\to K^+K^-K^-)} = (\ -9 \pm\ \ 9 \pm\ \ 6) \times 10^{-5} < 18\ (22) \times 10^{-5}\,. \end{aligned}$$

Thus, ignoring possible correlations of the uncertainties, the previous analysis gives

$$\mathcal{R} = \frac{R_{\Omega_b^-\to pK^-K^-}}{R_{\Xi_b^-\to pK^-K^-}} = (-34 \pm 34 \pm 23) \times 10^{-3}\,.$$

The updated result can be seen to be consistent with, and more precise than, the previous measurements (note here that the Run I (2011–2012) samples are correlated so comparison between results is not straight forward).

5.4 Amplitude analysis

Our goal in this section is to develop a fit model to fit the phase space distributions of $\Xi_b^- \to pK^-K^-$ decays, taking into account the variation of detection efficiency and background distribution over the phase space, using both Run I (2011–2012) and Run II (2015–2016) data simultaneously. Using this nominal model the phase space distribution of Ξ_b^- and Ξ_b^+ data samples is fitted simultaneously to extract the CP violation parameters. When conducting fits to establish the fit model, the total probability density function (PDF) of the phase space is defined as:

$$\mathcal{P}_{\rm tot}(\vec{\theta},\Omega) = \frac{1}{N_{\rm tot}}\left[N_{\rm sig}\mathcal{P}_{\rm sig}(\vec{\theta},\Omega) + N_{\rm comb}\mathcal{P}_{\rm comb}(\Omega) + N_{\rm crsfd}\mathcal{P}_{\rm crsfd}(\Omega)\right] \tag{5.6}$$

where subscripts 'sig', 'comb' and 'crsfd' respectively denote the signal, combinatorial and cross-feed components. The symbol N denotes the expected yield of the components in the signal region $m(\Xi_b^-) \pm 2.5\sigma$. These yields are fixed according to Table 5.8. Here $N_{\rm tot}$ is the

total expected yield and is given by $N_{\rm tot} = N_{\rm sig} + N_{\rm comb} + N_{\rm crsfd}$. The symbol $\mathcal{P}$ represents the PDF of each component, which are discussed in detail in this section, where $\vec{\theta}$ and Ω denote the parameters of the model and the phase space of $\Xi_b^- \to pK^-K^-$ decay, respectively.

The negative log likelihood that needs to be minimised after adding a Poisson term to Eq. 5.6 (as done previously in Eq. 5.7), to extract best estimates of the parameters of the model ($\hat{\vec{\theta}}$) is given by:

$$-\ln(\mathcal{L}(\vec{\theta},\Omega)) = N_{\rm tot} - n\ln(N_{\rm tot}) + \ln(n!) - \sum_{i}^{n}\ln(\mathcal{P}^i_{\rm tot}(\vec{\theta},\Omega)) \tag{5.7}$$

where n is the observed number of data candidates which is a constant and since an extended likelihood fit is not being conducted, $N_{\rm tot}$ is also a constant. These constants do not affect the minimisation procedure and can be ignored.

When conducting the fits to extract the CP violation parameters the total PDF is given by:

$$\begin{aligned}\mathcal{P}^q_{\rm tot}(\vec{\theta},\Omega) = \frac{1}{N_{\rm tot}}\Big[N_{\rm sig}\mathcal{P}^q_{\rm sig}(\vec{\theta},\Omega) + N_{\rm comb}\frac{(1-qA_{\rm comb})}{2}\mathcal{P}^q_{\rm comb}(\Omega) \\ + N_{\rm crsfd}\frac{(1-qA_{\rm crsfd})}{2}\mathcal{P}^q_{\rm crsfd}(\Omega)\Big]\end{aligned} \tag{5.8}$$

where $q = +1$ for Ξ_b^- sample and $q = -1$ for Ξ_b^+ sample. The likelihood for the above PDF is constructed in a similar manner to that in Eq. 5.7.

5.4.1 Signal PDF

The signal PDF mentioned in Eq. 5.6, to establish the baseline model is given by

$$\mathcal{P}_{\rm sig}(\Omega) = \frac{\epsilon(\Omega)\frac{d\Gamma}{d\Omega}}{\int_\Omega \epsilon(\Omega)\frac{d\Gamma}{d\Omega}d\Omega} \tag{5.9}$$

where $\epsilon(\Omega)$ is the detection efficiency (see in Section 5.2) and $\frac{d\Gamma}{d\Omega}$ is the differential decay density of $\Xi_b^- \to pK^-K^-$. Here $\epsilon(\Omega)$ is the average of the efficiency maps obtained for Ξ_b^- and Ξ_b^+ decays.

The signal PDF, mentioned in Eq. 5.8, to extract the CP violation parameters is given by

$$\mathcal{P}^q_{\rm sig}(\Omega) = \frac{\epsilon^q(\Omega)\frac{d\Gamma^q}{d\Omega}}{\int_\Omega\left(\epsilon^+(\Omega)\frac{d\Gamma^+}{d\Omega} + \epsilon^-(\Omega)\frac{d\Gamma^-}{d\Omega}\right)d\Omega} \tag{5.10}$$

where ϵ^+ and ϵ^- describe the selection efficiency variation over phase space for Ξ_b^- and Ξ_b^+ samples, respectively. Similarly $\frac{d\Gamma^+}{d\Omega}$ and $\frac{d\Gamma^-}{d\Omega}$ describe the differential decay density for Ξ_b^- and

Ξ_b^+ decays to the pKK final state respectively. The normalisation factor in Eq. 5.10 is chosen such that the signal PDF describing the decay dynamics of a given charged initial state is sensitive to the global asymmetry that exists between Ξ_b^- and Ξ_b^+ samples. In other words, the parameters of the amplitude model determine the asymmetry in the signal yield.

The $\Xi_b^- \to pK^-K^-$ decay is a three-body decay. One could in principle have quasi-two body contributions that appear in invariant mass distributions of $m(pK^-)_{\rm low}$, $m(pK^-)_{\rm high}$ and $m(K^-K^-)$ where $m(pK^-)_{\rm low} = \sqrt{\rm Min}(m(K_1^-p), m(K_2^-p))$ and $m(pK^-)_{\rm high} = \sqrt{\rm Max}(m(K_1^-p), m(K_2^-p))$. I show the square DP and conventional DP distributions for Run I (2011–2012) and Run II (2015–2016) data in the signal region in Figure 5.13. It can be seen here that there are no visible resonant contributions in the invariant mass spectrum of $m(pK^-)_{\rm high}$ and no exotic contributions in $m(K^-K^-)$. Therefore only resonance components, denoted by ${\rm Res}$, appearing in the $m(pK^-)_{\rm low}$ spectrum are initially considered. Since Ξ_b^- is an isospin $1/2$ particle, and since the resonance recoils against a kaon (isospin $1/2$), contributions through the strong decay of both Λ^* and Σ^* resonances are expected. Other resonances that would decay weakly to the pK^- final state are highly suppressed and are not considered.

The phase space of the decay of a spin-half particle to another spin-half particle and two pseudo-scalars is characterised by five variables. I define these 5 variables to be $\Omega = \left\{m(pK^-)^2_{\rm low}, \theta^{[\Xi_b]}_{\rm Res}, \phi^{[\Xi_b]}_{\rm Res}, \theta^{[\rm Res]}_p, \phi^{[\rm Res]}_p\right\}$. Here, $(\theta^{[\Xi_b]}_{\rm Res}, \phi^{[\Xi_b]}_{\rm Res})$ are the polar and azimuthal angles of the momentum vector of the resonance, defined in the frame of Ξ_b which is the helicity frame (z-axis is along the Ξ_b momentum). Similar angles $(\theta^{[\rm Res]}_p, \phi^{[\rm Res]}_p)$ for the proton are defined. A schematic diagram depicting these various angles is shown in Figure 5.14. The superscripts $[\Xi_b]$ and $[{\rm Res}]$ appearing in the angles denote quantities defined in the Ξ_b and ${\rm Res}$ rest frames, respectively. Any other variables (*e.g.* the angles defined in Sec. 4.3) related to the phase space of the decay can be written as a function of these 5 variables. For example the variable $m(pK^-)^2_{\rm high}$ is related to $\theta^{[\rm Res]}_p$ by the following relation

$$\cos(\theta^{[\rm Res]}_p) = \frac{(m(pK^-)^2_{\rm high})_{\rm max} + (m(pK^-)^2_{\rm high})_{\rm min} - 2m(pK^-)^2_{\rm high}}{(m(pK^-)^2_{\rm high})_{\rm max} - (m(pK^-)^2_{\rm high})_{\rm min}} \tag{5.11}$$

where ${\rm max}$ and ${\rm min}$ denote the upper and lower kinematic limits of $m(pK^-)^2_{\rm high}$ respectively.

To describe the differential decay density I use the helicity formalism [153] to parametrise the decay dynamics and the isobar approximation to coherently sum resonant and nonresonant contributions. The amplitude of $\Xi_b^- \to {\rm Res}(\to K^-p)K^-$ decay for specific values of

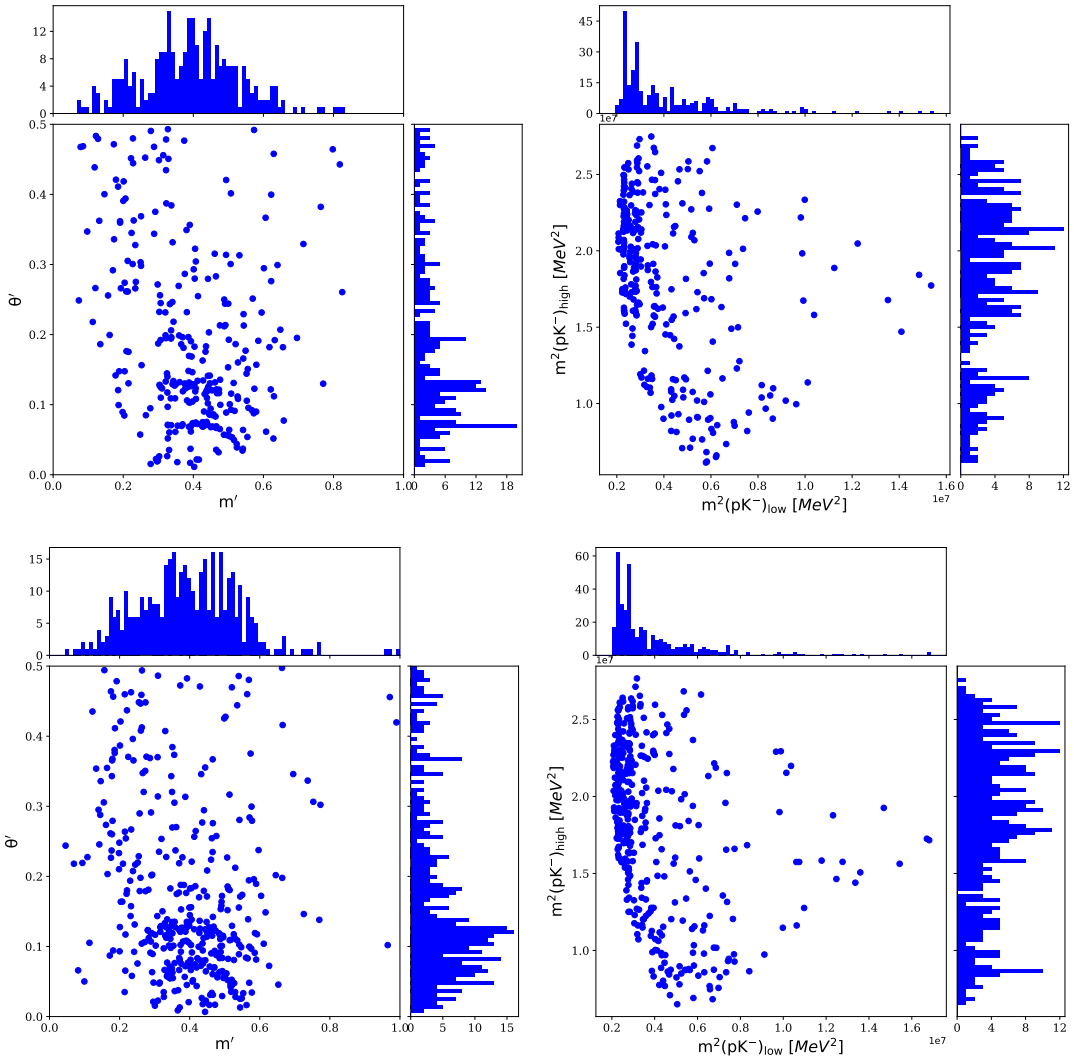

Figure 5.13: (Top) For Run I (2011–2012) and (bottom) Run II (2015–2016) data, in the signal region, the distribution of the candidates as a function of (left) square DP and (right) conventional DP. The X, Y projections of these plots are also shown.

$M_{\Xi_b}, \lambda_{\rm Res}, \lambda_p$ is given by

$$A_{M_{\Xi_b},\lambda_{\rm Res},\lambda_p}(\Omega) = D^{J_{\Xi_b}*}_{M_{\Xi_b},\lambda_{\rm Res}-\lambda_K}(\phi^{[\Xi_b]}_{\rm Res},\theta^{[\Xi_b]}_{\rm Res},0) D^{J_{\rm Res}*}_{M_{\rm Res},\lambda_p-\lambda_K}(\phi^{[\rm Res]}_p,\theta^{[\rm Res]}_p,0) \times a^{J_{\Xi_b}}_{\lambda_{\rm Res},\lambda_K} b^{J_{\rm Res}}_{\lambda_p,\lambda_K} R(m(pK^-)_{\rm low}) \quad (5.12)$$

where J, M, λ are eigenvalues of the spin, z-component of the spin and helicity operators for a given particle, which is denoted in the subscript. The terms appearing in the equation are

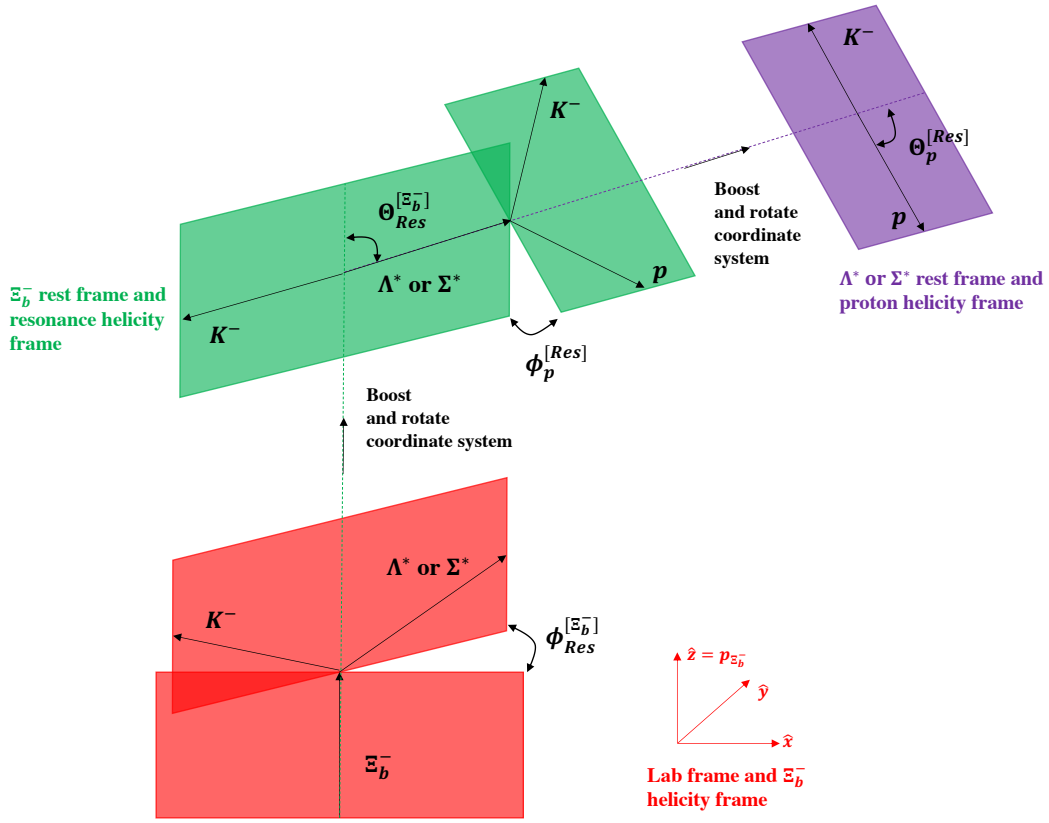

Figure 5.14: The schematic diagram showing the various angles, $\theta_{\rm Res}^{[\Xi_b]}$, $\phi_{\rm Res}^{[\Xi_b]}$, $\theta_p^{[\rm Res]}$, $\phi_p^{[\rm Res]}$, which along with $m(pK^-)_{\rm low}$ form the five phase space variables of $\Xi_b^- \to pK^-K^-$ decay.

discussed in the following sub-sections.

Note that Eq. 5.12 must be symmetrised on the exchange of the four momentum of the two final state kaons, since they are identical bosons. Therefore my symmetrised amplitude becomes[2]

$$A^{M_{\Xi_b},\lambda_{\rm Res},\lambda_p}(\Omega,\Omega') = A_{M_{\Xi_b},\lambda_{\rm Res},\lambda_p}(\Omega) + \sum_{\lambda_p'=\pm 1/2} d^{1/2}_{\lambda_p',\lambda_p}(\theta_{\rm rot}) \times A_{M_{\Xi_b},\lambda_{\rm Res},\lambda_p'}(\Omega')\,. \quad (5.13)$$

Here Ω consists of the 5 previously mentioned phase space variables for the case when the resonance state decays to a proton and a kaon. Ω' consists of the same 5 phase space variables but for the case when the resonance state decays to a proton and the other kaon. I have chosen to express my symmetric amplitude in terms of the proton helicity state that is used in the definition

[2]Thanks to Mikhail Mikhasenko for verifying this equation.

of the amplitude that appears in the first term of Eq. 5.13, *i.e.* $|\lambda_p\rangle$. Since the amplitude in the second term of Eq. 5.13 is expressed in terms of a proton helicity state defined in a different reference frame ($|\lambda'_p\rangle$), it needs to be rotated as under

$$|\lambda'_p\rangle = \sum_{\lambda'_p=\pm 1/2} d^{1/2}_{\lambda'_p,\lambda_p}(\theta_{\rm rot})|\lambda_p\rangle \tag{5.14}$$

where $\theta_{\rm rot}$ is the angle between the quantisation axis of protons defined in $m(pK^-)_{\rm high}$ and $m(pK^-)_{\rm low}$ reference frames and d is the small Wigner d-matrix. The angle $\theta_{\rm rot}$ is visualised in Figure 18 of Ref. [21]. Therefore the second term of Eq. 5.13 acquires this additional term.

5.4.1.1 Wigner D-Matrix elements

In Eq. 5.12, the complete angular dependence of the decay is captured by Wigner D-Matrix elements ($D^{J*}_{M,M'}$) [153, 154] which are given as

$$D^{J*}_{M,M'}(\alpha,\beta,\gamma) = e^{iM\alpha} d^J_{M,M'}(\beta) e^{iM'\gamma} \tag{5.15}$$

where α, β, γ are the standard Euler angles, and $d^J_{M,M'}(\beta)$ are the small Wigner d-matrix elements [153, 154]. The Wigner D-Matrix elements in Eq. 5.12 also impose angular momentum conservation giving rise to two conditions $J_{\Xi_b} \geq |\lambda_{\rm Res} - \lambda_K|$ and $J_{\rm Res} \geq |\lambda_p - \lambda_K|$. Using the values $J_{\Xi_b} = 1/2$ and $\lambda_K = 0$, these conditions imply that only $\lambda_{\rm Res} = \pm 1/2$ helicity states contribute to the amplitude for resonances with any half-integer spin values. The definitions for small Wigner d-matrix elements for $(M, M') = (1/2, 1/2), (-1/2, 1/2)$ are given in Table 5.9 with the property $d^J_{M,M'} = (-1)^{M-M'} d^J_{M',M} = d^J_{-M,-M'}$.

Table 5.9: Definition of small Wigner d-matrix elements ($d^J_{M,M'}(\beta)$) for $(M, M') = (1/2, 1/2), (-1/2, 1/2)$ which exhibit the property $d^J_{M,M'} = (-1)^{M-M'} d^J_{M',M} = d^J_{-M,-M'}$.

Spin (J)	$d^J_{1/2,1/2}(\beta)$	$d^J_{1/2,-1/2}(\beta)$
1/2	$\cos(\frac{\beta}{2})$	$-\sin(\frac{\beta}{2})$
3/2	$\frac{1}{2}\cos(\frac{\beta}{2})(3\cos(\beta)-1)$	$-\frac{1}{2}\sin(\frac{\beta}{2})(3\cos(\beta)+1)$
5/2	$\frac{1}{2}\cos(\frac{\beta}{2})(5\cos(\beta)^2+2\cos(\beta)-1)$	$-\frac{1}{2}\sin(\frac{\beta}{2})(5\cos(\beta)^2+2\cos(\beta)-1)$

5.4.1.2 Helicity coupling amplitudes

In Eq. 5.12, the complex terms $a^{J_{\Xi_b}}_{\lambda_{\rm Res},\lambda_K}$ and $b^{J_{\rm Res}}_{\lambda_p,\lambda_K}$ are rotationally invariant helicity coupling amplitudes whose phase and magnitude are the free parameters of the model that are inferred

from the fit to data.

Since the decay $\mathrm{Res} \to K^- p$ is a strong decay, it conserves parity. This property leads to the following condition [153, 154]

$$b^{J_{\mathrm{Res}}}_{\lambda_p,\lambda_K} = (-1)^{j_p + j_K - j_{\mathrm{Res}}} \eta_{\mathrm{Res}} \eta_K \eta_p b^{J_{\mathrm{Res}}}_{-\lambda_p,-\lambda_K}, \tag{5.16}$$

which reduces the number of free parameters in the model when summing over $\lambda_p = \pm 1/2$ (as the proton spin is not measured and $\lambda_K = 0$). There is an ambiguity in whether to write the amplitude in terms of $b^{J_{\mathrm{Res}}}_{+1/2,0}$ or $b^{J_{\mathrm{Res}}}_{-1/2,0}$. In my case, I have chosen to write the amplitude in terms of $b^{J_{\mathrm{Res}}}_{+1/2,0}$ only. The free parameters can be further reduced by absorbing the magnitude and phase of $b^{J_{\mathrm{Res}}}_{\lambda_p,0}$ into a redefinition of $a^{J_{\Xi_b}}_{\lambda_{\mathrm{Res}},\lambda_K}$. This procedure essentially halves the free parameters of the model, however it does not affect the model dependence on the number of physical degrees of freedom. Therefore, adding a single resonance (Res) adds 2 complex helicity couplings ($a^{\pm} = a^{J_{\Xi_b}=1/2}_{\lambda_{\mathrm{Res}}=\pm 1/2,\lambda_K=0}$), corresponding to $\lambda_{\mathrm{Res}} = \pm 1/2$, and hence adds 4 free parameters to the model. This number is, however, increased by a further factor of 2 when allowing the most generic CP violation effects.

The helicity couplings in the CP violation fits is parametrised as follows:

$$a^{\pm} = (x^{\pm} + \delta x^{\pm}) + i(y^{\pm} + \delta y^{\pm}) \tag{5.17}$$

$$\overline{a^{\pm}} = (x^{\pm} - \delta x^{\pm}) + i(y^{\pm} - \delta y^{\pm}) \tag{5.18}$$

where $a^{\pm}$ and $\overline{a^{\pm}}$ are helicity couplings for Ξ_b^- and Ξ_b^+ models, respectively. Here $x^{\pm}, y^{\pm}$ are CP conserving values and $\delta x^{\pm}, \delta y^{\pm}$ are CP violating values. When establishing the baseline model I set the CP violating parameters to zero.

The CP asymmetry in the magnitude of each contributing resonance is measured and is given by:

$$A^{CP}_{\pm} = \frac{|a^{\pm}|^2 - |\overline{a^{\pm}}|^2}{|\overline{a^{\pm}}|^2 + |a^{\pm}|^2} \tag{5.19}$$

5.4.1.3 Resonance lineshape

The term $R(M_{pK} = m(pK^-)_{\mathrm{low}})$ in Eq. 5.12 is a rotationally invariant amplitude that describes the lineshape of each resonant or nonresonant contribution. Resonances are parametrised with relativistic Breit–Wigner functions [155, 156] that are modified by Blatt-Weisskopf form

factors [157] and are given by

$$R(M_{Kp}) = B'_{L^{\rm Res}_{\Xi_b}}(p, p_0, d) \left(\frac{p}{M_{\Xi_b}}\right)^{L^{\rm Res}_{\Xi_b}} \times BW(M_{Kp}|M_0^{\rm Res}, \Gamma_0^{\rm Res}) \times B'_{L_{\rm Res}}(q, q_0, d) \left(\frac{q}{M_0^{\rm Res}}\right)^{L_{\rm Res}} . \tag{5.20}$$

Here, p is the momentum of Res in the Ξ_b rest frame ($p = |\vec{p}^{\,[\Xi_b]}_{\rm Res}|$). Similarly, q is the K momentum in the Res rest frame ($q = |\vec{p}^{\,[\rm Res]}_K|$). The symbols p_0 and q_0 denote values of these quantities at the resonance peak ($M_{Kp} = M_0^{\rm Res}$). The orbital angular momentum between the Res and K^- particles in the Ξ_b decay is denoted as $L^{\rm Res}_{\Xi_b}$. Similarly, $L_{\rm Res}$ is the orbital angular momentum between the p and K in the Res decay. The orbital angular momentum barrier factors, $p^L\, B'_L(p, p_0, d)$, involve the Blatt–Weisskopf functions [157]:

$$B'_0(p, p_0, d) = 1\,, \tag{5.21}$$

$$B'_1(p, p_0, d) = \sqrt{\frac{1 + (p_0\, d)^2}{1 + (p\ \ d)^2}}\,, \tag{5.22}$$

$$B'_2(p, p_0, d) = \sqrt{\frac{9 + 3(p_0\, d)^2 + (p_0\, d)^4}{9 + 3(p\ \ d)^2 + (p\ \ d)^4}}\,, \tag{5.23}$$

$$B'_3(p, p_0, d) = \sqrt{\frac{225 + 45(p_0\, d)^2 + 6(p_0\, d)^4 + (p_0\, d)^6}{225 + 45(p\ \ d)^2 + 6(p\ \ d)^4 + (p\ \ d)^6}}\,, \tag{5.24}$$

and account for difficulty in creating the orbital angular momentum L, which depends on the momentum of the decay products p (in the rest frame of the decaying particle) and on the size of the decaying particle given by the d constant. I set $d = 5.0\ \mathrm{GeV}^{-1}$ for Ξ_b and $1.5\,\mathrm{GeV}^{-1}$ for resonances(as done is Ref. [126], however a systematic will be assigned by varying these parameters). The relativistic Breit–Wigner (RBW) amplitude is given by [155, 156]:

$$BW(M|M_0, \Gamma_0) = \frac{1}{M_0^2 - M^2 - iM_0\Gamma(M)}\,, \tag{5.25}$$

where

$$\Gamma(M) = \Gamma_0 \left(\frac{q}{q_0}\right)^{2\,L_{\rm Res}+1} \frac{M_0}{M} B'_{L_{\rm Res}}(q, q_0, d)^2\,. \tag{5.26}$$

In the case of the $\Lambda(1405)$ resonance, which peaks below the pK^- threshold, the M_0 is replaced by an "effective mass" in the kinematically allowed region, calculated using the ad-hoc

formula,

$$M_0^{\text{eff}} = M^{\min} + (M^{\max} - M^{\min}) \left(1 + \tanh\left(\frac{M_0 - \frac{M^{\max}+M^{\min}}{2}}{M^{\max} - M^{\min}}\right) \right) \tag{5.27}$$

where $M^{\max}$ and $M^{\min}$ are the upper and lower limits of the kinematically allowed range, respectively. The q_0 value in Eq. 5.26 is then the value of q at $M = M_0^{\text{eff}}$. This parameterisation ensures that only the tail of the RBW function enters the DP as a virtual contribution. Note that in the pentaquark analysis of $\Lambda_b^0 \to J/\psi\, pK^-$ decays [21], the $\Lambda(1405)$ resonance is modelled with a two-component width (one corresponding to the pK^- channel and the other corresponding to the dominant $\Sigma^+\pi^-$ channel) equivalent to the Flatté parameterisation [156]. This shape will be considered for this component when evaluating the systematic uncertainty due to the model dependence.

Angular momentum conservation in the weak decay $\Xi_b^- \to \text{Res}\, K^-$ imposes, $\max(J_{\text{Res}} - (J_{\Xi_b} + J_{K^-})\,,\, 0) \leq L_{\Xi_b}^{\text{Res}} \leq J_{\text{Res}} + (J_{\Xi_b} + J_{K^-})$. I assume the minimal value of $L_{\Xi_b}^{\text{Res}} = J_{\text{Res}} - 1/2$ when calculating $R(M_{Kp})$. Angular momentum conservation in the decay of $\text{Res} \to pK^-$ also limits L_{Res} to $J_{\text{Res}} \pm \frac{1}{2}$, which is then uniquely defined by parity conservation in the Res decay, $P_{\text{Res}} = (-1)^{L_{\text{Res}}+1}$.

If no final state interaction is assumed, then the distribution of the nonresonant contribution is uniform over $m(pK^-)_{\text{low}}$. However this is almost never the case and therefore to parametrise the nonresonant component, I use the exponential nonresonant line shape as defined below [126, 156]:

$$R_{\text{NR}\,\text{exp}}(M_{Kp}) = \left(\frac{p}{M_{\Xi_b}}\right)^{L_{\Xi_b}^{\text{Res}}} \left(\frac{q}{M_0^{\text{Res}}}\right)^{L_{\text{Res}}} \exp(-\alpha M_{Kp}^2), \tag{5.28}$$

where α is a slope parameter that is floated in the fit. An alternate polynomial shape for the nonresonant components will be investigated when studying systematic uncertainty introduced by the assumptions made about the fit model.

5.4.1.4 Differential decay density

To get the differential decay density for $\Xi_b^- \to \text{Res}(\to K^-p)K^-$ decay as a function of Ω, the amplitude shown in Eq. 5.12 is summed coherently over all n resonant and nonresonant contributions and resonance helicities, since they lead to the same final state with the same

helicities, and incoherently over all helicities of the initial and final states. It is given by

$$\frac{d\Gamma}{d\Omega} = \sum_{M_{\Xi_b},\lambda_p} (\frac{1+2M_{\Xi_b}P_{\Xi_b}}{2}) \left| \sum_{n,\lambda_{\rm Res}} A^n_{M_{\Xi_b},\lambda_{\rm Res},\lambda_p}(\Omega) \right|^2 \tag{5.29}$$

where P_{Ξ_b} is the difference in probability of Ξ_b^- to be in the initial spin state of $1/2$ and $-1/2$. When Ξ_b^- is unpolarised ($P_{\Xi_b} = 0$) the density only depends on $m(pK^-)^2_{\rm low}$ and $\theta_p^{[\rm Res]}$ (or $m(pK^-)^2_{\rm high}$). As discussed previously in Sec. 5.2.2, in the case when Ξ_b^- is polarised, it is shown in Ref. [126] that, if the detection efficiency is symmetric over $\cos(\theta^{[\Xi_b]}_{\rm Res})$ (which it is, to a reasonable approximation, as shown in Figure5.8), the polarisation effect cancels out and the density still depends only on the two DP variables. Therefore, in this study I only consider dependence of the density on the normal DP variables $m(pK^-)^2_{\rm low}$ and $m(pK^-)^2_{\rm high}$ and either integrate out the other angles or fix their values. In this analysis I choose to fix the angles defined in the previous chapter in Sec.4.3, which describe the orientations of different decay planes and are functions of the polar and azimuthal angles defined in this section, to zero.

In an example decay of $P \to {\rm Res}(\to ab)c$, I show the decay density defined using the non-symmetric amplitude as given in Eq. 5.12, as a function of the cosine of the helicity angle of the resonance ($\cos(\theta_{ab})$) and the invariant mass m_{ab} for a resonance with positive parity and various spin hypotheses $J = 1/2, 3/2, 5/2, 7/2$ in Figure 5.15. In this example, the resonance mass is set to be 2.35 GeV and its width is set to be 0.7 GeV. When the symmetric amplitude defined in Eq. 5.12 is used, the $\cos(\theta_{ab})$ distribution depend on the phase space location of the resonance.

I also define here the fit fraction for the resonance labelled i as

$$FF^i = \frac{\int_\Omega \sum_{M_{\Xi_b},\lambda_p} (\frac{1+2M_{\Xi_b}P_{\Xi_b}}{2}) \left| \sum_{\lambda_{\rm Res}} A^i_{M_{\Xi_b},\lambda_{\rm Res},\lambda_p} \right|^2 d\Omega}{\int_\Omega \frac{d\Gamma}{d\Omega} d\Omega} \tag{5.30}$$

Note that the sum of fit fractions (FF) for all the resonances and nonresonant components in a model does not need to be 100%, due to interference effects that are nonzero after integrating over the phase space. Such effects are quantified by interference fit fractions. When investigating CP violation effects, I also define a fit fraction asymmetry as follows:

$$\text{FF asymmetry} = \frac{FF - \overline{FF}}{\overline{FF} + FF} \tag{5.31}$$

where FF and $\overline{FF}$ are fit fractions defined for the Ξ_b^- and Ξ_b^+ models, respectively.

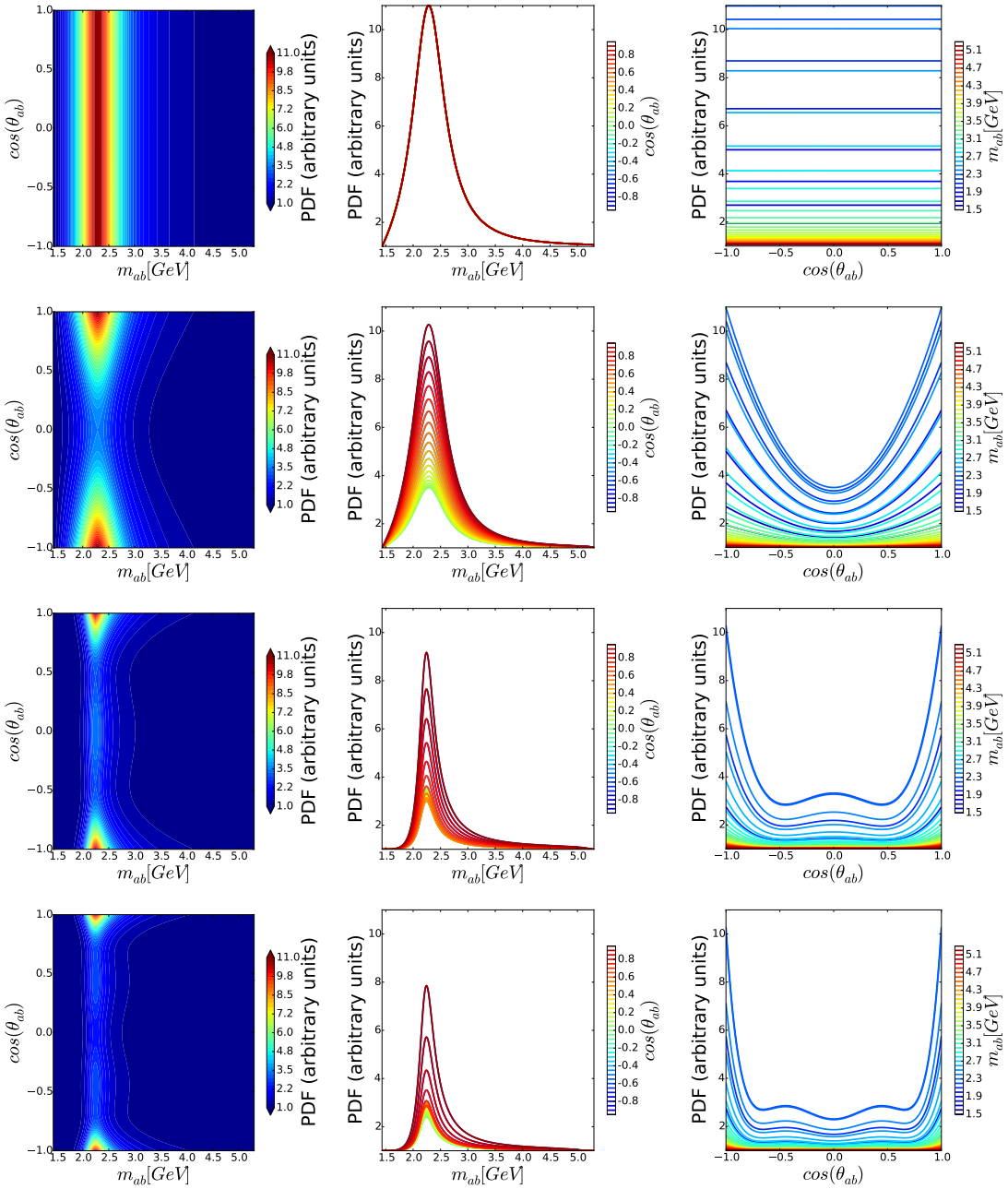

Figure 5.15: In an example decay of $P \to \mathrm{Res}(\to ab)c$, with resonance of positive parity and (top to bottom) spin hypothesis $J = 1/2, 3/2, 5/2, 7/2$, (left) the decay density defined using the non-symmetric amplitude as given in Eq. 5.12 as functions of the helicity angle of the resonance (θ_{ab}) and invariant mass m_{ab}. The projections of PDF onto the (left) x axis in slices along the y axis and similarly (right) y axis in slices along the x axis. In this example the resonance mass is set to be 2.35 GeV and its width is set to be 0.7 GeV. When the symmetric amplitude shown in Eq. 5.12 is used in the definition of the decay density, the angular distributions depends on the phase space location of the resonance.

5.4.2 Combinatorial background PDF

To model the combinatorial background PDF, usually the candidates in the sideband above the signal region (for my case $m(pKK) > 5890\,\mathrm{MeV}$) are used to obtain the DP distribution. However, this procedure can introduce some bias into the combinatorial background PDF, since going higher in $m(pKK)$ leads to an increase in phase space and therefore the area of the unconstrained DP also increases. For conducting fits using the candidates in the signal region, I use the constrained DP, in which $m(pKK) = m_{\Xi_b^-}$ is enforced, as this improves the resolution for signal decays. Application of this constraint to candidates away from the signal region however causes some distortion which could be important if, for example, decays of narrow resonances are present in the background.

To get around this problem, one could consider use of the same-sign background sample (*viz.* $\overline{p}K^-K^-$ sample) to obtain the DP distribution in the signal region. However, this same-sign sample is not expected to accurately represent the combinatorial background, which can in principle contain real Λ^* resonances. Therefore, the PDF obtained using the same-sign sample alone does not capture the true background PDF.

To overcome these issues, I have investigated two methods, one based on Gaussian processes (GPy) and the other on neural networks (NN), of extrapolating the DP distribution from the RSB to obtain the combinatorial background PDF. The GPy method is a binned approach documented in Refs [137, 158]. The second approach is an unbinned approach which relies on training neural networks using the right side band (RSB) candidates. In both of these approaches I use candidates above $m(pKK) > 5.89\,\mathrm{GeV}$ to conduct 3D fits in $(m(pKK), m(pK^-)^2_{\mathrm{low}}, m(pK^-)^2_{\mathrm{high}})$ space and extract the corresponding model parameters.

The Gaussian process method (GPy) has been implemented in the GPy python package [159]. The parameters of the GPy model are obtained by fitting the RSB $(m(pKK), m(pK^-)^2_{\mathrm{low}}, m(pK^-)^2_{\mathrm{high}})$ space in uniform bins of $(10, 5, 5)$ respectively. Its prediction of the DP distribution in the signal region is then obtained.

In the neural network method (NN), I assume that the PDF function describing the $\omega = (m(pKK), m(pK^-)^2_{\mathrm{low}}, m(pK^-)^2_{\mathrm{high}})$ space is given by:

$$\begin{aligned} F(\omega) = \big| & f_0(m(pK^-)^2_{\mathrm{low}}, m(pK^-)^2_{\mathrm{high}}) \\ & + \exp(-\alpha m(pKK)) \times f_1(m(pK^-)^2_{\mathrm{low}}, m(pK^-)^2_{\mathrm{high}}) \big|^2 \end{aligned} \tag{5.32}$$

where α is a free parameter to be determined from the fit. The functions f_0 and f_1 are modelled as neural networks that are trained on the RSB data with 2 input variables ($m(pK^-)^2_{\mathrm{low}}$, $m(pK^-)^2_{\mathrm{high}}$), 2 hidden layers with 8 nodes each, sigmoid activation functions and using the negative log likelihood as the cost function to be minimised to determine various parameters of

the neural network (weights and biases); and its prediction of the DP distribution in the signal region is then obtained.

It was observed that the presence of real Λ^* resonances, as seen in the DP distribution for the candidates lying in the RSB region of $m(pKK)$ is modelled well using the NN method. Therefore the PDF obtained from this method is used to describe the combinatorial background shape in the amplitude analysis. The background distributions obtained using this method are shown in Figure 5.16 for Run I (2011–2012) and Run II (2015–2016).

I also investigated using this method to obtain combinatorial background PDFs split by the charge of the initial state. However, due to lack of training data the background shapes split by initial state charge do not model the shapes correctly. Therefore, in the baseline fit I assume that any asymmetry in the combinatorial background is negligible and use the PDF extrapolated from the RSB of the combined Ξ_b^- and Ξ_b^+ sample. Possible background asymmetry will be a source of systematic uncertainty, which I plan to evaluate by obtaining separate PDFs after loosening the MVA output requirement in the RSB region, which should allow to obtain reliable background shapes from the neural network method. This choice should not induce any bias as no significant correlation is observed between the MVA output and DP variables.

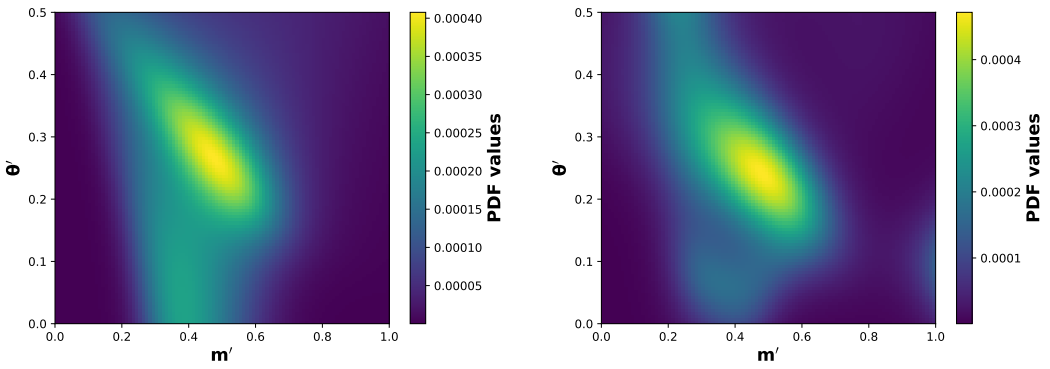

Figure 5.16: The distribution of combinatorial background in the signal region as a function of square DP position for (left) Run I (2011–2012) and (right) Run II (2015–2016).

5.4.3 Crossfeed $\Xi_b^- \to pK^-\pi^- \to pKK$ background PDF

The DP distribution of the $\Xi_b^- \to pK^-\pi^-$ decay is not known. For the fit to $m(pKK)$, I weight the $\Xi_b^- \to pK^-\pi^-$ MC sample to match the square DP distribution to that of a cocktail model, which consists of $\Lambda(1405)$, $\Lambda(1520)$, $\Lambda(1690)$, $N(1440)$, $N(1520)$, $N(1535)$ and $N(1650)$ resonances. The background DP distribution with a π reconstructed as a K is then obtained with this weighted MC sample. The background DP distribution of $\Xi_b^- \to pK^-\pi^- \to pKK$ samples

is shown in Fig 5.17 for Run I (2011–2012) and Run II (2015–2016). Systematic uncertainties will be evaluated due to the model dependence by adding and removing components from the cocktail DP model.

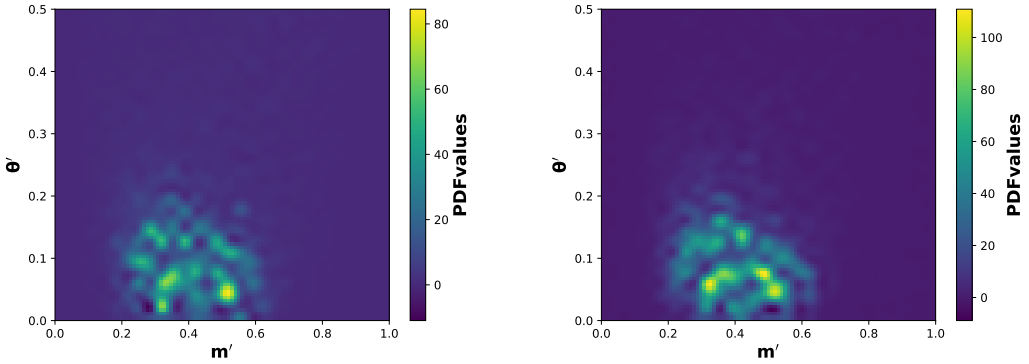

Figure 5.17: (Left) For Run I (2011–2012) and (right) Run II (2015–2016) the distribution of the $\Xi_b^- \to pK^-\pi^-$ cross-feed background as a function of square DP position in the signal region.

The asymmetry for the $\Xi_b^- \to pK^-\pi^-$ cross-feed mode is not known a priori. In the baseline CPV fits, I assume zero local asymmetry for this component. Since the fraction of this component in the signal region is low, negligible effect is expected on my result due to this assumption. To evaluate the associated systematic uncertainty, I allow a global asymmetry ($A_{\rm crsfd}$ defined in Eq. 5.8) of $\pm 20\%$ in the crossfeed background, which is considered to be conservative.

5.4.4 Establishing the baseline model

For the amplitude analysis, I conduct an unbinned maximum-likelihood fit to both Run I (2011–2012) and Run II (2015–2016) data simultaneously with the corresponding efficiency maps and background PDFs. Note that before combining the data corresponding to Run I, it was checked that no significant differences were observed between the relative efficiency maps between 2011 and 2012 data taking periods. A similar study was conducted for the Run II (2015–2016) data sample.

As mentioned previously, only resonances in pK^- are possible in the $\Xi_b^- \to pK^-K^-$ decay through the $\Delta I = 0$ weak decay $\Xi_b^- \to {\rm Res}\, K^-$, where ${\rm Res} = \Lambda^*, \Sigma^*$. There are no significant $\Delta I > 0$ processes that could contribute to the pKK final state. The Particle Data Group (PDG) [151] reports a large number of Λ^* and Σ^* states; those that are sufficiently well-established to be ranked **** or *** are summarised in Table 5.10 (a single * means its has been confirmed by an experiment). For many of these states, the PDG does not report masses and

widths with central values and uncertainties, but does give the real and imaginary parts of the pole position in that format. This reflects the fact that a simple Breit–Wigner parameterisation of these resonances does not fully describe their lineshapes; however, more sophisticated parametrisations are beyond the scope of the current analysis with limited statistics.

Table 5.10: Summary of the Λ^* and Σ^* resonances reported in the PDG and ranked **** or *** [151]. Note that the pK^- threshold is at $1432\,\mathrm{MeV}$. For $\Sigma(2250)$, the spin parity is not known and I have assumed it to be $\frac{3}{2}^+$. Contributions from states with $J \geq \frac{7}{2}$ are suppressed and hence not considered in the fit model.

Components that are ranked ****				
Name	J^P	Mass $(\,\mathrm{MeV})$	Width $(\,\mathrm{MeV})$	Main decay channels
$\Lambda(1405)$	$\frac{1}{2}^-$	$1405.1\,^{+1.3}_{-1.0}$	50.5 ± 2.0	$\Sigma\pi$
$\Lambda(1520)$	$\frac{3}{2}^-$	1519.5 ± 1.0	15.6 ± 1.0	$N\overline{K}$, $\Sigma\pi$
$\Lambda(1670)$	$\frac{1}{2}^-$	1660 to 1680	25 to 50	$N\overline{K}$, $\Sigma\pi$, $\Lambda\eta$
$\Lambda(1690)$	$\frac{3}{2}^-$	1685 to 1695	50 to 70	$N\overline{K}$, $\Sigma\pi$, $\Lambda\pi\pi$, $\Sigma\pi\pi$
$\Lambda(1820)$	$\frac{5}{2}^+$	1815 to 1825	70 to 90	$N\overline{K}$
$\Lambda(1830)$	$\frac{5}{2}^-$	1810 to 1830	60 to 110	$\Sigma\pi$
$\Lambda(1890)$	$\frac{3}{2}^+$	1850 to 1910	60 to 200	$N\overline{K}$
$\Lambda(2100)$	$\frac{7}{2}^-$	2090 to 2110	100 to 250	$N\overline{K}$
$\Sigma(1670)$	$\frac{3}{2}^-$	1665 to 1685	40 to 80	$\Sigma\pi$
$\Sigma(1775)$	$\frac{5}{2}^-$	1770 to 1780	105 to 135	$N\overline{K}$, $\Lambda^{(*)}\pi$
$\Sigma(1915)$	$\frac{5}{2}^+$	1900 to 1935	80 to 160	not clear
$\Sigma(2030)$	$\frac{7}{2}^+$	2025 to 2040	150 to 200	$N\overline{K}$, $\Lambda^{(*)}\pi$, $\Delta\overline{K}$
Components that are ranked ***				
$\Lambda(1600)$	$\frac{1}{2}^+$	1560 to 1700	50 to 250	$N\overline{K}$, $\Sigma\pi$
$\Lambda(1800)$	$\frac{1}{2}^-$	1720 to 1850	200 to 400	$N\overline{K}^{(*)}$, $\Sigma\pi$, $\Lambda\eta$
$\Lambda(1810)$	$\frac{1}{2}^+$	1750 to 1850	50 to 250	$N\overline{K}^{(*)}$, $\Sigma\pi$, $\Lambda\eta$, ΞK
$\Lambda(2110)$	$\frac{5}{2}^+$	2090 to 2140	150 to 250	$N\overline{K}^{(*)}$, $\Sigma\pi$, $\Lambda\omega$
$\Lambda(2350)$	$\frac{9}{2}^+$	2340 to 2370	100 to 250	$N\overline{K}$, $\Sigma\pi$, $\Lambda\omega$
$\Sigma(1660)$	$\frac{1}{2}^-$	1630 to 1690	40 to 200	$N\overline{K}$, $\Sigma\pi$, $\Lambda\pi$
$\Sigma(1750)$	$\frac{1}{2}^-$	1730 to 1800	60 to 160	$N\overline{K}$, $\Sigma\pi$, $\Lambda\pi$, $\Sigma\eta$
$\Sigma(1940)$	$\frac{3}{2}^-$	1900 to 1950	150 to 300	$N\overline{K}$, $\Sigma\pi$, $\Lambda\pi$
$\Sigma(2250)$	$?^?$	2210 to 2280	60 to 150	$N\overline{K}$, $\Sigma\pi$, $\Lambda\pi$

To build the signal model, I consider only the Λ^* and Σ^* resonances listed in Table 5.10. If these states do not give a good description of the data then less well-established (ranked ** in the PDG) resonances could be considered. This is found to be unnecessary. Furthermore, the spin

$7/2$ states, *i.e.* $\Lambda(2100)$ and $\Sigma(2030)$, are excluded from consideration since the minimal $L^{\rm Res}_{\Xi_b}$ for spin $7/2$ equals $J_{\rm Res} - (J_{\Xi_b} + J_{K^-}) = 3$ (F-wave). A value of $L = 3$ is the highest orbital angular momentum observed, with a very small rate, in decays of B mesons with $Q = 1.78\,\text{GeV}$ [160]. In my case an additional suppression from the spin counting factors present in $\Lambda(2100)$ and $\Sigma(2030)$ states [21] is present. Hence, even though a larger phase-space is available ($Q = 3.2\,\text{GeV}$ for $\Lambda(2100)$ and $Q = 3.27\,\text{GeV}$ for $\Sigma(2030)$), contributions from such spin $7/2$ states are suppressed.

In order to obtain a nominal fit model and conduct CPV studies, the helicity couplings of all resonances with respect to a contributing reference resonance are measured. From Figure 5.13, it we deduce that the resonance with the largest fit fraction would be $\Lambda(1520)$. Therefore, $\Lambda(1520)$ is taken to be the reference resonance. As a result, I fix the values of the two helicity couplings (a^+, a^- corresponding to $\lambda_{\rm Res} = \pm 1/2$) of the $\Lambda(1520)$ reference resonance to:

- $Re(a^+) = 1$ and $Im(a^+) = 0$, when conducting fits to Ξ_b and $\overline{\Xi}_b$ samples, the CP violating part of $Re(a^+)$ is floated to allow for CP violation in $\Xi_b^- \to \Lambda(1520)K^-$ decay.
- $Re(a^-) = 0$, since the analysis is insensitive to this value (confirmed using toys by making scans of the negative log likelihood (NLL) as a function of this variable).
- $Im(a^-) = 0$, since the relative phase between a^+ and a^- is arbitrary. This arbitrariness is introduced since the angular dependence from the decay density is removed which essentially leads to incoherent addition of amplitudes corresponding to the two helicities of the intermediate resonances (see Eq. 9 and 10 in Ref. [126]).

The helicity couplings of all other resonances are left floating in the fit to data.

I have employed three approaches to establish the nominal fit model which are discussed in the following subsections.

5.4.4.1 Bottom-up approach

The first, referred to as the "bottom-up" approach, is where I start by including only the reference resonance $\Lambda(1520)$. Then each possible component from all possible states shown in Table 5.10 is added, including nonresonant components with $J^P = \frac{1}{2}^+, \frac{1}{2}^-, \frac{3}{2}^+, \frac{3}{2}^-$ but excluding $\Lambda(2100), \Sigma(2030)$ states. The additional component that yields the largest improvement in NLL value is then added to the model, and the process repeated so long as the change in NLL satisfies $-\Delta\text{NLL} > 4.5 \equiv \sqrt{-\Delta 2\text{NLL}} > 3$.[3] The process ends once addition of extra components and

[3] The $\sqrt{-\Delta 2\text{NLL}}$ is the approximate Gaussian significance, but does not take into account the change in the number of degrees-of-freedom (DoF) between the models being compared. While comparing models that have Δ DoF $= 4$ or 5, the $\sqrt{-\Delta 2\text{NLL}}$ is an overestimate of the significance. When I converge to a baseline model, I evaluate the significance of any marginal components taking into account the change in DoF.

removal of existing components leads to no significant change in the NLL value.

Note that a fit with a given model can suffer from multiple solutions. In an attempt to ensure that the fit converges to the global minimum, the fit is conducted multiple times randomising every time the starting values of the helicity couplings and picking the smallest NLL (also requiring that the fit converges with MINUIT status = 3, assuring the covariance matrix is accurate).

Components with different J^P quantum numbers should have zero interference fit fractions due to the orthogonality relation satisfied by the Wigner-D matrix elements. (In practice, the symmetrisation of the Dalitz plot can lead to non-zero values for such interference fit fractions in this analysis.) However the interference between overlapping components with the same J^P quantum numbers can yield unphysically large interference fit fractions ($> 50\%$). When such a case occurs, I discard the interfering component that is non-significant.

In Table 5.11, I show the components of the baseline model and their significance obtained with the "bottom-up" approach. In this table we quote the "Gaussian significance" that is obtained by translating the $-2\Delta\text{NLL}$ into a corresponding p-value for a χ^2 PDF with Δ DoF as the number of degrees of freedom. The quoted value is then evaluated as $\sqrt{2}\,\text{erfc}^{-1}(p)$. In Figure 5.18, I show the results of the fit to data with this model. The fit fractions and interference fit fractions obtained from this fit are shown in Figure 5.19.

Table 5.11: A rough estimate of the significance of the components of the baseline model reached through the "bottom-up" approach. Please refer to the text for more details on how these significances were evaluated.

Components	Δ DoF	-2ΔNLL	Gaussian significance
$\Lambda(1520)$ (reference)	-	-	-
$\Lambda(1405)$	4	31.8	4.7
$\Lambda(1670)$	4	85.6	8.5
$\text{NR}(\frac{1}{2}^+)$	5	37.3	5.0
$\text{NR}(\frac{3}{2}^-)$	5	28.5	4.1

5.4.4.2 Top-down approach

I also performed a "top-down" approach, where I start by including all the states shown in Table 5.10, including nonresonant components with $J^P = \frac{1}{2}^+, \frac{1}{2}^-, \frac{3}{2}^+, \frac{3}{2}^-$ but excluding the $\Lambda(2100), \Sigma(2030)$ components. Models are then constructed by removing each component one by one, and fits performed. The model that gives the smallest increase in NLL forms

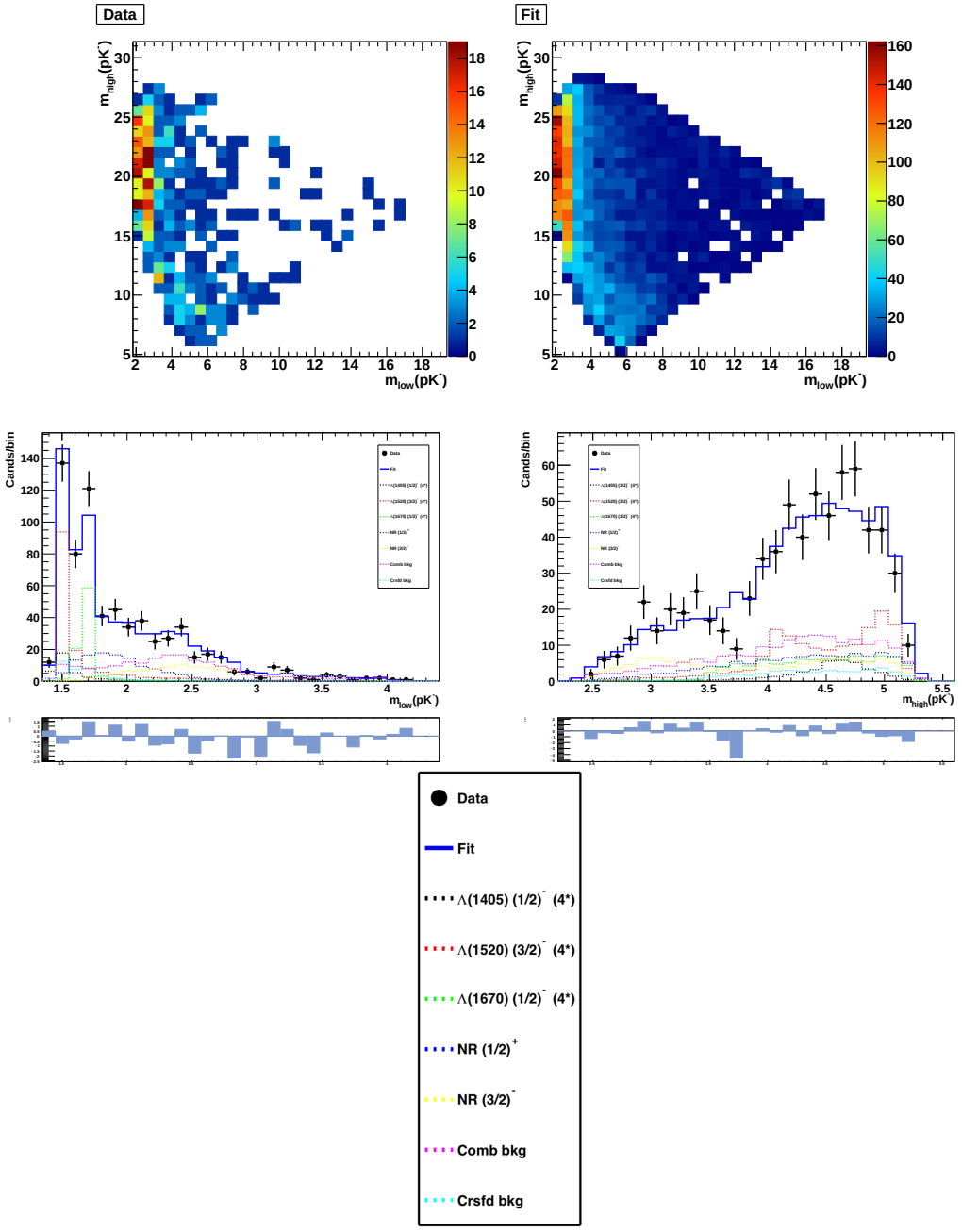

Figure 5.18: Result of the "bottom-up" approach: Run I (2011–2012) + Run II (2015–2016) (top left) data sample and (top right) fit result. Shown also are the fit projections along (middle-left) $m(pK^-)_{\text{high}}$ and (middle-right) $m(pK^-)_{\text{low}}$ together with the data. The zoomed-in (bottom) legend is also shown.

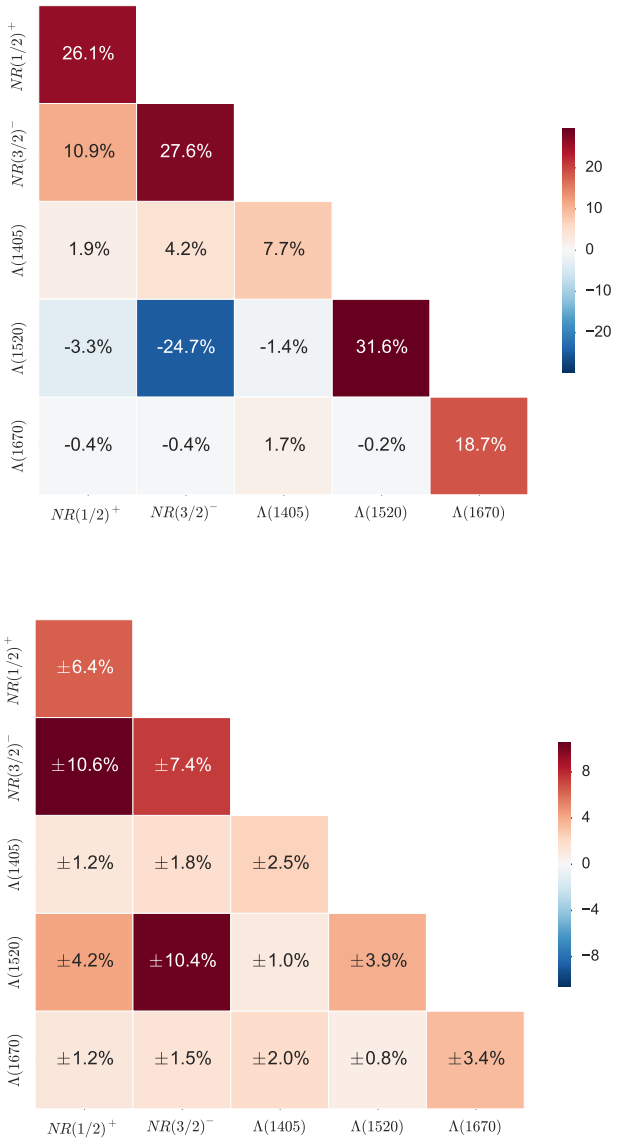

Figure 5.19: (Top) Fit fractions and interference fit fractions (%) from the fit to the Run I (2011–2012) + Run II (2015–2016) data sample using the baseline fit model obtained using the "bottom-up" approach. (Bottom) Corresponding errors on these quantities determined by conducting a fit to 100 pseudo-experiments, each generated with as many candidates as seen in data. A negative sign on the fit fraction indicates destructive interference. The sum of all fit fractions and interference fit fractions gives 100% by definition. The sum of diagonal elements gives 111.7%. The fit fractions from toys lead to Gaussian distributions, however due to a floated nonresonant shape parameters the interference fit fractions do not entirely follow Gaussian distribution.

the baseline model for the next iteration of the fit, so long as the change in NLL satisfies $-\Delta\text{NLL} < 4.5 \equiv \sqrt{-\Delta 2\text{NLL}} < 3$. The process ends once removal of any further component from the model causes too large a change in the NLL value. When compared to the bottom-up approach, it is highly likely that in the top-down approach multiple components with same J^P quantum numbers could give rise to large unphysical interference fit fractions ($> 50\%$). In such a case, the two components that give the largest interference fit fraction are investigated and the component that when removed gives the minimal NLL is discarded. As done in the bottom-up approach, I perform multiple fits in each configuration to try to reach the global NLL minimum.

In Table 5.12, I show the components of the baseline model and their significance obtained through the "top-down" approach. In Figure 5.20, I show the plots obtained from the fit to data. The fit fractions and interference fit fractions obtained from the result of the fit are shown in Figure 5.21.

Table 5.12: A rough estimate of the significance of the components of the baseline model reached through the "top-down" approach. Please refer to Sec 5.4.4.1, for more details on how these significances were evaluated.

Components	Δ DoF	$-2\Delta\text{NLL}$	Gaussian significance
$\Lambda(1520)$ (reference)	-	-	-
$\Lambda(1405)$	4	46.1	5.9
$\Lambda(1670)$	4	130.9	10.8
$\Sigma(1775)$	4	26.7	4.2
$\text{NR}(\frac{3}{2}^+)$	5	61.1	6.8

5.4.4.3 LASSO regularisation approach

The LASSO regularisation method [161, 162] adds a penalty term to $-\text{NLL}$ for every four parameters of the model (magnitudes and phases of two amplitudes corresponding to $\lambda_{\text{Res}} = \pm 1/2$),[4] therefore the function to minimise changes as follows:

$$-2\ln(\mathcal{L}) \rightarrow -2\ln(\mathcal{L}) + \lambda \times \sum_{M_{\Xi_b},\lambda_p,i} \sqrt{\int \left| A^i_{M_{\Xi_b},\lambda_{\text{Res}}=1/2,\lambda_p}(\Omega) + A^i_{M_{\Xi_b},\lambda_{\text{Res}}=-1/2,\lambda_p}(\Omega) \right|^2 d\Omega} \tag{5.33}$$

[4] A toy study was conducted in which a penalty term for every two parameters of the model as done in Ref. [161, 162] was also investigated instead of four parameters, however this did not affect the results of the study.

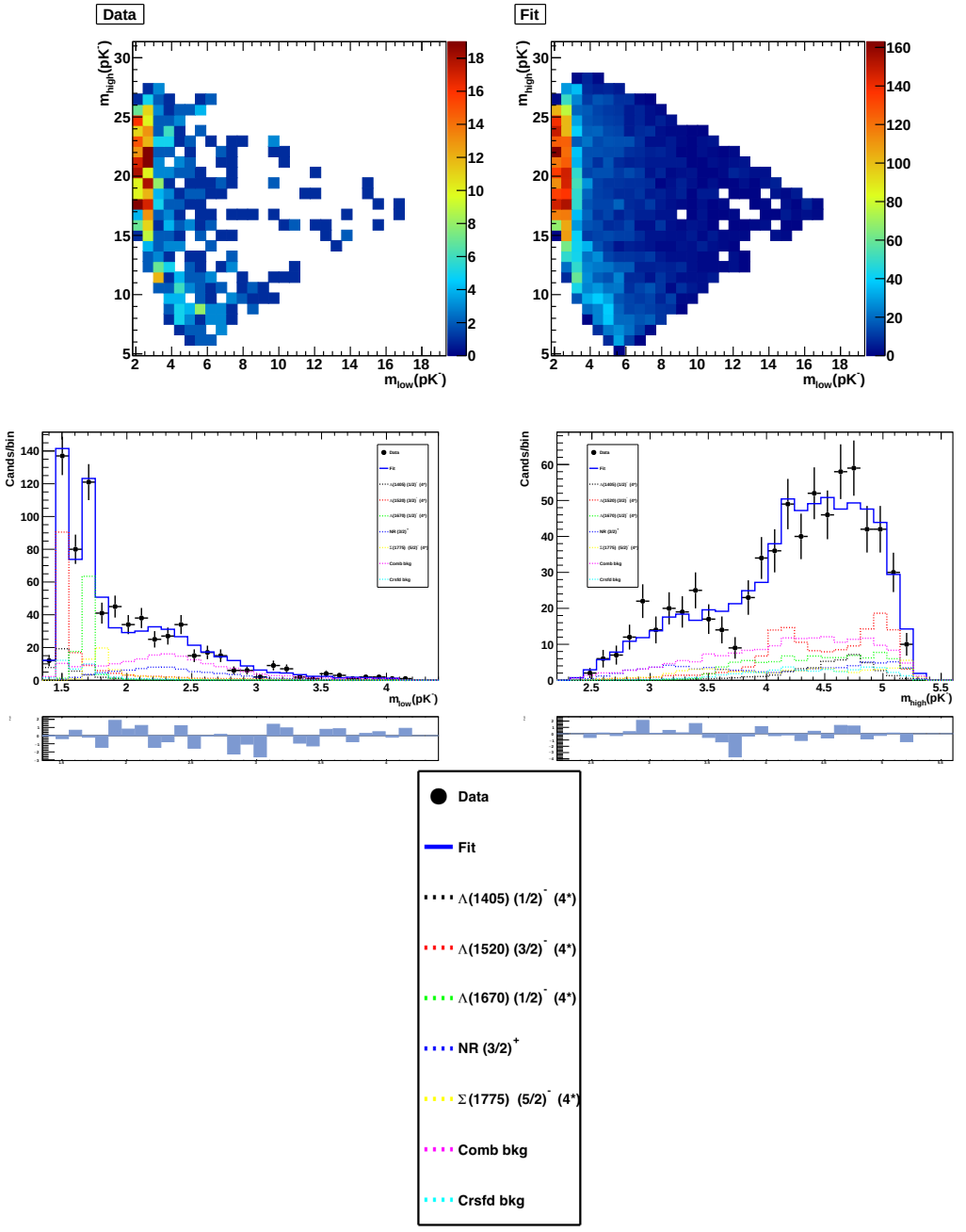

Figure 5.20: Result of the "top-down" approach: Run I (2011–2012) + Run II (2015–2016) (top left) data sample and (top right) fit result. Shown also are the fit projections along (middle-left) $m(pK^-)_{\text{high}}$ and (middle-right) $m(pK^-)_{\text{low}}$ together with the data. The zoomed-in (bottom) legend is also shown.

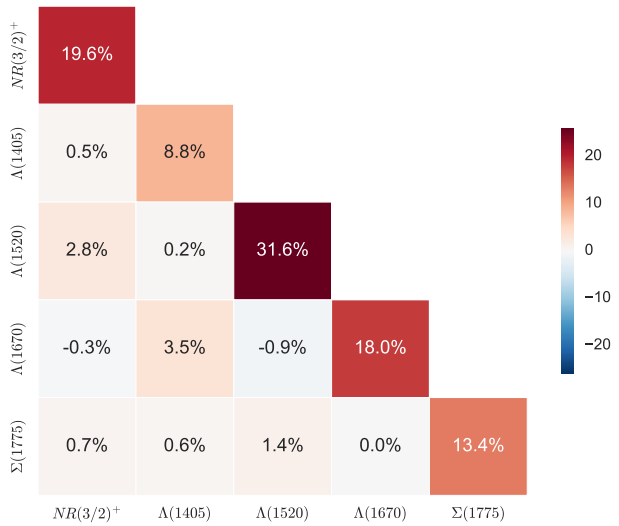

Figure 5.21: Fit fractions and interference fit fractions (%) from the fit to the Run I (2011–2012) + Run II (2015–2016) data sample using the baseline fit model obtained using the "top-down" approach. A negative sign indicates destructive interference. The sum of all fit fractions and interference fit fractions gives 100% by definition. The sum of diagonal elements gives 91.5%. The fit fractions from toys lead to Gaussian distributions, however due to a floated nonresonant shape parameters the interference fit fractions do not entirely follow Gaussian distribution.

A proper choice of λ reduces the number of amplitudes contributing to the fit result, similarly to what is done explicitly with the *ad hoc* procedures discussed in the previous sections. Selecting an optimal λ leads to a model with significantly non-zero fit fractions only for components that are present, at significant levels, in the underlying true physical model. An optimal λ is selected as the value that minimises the BIC criterion, as done in Ref. [161, 162], which is expressed as follows:

$$\mathrm{BIC} = -2\ln(\mathcal{L}) + 2r\ln(n). \tag{5.34}$$

Here n is the number of candidates in the data sample and r is the number of amplitudes in the model that have a fit fraction less than 0.1%.

The LASSO regularisation method was validated using a toy study, in which a finite number of amplitudes were used to generate data which was then fitted for different values of λ using a model that included not only the true components but a large number of other extraneous components. Minimising the BIC criterion led to an optimal value of λ which gave a model that contained true components and a small set of extraneous components with non-zero fit fraction ($\sim 5\%$). Applying the "top-down" approach (with $\lambda = 0$) to this small set of components gave us the true model back.

I then applied the same procedure to the collision data sample where the fit was carried out, for different values of λ, using all the states shown in Table 5.10. Figure 5.22 shows a scan of the BIC criterion with respect to the LASSO parameter λ; the optimal value is found to be $\lambda = 0.1$. I then employ a "top-down" approach (with $\lambda = 0$) to this model to keep only significant components that leads us to a baseline model obtained through the LASSO procedure.

In Table 5.13, I show the components of the baseline model and their significance obtained through the LASSO approach. In Figure 5.23, I show the plots obtained from the fit to data (with $\lambda = 0$). The fit fractions and interference fit fractions obtained from the result of the fit are shown in Figure 5.24.

5.4.4.4 Discussion of all approaches

A summary of all the models obtained from each of the approaches is presented in Table 5.14. It can be seen that all three models contain in common some low mass resonances, however there is uncertainty in establishing the broad components at higher mass. It can be seen from the sum of fit fractions that none of the models causes particularly large (and thus likely to be unphysical) net interference between the components.

The model from the "top-down" approach has one less free parameter compared to the model obtained through the "bottom-up" approach. This is due to the fact that the top-down model contains $\Sigma(1775)$, with $J^P = 5/2^-$, instead of an exponential $\mathrm{NR}(\frac{1}{2}^+)$ component that

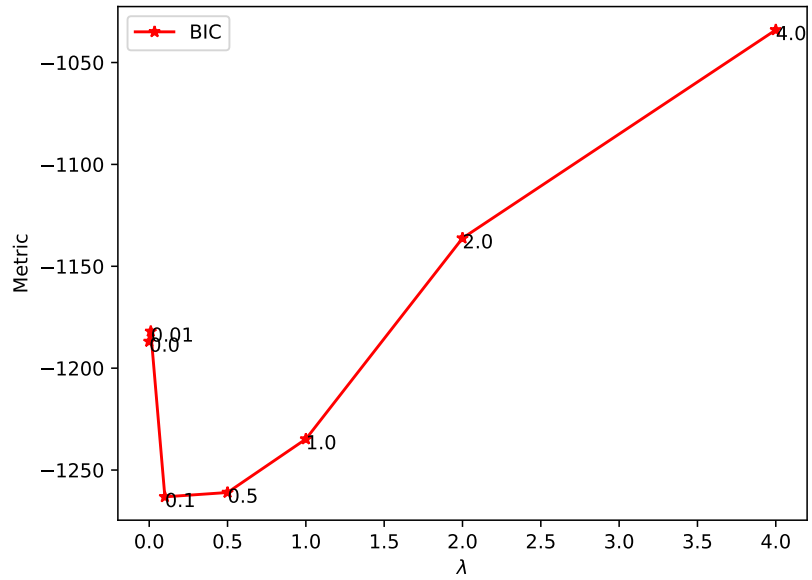

Figure 5.22: The BIC criterion with respect to LASSO parameter λ. The optimal value is found to be $\lambda = 0.1$.

Table 5.13: A rough estimate of the significance of the components of the baseline model reached through the LASSO approach. Please refer to Sec 5.4.4.1, for more details on how these significances were evaluated.

Components	Δ DoF	-2ΔNLL	Gaussian significance
$\Lambda(1520)$ (reference)	-	-	-
$\Lambda(1405)$	4	30.6	4.6
$\Lambda(1670)$	4	91.3	8.8
$\Lambda(1600)$	4	25.0	4.0
$\Lambda(1890)$	4	17.7	3.2
NR($\frac{3}{2}^-$)	5	31.9	4.5

is present in the bottom-up model. An extra free parameter in the bottom-up model accounts for the shape of the nonresonant component. The NLL value of the top-down model, however, is slightly larger compared to the bottom-up model. The LASSO model, on the other hand, contains four additional DoF when compared to the bottom-up model and does not lead to a significantly lower NLL value. Due to these reasons, I have chosen the model obtained from the "bottom-up" approach to be my nominal model. The largest fit fraction ($\sim 32\%$) in this nominal model is for

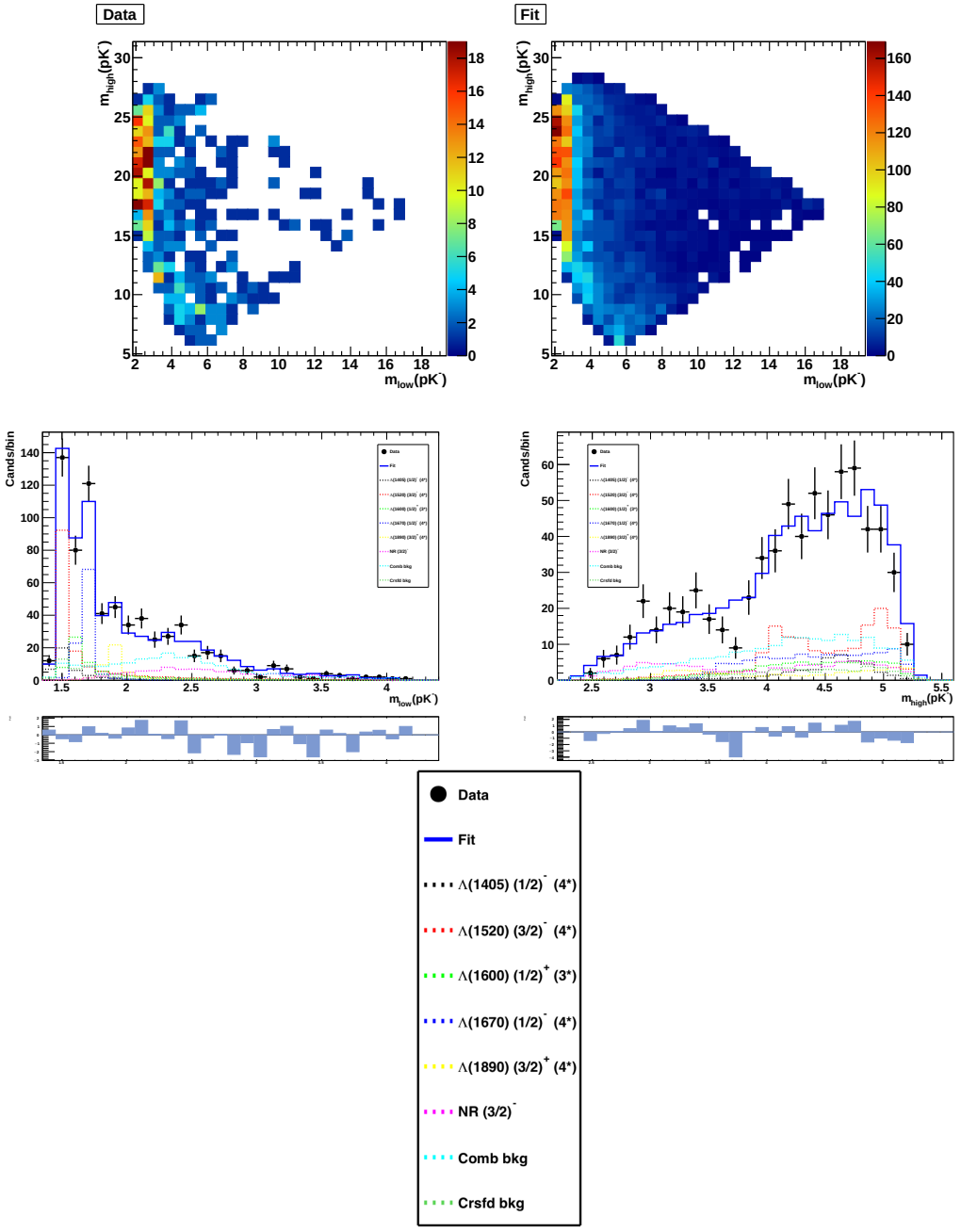

Figure 5.23: Result of the LASSO approach: Run I (2011–2012) + Run II (2015–2016) (top left) data sample and (top right) fit result. Shown also are the fit projections along (middle-left) $m(pK^-)_{\rm high}$ and (middle-right) $m(pK^-)_{\rm low}$ together with the data. The zoomed-in (bottom) legend is also shown.

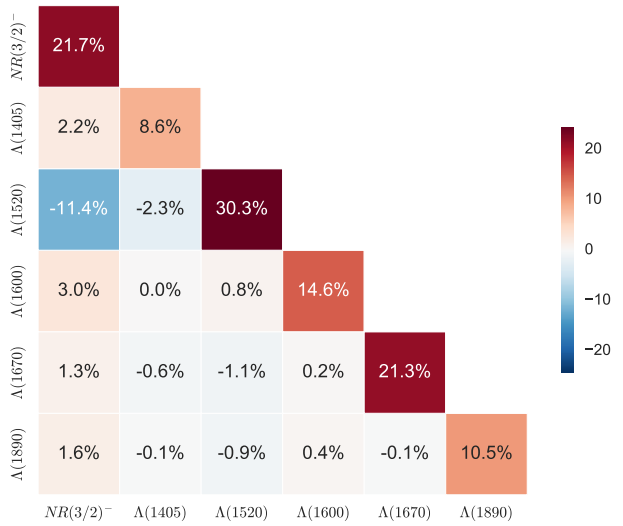

Figure 5.24: Fit fractions and interference fit fractions (%) from the fit to the Run I (2011–2012) + Run II (2015–2016) data sample using the baseline fit model obtained using LASSO approach. A negative sign indicates destructive interference. The sum of all fit fractions and interference fit fractions gives 100%. The sum of diagonal elements gives 107%.

Table 5.14: Summary of the three approaches ("top-down", "bottom-up" and LASSO) used to establish the baseline amplitude model. Shown here are the number of components in the model, the sum of fit fractions obtained, the number of degrees of freedom in the model (DoF), and the minimum NLL obtained from the fit to the data sample. The † represents the chosen nominal model.

Approaches	Components	Sum of FF (%)	DoF	Min $-\Delta$NLL
"Top-down"	$\Lambda(1520)$, $\Lambda(1405)$, $\Lambda(1670)$, $\Sigma(1775)$, NR($\frac{3}{2}^+$)	91.5%	17	-690
"Bottom-up"†	$\Lambda(1520)$, $\Lambda(1405)$, $\Lambda(1670)$, NR($\frac{3}{2}^-$), NR($\frac{1}{2}^+$)	111.7%	18	-694
LASSO	$\Lambda(1520)$, $\Lambda(1405)$, $\Lambda(1670)$, $\Lambda(1600)$, $\Lambda(1890)$, NR($\frac{3}{2}^-$)	107.1%	21	-696

the reference $\Lambda(1520)$ resonance and the largest interference fit fraction ($\sim -25\%$) is due to the destructive interference between $\Lambda(1520)$ and NR($\frac{3}{2}^-$). To inspect the fit quality in different regions of the phase space, I show fit projections in the slices of $m(pK^-)_{\text{low}}$ (Figure 5.25) and $m(pK^-)_{\text{high}}$ (Figure 5.26). Generally a good agreement between the fit and data can be seen from these projections. We have also evaluated the p-value using the two unbinned goodness-of-fit

tests: mixed sample and point-to-point dissimilarity tests as defined in Ref. [163]. The p-value evaluated using mixed sample test for the combined Run I (2011–2012) and Run II (2015–2016) data is 0.1, whereas the one evaluated using point-to-point dissimilarity test yields a p-value 0.3. Therefore, this nominal model is compatible with data.

Initially in the "top-down" and "bottom-up" approaches, the option to float the masses and widths of the resonances was considered. However this rendered the fit unstable and non-convergent. Therefore the masses and widths of the resonances were fixed to values from the PDG in all of these approaches. Uncertainties in these values will be considered as a source of systematic uncertainty.

In results of fits with models obtained from all the considered approaches, there is an apparent discrepancy between the model and the data at $m(pK^-)_{\rm high} \sim 3.7\,{\rm GeV}$, where the sum of components is above the data. As there is also a single bin excess, where the data is above the model, at $m(pK^-)_{\rm high} \sim 3.4\,{\rm GeV}$, it is tempting to think that this structure may be caused by a new resonance. Even though it seems unlikely that such a high mass state could be narrow, and that there is no evident signal for such a state in the Dalitz plot distribution, I investigate the impact of adding an extra Breit–Wigner component, with various spin and parity, to the model as shown in Table 5.15.

The mass of the Breit–Wigner components is floated between 2.3 and $4.1\,{\rm GeV}$ and the width is floated between 25 and $700\,{\rm MeV}$. In the case that yields the largest $-2\Delta{\rm NLL}$ value, *i.e.* for $X1(\frac{7}{2}^+)$, the mass and width converge to $m_{X1(\frac{7}{2}^+)} = (3.21 \pm 0.05)\,{\rm GeV}$ and $\Gamma_{X1(\frac{7}{2}^+)} = (0.15 \pm 0.08)\,{\rm GeV}$, respectively. It can be seen from Table 5.15 that none of the components lead to a significant improvement in NLL value. Additionally, these values do not take into account the systematic uncertainty in the mis-modelling of the combinatorial background, which is the largest component at $m(pK^-)_{\rm high} \sim 3.7\,{\rm GeV}$. Therefore, the discrepancy seen in the projection plots in $m(pK^-)_{\rm high}$ is not supportive of an additional resonance. We believe this to be statistical fluctuation and is not investigated further.

The conclusion of this section is that I have established as my baseline a model that includes the $\Lambda(1520)$, $\Lambda(1405)$ and $\Lambda(1670)$ resonances as well as ${\rm NR}(\frac{1}{2}^+)$ and ${\rm NR}(\frac{3}{2}^-)$ components. I investigate, in the next section, the sensitivity that can be achieved on the CP violation observables with the current $\Xi_b^- \to pK^-K^-$ data sample.

Table 5.15: Impact of adding an extra Breit–Wigner component in $m(pK^-)_{\rm high}$ to the nominal model ("bottom-up"). The "Gaussian significance" is obtained by translating the $-2\Delta{\rm NLL}$ into a corresponding p-value for a χ^2 PDF with Δ DoF as the number of degrees of freedom. The quoted value is then evaluated as $\sqrt{2}\,{\rm erfc}^{-1}(p)$. This is expected to be an overestimate of the significance, since the mass and width of the resonance are floated. A proper evaluation would require detailed studies.

Components	Δ DoF	-2ΔNLL	Gaussian significance
$X1(\frac{1}{2}^-)$	6	14.1	2.2
$X1(\frac{1}{2}^+)$	6	10.1	1.5
$X1(\frac{3}{2}^-)$	6	13.7	2.1
$X1(\frac{3}{2}^+)$	6	13.8	2.1
$X1(\frac{5}{2}^-)$	6	13.3	2.0
$X1(\frac{5}{2}^+)$	6	12.2	1.8
$X1(\frac{7}{2}^-)$	6	14.4	2.2
$X1(\frac{7}{2}^+)$	6	16.5	2.5

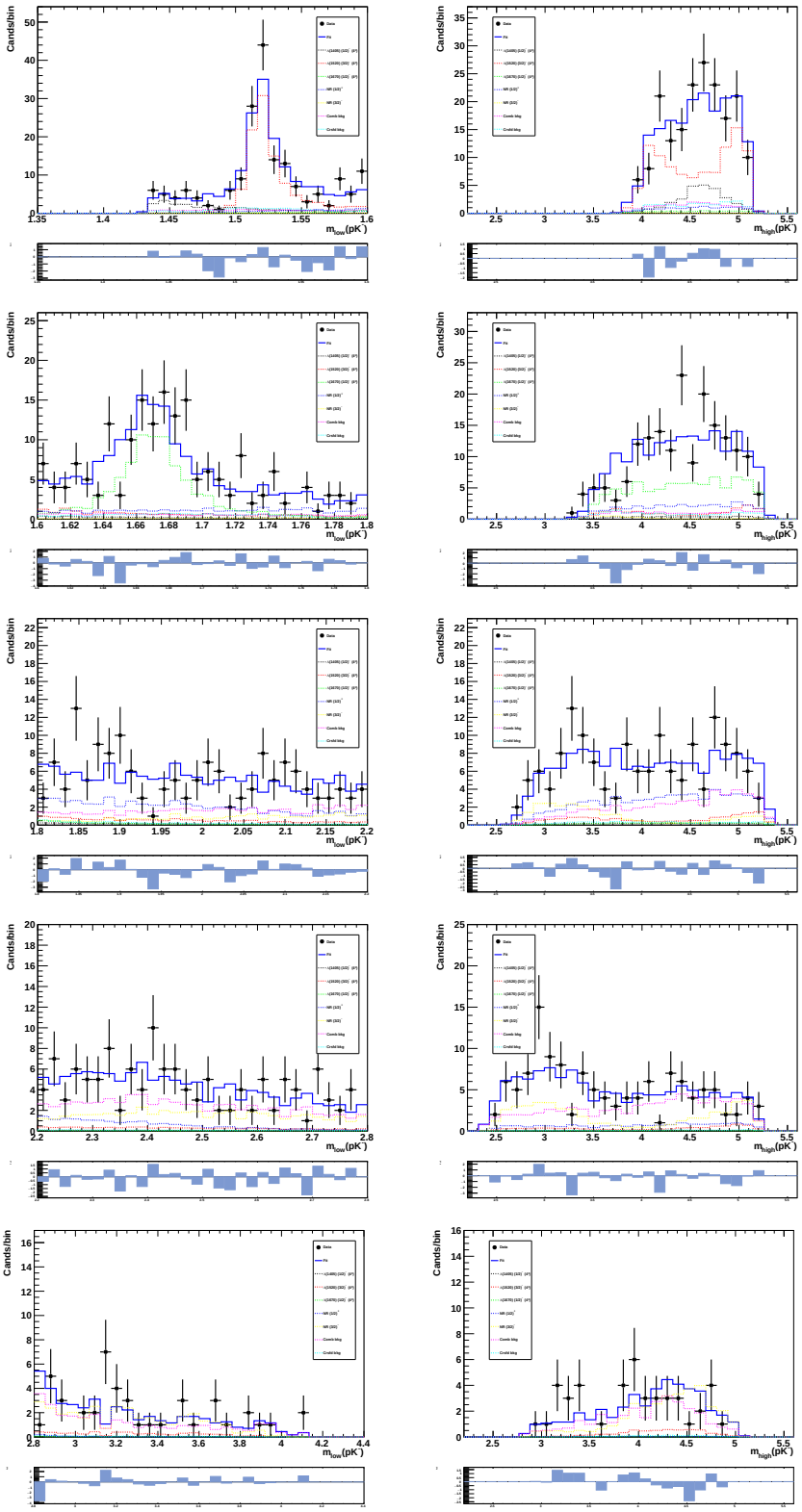

Figure 5.25: Projections of the result of the fit into $m(pK^-)_{\rm low}$ and $m(pK^-)_{\rm high}$, in slices of $m(pK^-)_{\rm low}$. From top to bottom row: $m(pK^-)_{\rm low} < 1.6\,{\rm GeV}$, $1.6\,{\rm GeV} < m(pK^-)_{\rm low} < 1.8\,{\rm GeV}$, $1.8\,{\rm GeV} < m(pK^-)_{\rm low} < 2.2\,{\rm GeV}$, $2.2\,{\rm GeV} < m(pK^-)_{\rm low} < 2.8\,{\rm GeV}$ and $m(pK^-)_{\rm low} > 2.8\,{\rm GeV}$.

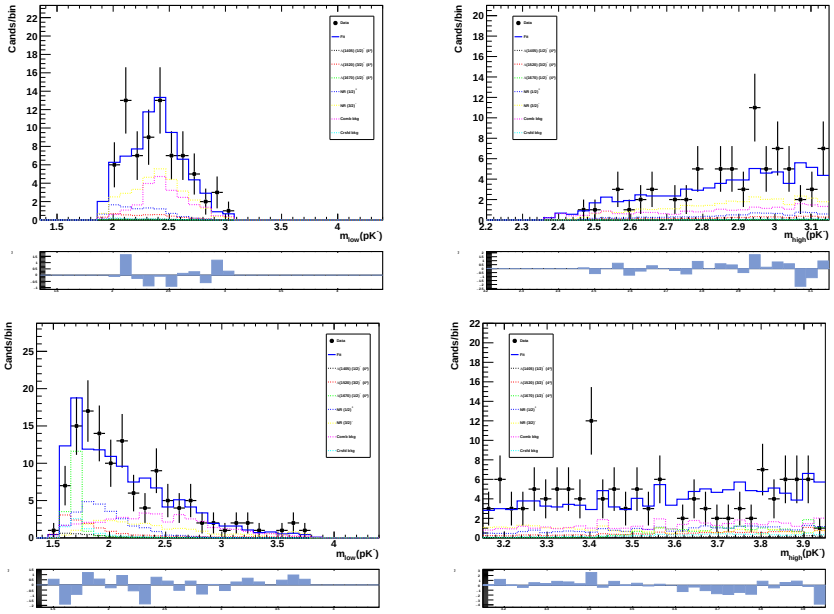

Figure 5.26: Projections of the result of the fit onto $m(pK^-)_{\rm low}$ and $m(pK^-)_{\rm high}$, in slices of $m(pK^-)_{\rm high}$. From top to bottom row: $m(pK^-)_{\rm high} < 3.15\,\text{GeV}$ and $3.15\,\text{GeV} < m(pK^-)_{\rm low} < 3.9\,\text{GeV}$.

5.5 Sensitivity to CP violation parameters

In the previous chapter, I established a baseline model that includes the $\Lambda(1520)$, $\Lambda(1405)$ and $\Lambda(1670)$ resonances as well as $\mathrm{NR}(\frac{1}{2}^+)$ and $\mathrm{NR}(\frac{3}{2}^-)$ components. This nominal model is parametrised by four CP conserving and 4 CP violation parameters associated to each of the resonant and nonresonant components, except for the reference $\Lambda(1520)$ resonance that is associated with only one CP violation parameter. It also has one shape parameter associated to each of the nonresonant components. Therefore, 17 CP violation parameters, 16 CP conserving parameters and 2 shape parameters (*i.e.* a total of 35 free parameters) are required to be inferred from the simultaneous fit to the $\Xi_b^- \to pK^-K^-$ and $\Xi_b^+ \to \overline{p}K^+K^+$ data samples.

In this section, I investigate the sensitivity that can be achieved in the measurement of the CP violation parameters associated to the components present in the nominal amplitude model. To do this, I generate 100 pseudo-experiments using Ξ_b and $\overline{\Xi}_b$ models with the efficiency and the background information discussed in the previous chapter. Each of the pseudo-experiments consists of the same amount of candidates as seen in Run I (2011–2012) and Run II (2015–2016) data sample. No implicit CP violation has been introduced (*i.e.* $\delta x^{\pm}, \delta y^{\pm} = 0$ in Eq. 5.17) in the generation of the pseudo-experiments. Simultaneous fits to each Ξ_b and $\overline{\Xi}_b$ toy sample have been conducted 20 times randomising the starting values of the free parameters to pick the best fit result. The fit to each of the pseudo-experiments has been conducted in two steps to facilitate stable and fast fit convergence by MINUIT. The first step involves fitting the sample by fixing the CP violation parameters to zero and floating only the CP conserving parameters. The second step involves floating both the CP violation and CP conserving parameters in the fit to the sample, where the initial starting values for CP conserving parameters are fixed to the results from step one.

The fit fractions (FFs) of the different resonance states obtained from the fit to Ξ_b^- and Ξ_b^+ toy samples are shown in Figure 5.27 and Figure 5.28, respectively. In these figures, I also show the sensitivity achieved on each of these observables. It can be seen from these fits that the model does not exhibit any large bias in the fit fractions.

The interference FFs of the different resonance states from the fit to the Ξ_b^- and Ξ_b^+ toy samples are shown in Figure 5.29 and Figure 5.30, respectively. These plots show, as expected, that there is significant net interference between components with same J^P (*e.g.* as seen in the interference FF of $\Lambda(1520)$ & $\mathrm{NR}(\frac{3}{2}^-)$). The components with different J^P quantum numbers should have interference FF consistent with zero due to the orthogonality relation satisfied by the Wigner-D matrix elements (*e.g.* as seen in interference FF plot of $\Lambda(1520)$ & $\Lambda(1670)$). Several biases are evident in the interference FF plots. The most prominent are those between the $\Lambda(1405)$ and $J^P = 3/2^-$ components (*viz.* $\Lambda(1520)$ and $\mathrm{NR}(\frac{3}{2}^-)$). These seem to be due to low

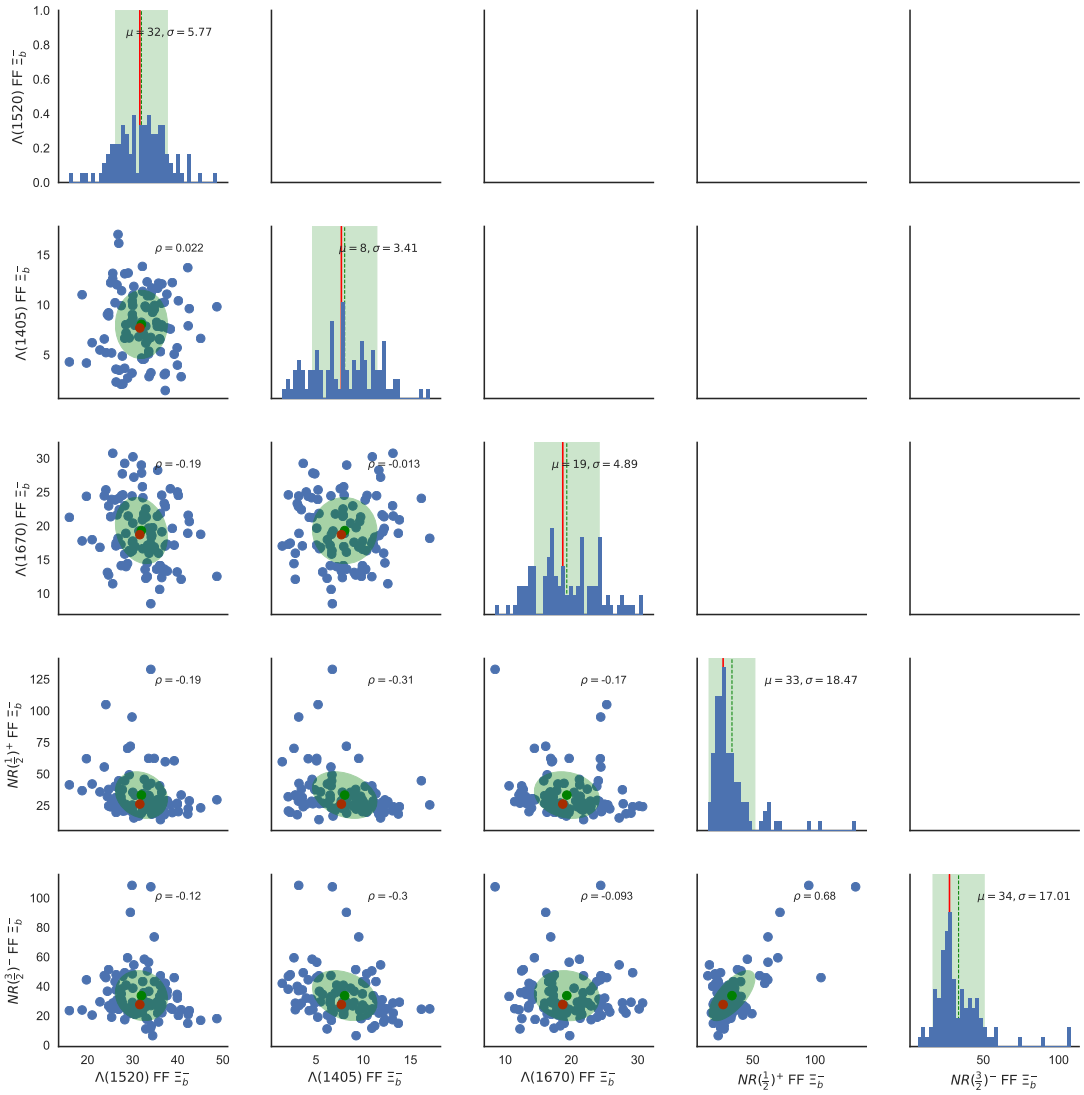

Figure 5.27: The correlation matrix of the fit fraction of different resonance components contributing in the Ξ_b^- toy samples. The red lines and red markers show the true value of the parameter. The green dotted lines and markers shows the mean of the sample. The dark green bands and ellipses show the 1σ contours, which indicates the spread of the results (the uncertainty on the mean is $\sqrt{N} = 10$ times smaller, where N is the number of pseudo-experiments in the ensemble).

statistics in the generated pseudo-experiments, since it is confirmed that they do not become more significant in high statistics pseudo-experiments (*i.e.* when I generate 100 times the observed data for each pseudo-experiment). Another interesting feature is the non-Gaussian distribution of the

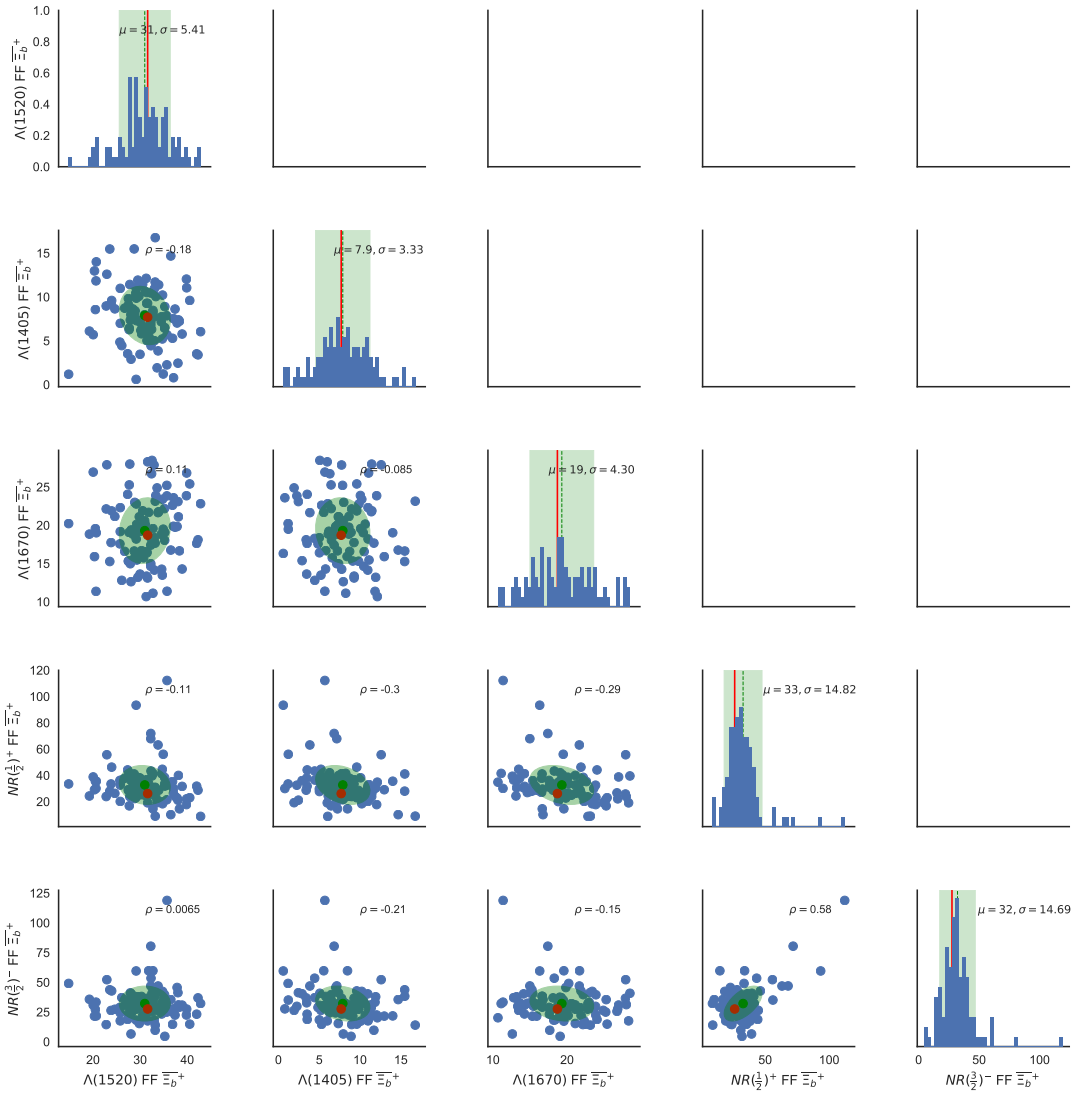

Figure 5.28: The correlation matrix of the fit fraction of different resonance components contributing in the Ξ_b^+ toy samples. The red lines and red markers show the true value of the parameter. The green dotted lines and markers shows the mean of the sample. The dark green bands and ellipses show the 1σ contours, which indicates the spread of the results (the uncertainty on the mean is $\sqrt{N} = 10$ times smaller, where N is the number of pseudo-experiments in the ensemble).

interference FF between the $\mathrm{NR}(\frac{1}{2}^+)$ and $\mathrm{NR}(\frac{3}{2}^-)$, and of some other interference FFs involving NR components. This occurs since the nonresonant shape parameters are floated and there is an instability related to a correlation between the shape parameters and the interference fit fractions.

Moreover the employed shapes for the nonresonant components are just effective descriptions of the underlying physics. Therefore I will be evaluating systematic uncertainty by choosing an alternate shape for the nonresonant component such as a polynomial lineshape instead of an exponential.

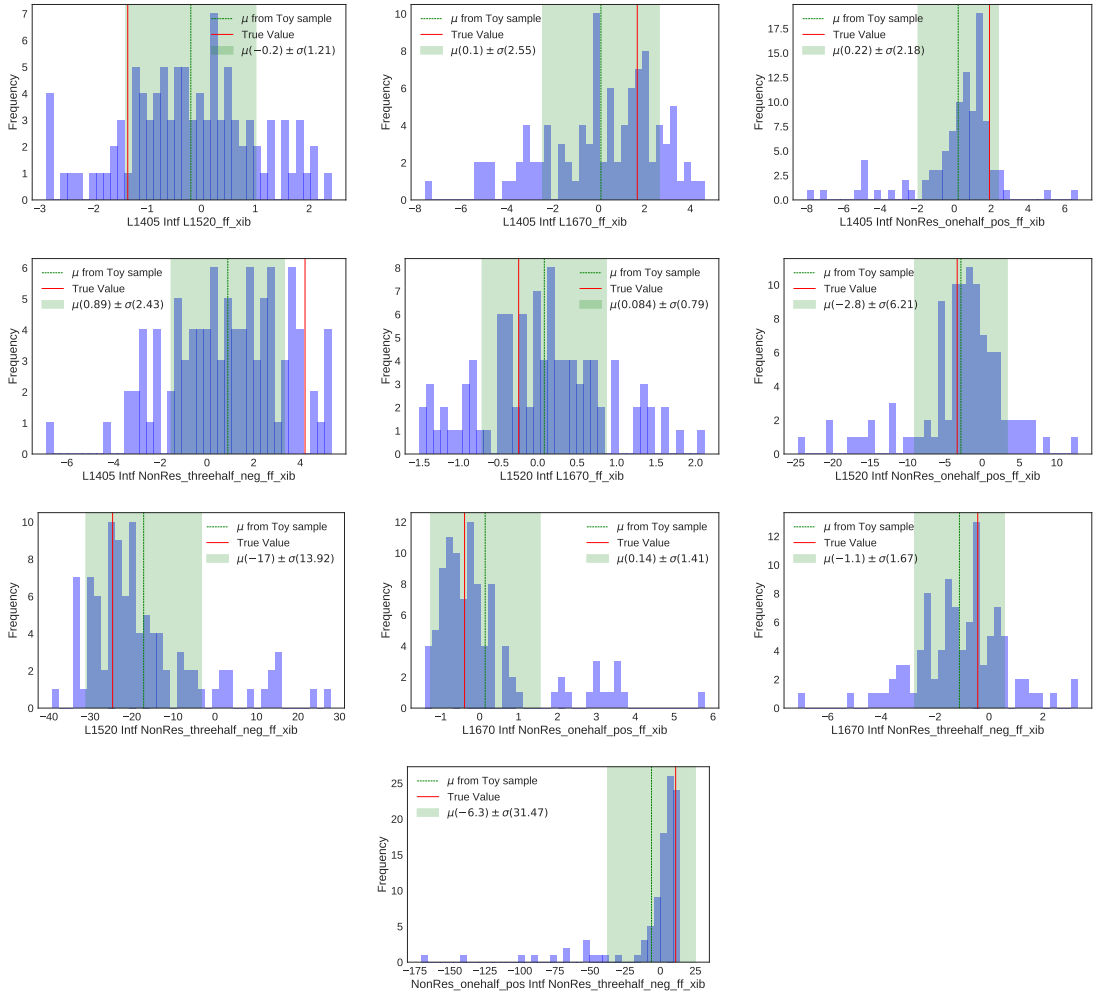

Figure 5.29: The interference fit fractions (%) between resonance components for Ξ_b samples.

The FF asymmetry (Eq. 5.31) and asymmetry in the magnitude of the helicity coupling, $A_{\pm}^{CP}$ (Eq. 5.19), for different components of the fit can be seen in Figures 5.31–5.35. Note that the asymmetry parameters can only take values between -1 to $+1$ (In the fit these parameters are only allowed to float between these two values too.). The sensitivity on each of these observables with the current data sample is also quoted on the figures. It can be seen from these plots that, with the current data sample, out of all the asymmetry observables the best sensitivity to asymmetry

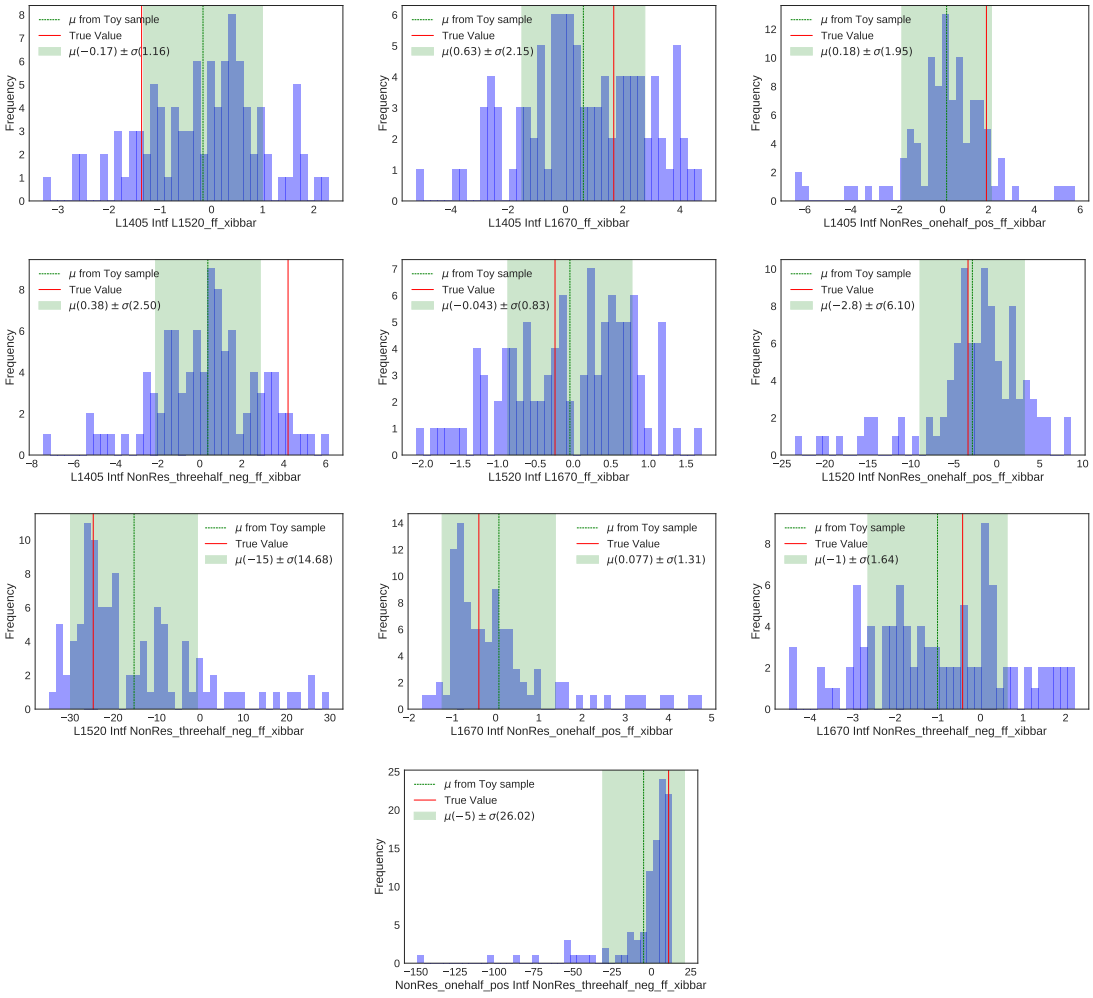

Figure 5.30: The interference fit fractions (%) between resonance components for $\overline{\Xi}_b$ samples.

observables related to $\Lambda(1520)$ is achieved. The analysis is also sensitive to the asymmetry in the fit fraction, in order of decreasing sensitivity, for $\mathrm{NR}(\frac{1}{2}^+)$, $\Lambda(1670)$, $\mathrm{NR}(\frac{3}{2}^-)$ and $\Lambda(1405)$ components. It is evident, with the current size of the data sample that, one cannot achieve a good sensitivity to the asymmetry parameters related to the helicity couplings of all other resonance components except for $\Lambda(1520)$. This can be attributed to the fact that the fit can determine CP violation in a component, but cannot determine in which helicity coupling it appears (except for $\Lambda(1520)$ where a^- is fixed to zero).

With the current data sample, the analysis is sensitive to the fit fractions of the components of the model only, with a 5 fb^{-1} of data sample collected by LHCb until 2016. LHCb will have collected around 9 fb^{-1} of data by the end of 2018. The experiment is then planned to undergo

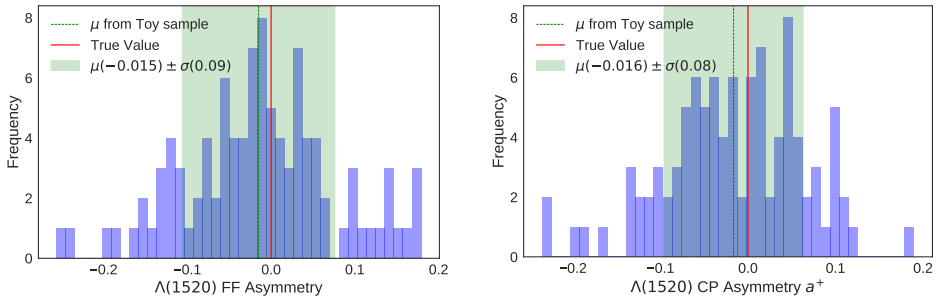

Figure 5.31: For $\Lambda(1520)$ resonance: (Left) fit fraction asymmetry and (right) CP-asymmetry in resonance coupling a^+.

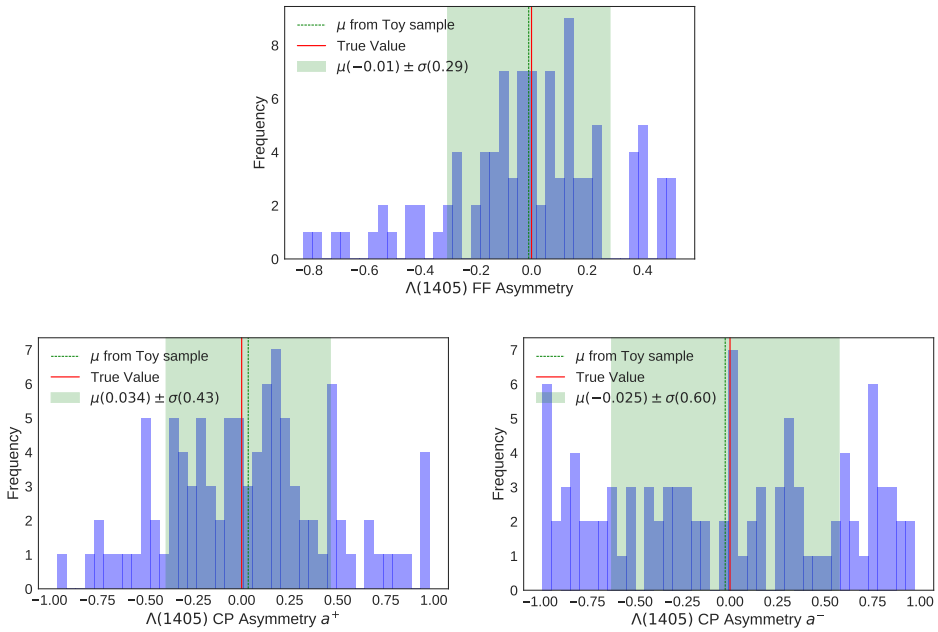

Figure 5.32: For $\Lambda(1405)$ resonance: (Top) fit fraction asymmetry, (bottom left) CP-asymmetry in resonance coupling a^+ and (bottom right) CP-asymmetry in resonance coupling a^-.

a phase I upgrade which will allow operation at luminosities of $2 \times 10^{33}\mathrm{cm}^2\mathrm{s}^{-1}$, *i.e.* ten times the current design luminosity (five times the current operational luminosity). After the phase I upgrade, LHCb plans to continue data taking and hopes to achieve an integrated luminosity of at least 50 fb^{-1} by 2029. As a result, the signal yield of $\Xi_b^- \to pK^-K^-$ will increase by as much as a factor of 10 (I ignore here any gain in the signal selection efficiency that would most definitely be achieved through an improved detector design and operation). The sensitivity on a given

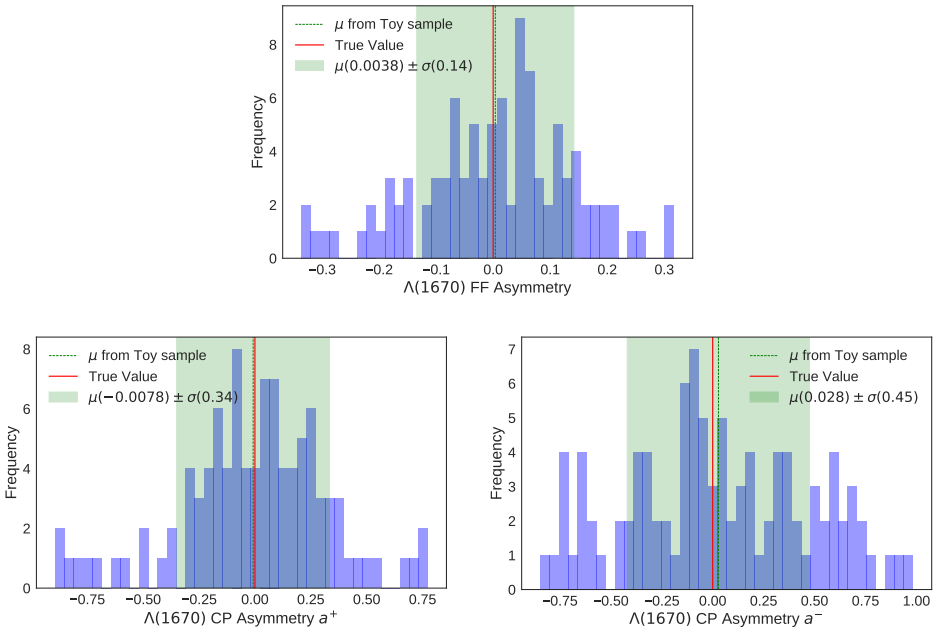

Figure 5.33: For $\Lambda(1670)$ resonance: (Top) fit fraction asymmetry, (bottom left) CP-asymmetry in resonance coupling a^+ and (bottom right) CP-asymmetry in resonance coupling a^-.

observable in this analysis roughly scales, ignoring correlation of the observable under study with other observables of the model, as $1/\sqrt{N_{\rm data}}$, where $N_{\rm data}$ is the size of the $\Xi_b^- \to pK^-K^-$ data sample. Therefore, by 2029 the statistical error on all the asymmetry observables should decrease by a factor $1/\sqrt{10}$ and sensitivity to all the asymmetry observables discussed in this section could be achieved.

5.6 Future work

This analysis is currently blind to the CP asymmetry observables. Unblinding will happen once the review committee appointed by the LHCb collaboration is satisfied with the procedure employed in this analysis. After unblinding the main results of the analysis are expected to be dominated by statistical uncertainties, which have already been evaluated in the previous section. However, inaccuracy in the experimental inputs used in the fit and the choice of the nominal parameterisation can, in principle, introduce significant systematic uncertainties. Therefore, these systematic uncertainties must be evaluated and will be quoted along with the statistical uncertainty on the results of the analysis after unblinding.

To evaluate systematic uncertainty on the update of $\mathcal{R}$, presented in Sec. 5.3.1, I follow

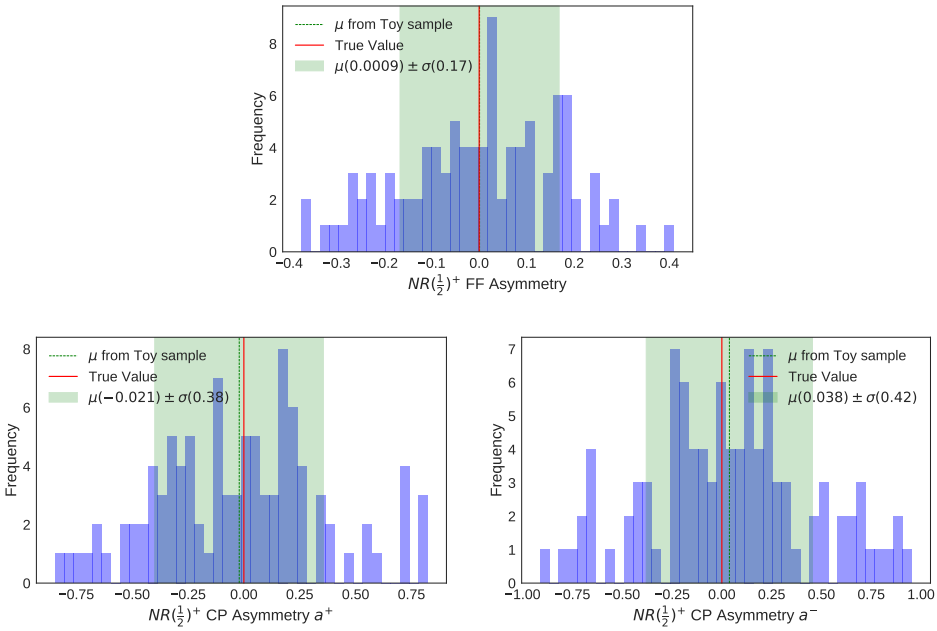

Figure 5.34: For NR($\frac{1}{2}^+$) resonance: (Top) fit fraction asymmetry, (bottom left) CP-asymmetry in resonance coupling a^+ and (bottom right) CP-asymmetry in resonance coupling a^-.

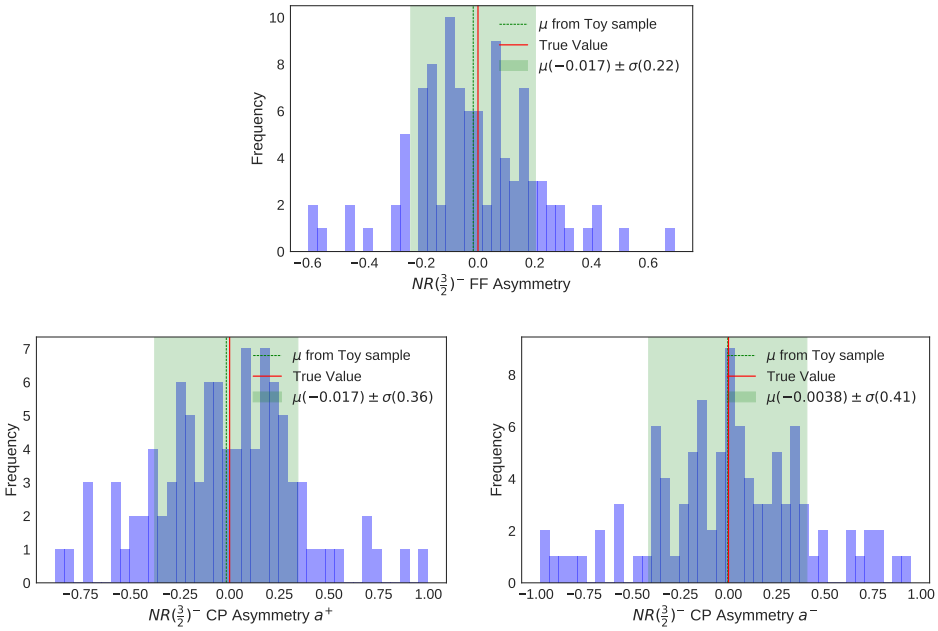

Figure 5.35: For NR($\frac{3}{2}^-$) resonance: (Top) fit fraction asymmetry, (bottom left) CP-asymmetry in resonance coupling a^+ and (bottom right) CP-asymmetry in resonance coupling a^-.

the same method as employed in the previous chapter in Sec. 4.6 and therefore the details are not discussed here. Below I discuss the sources and evaluation of systematic uncertainties related to the measurement of the asymmetry observables (defined in Eq. 5.19 and 5.31) obtained through amplitude analysis.

Mass fits: The values of the signal and background fit fractions are obtained from the fit to $m(pKK)$ and are used in the PDF definitions in Eq. 5.6 and 5.8. Since the estimated fractions are dominated by statistical uncertainty, the systematic uncertainty on the asymmetry observables originating from the $m(pKK)$ fit model will be evaluated by varying the yields using the covariance matrix obtained from the fit and repeating the amplitude fit.

Background shapes: The distribution of the combinatorial background across the phase space is obtained by extrapolating from the sideband region. This method introduces uncertainties related to the available yield in the sideband and the extrapolation procedure itself. Repeating the amplitude fit while varying the background shape within these uncertainties will allow the associated systematic uncertainty to be evaluated. For $\Xi_b^- \to pK^-\pi^-$ cross-feed background, the shape will be altered by assuming a different cocktail model as discussed in Sec. 5.4.3.

Background asymmetry: In the baseline fit, it is assumed that there is no asymmetry in the combinatorial background. The associated systematic uncertainty will be evaluated by allowing separate background PDFs for Ξ_b^- and Ξ_b^+, as discussed in Sec. 5.4.2. Global asymmetries in the combinatorial background will be inferred directly from the fit, whereas for the crossfeed background the asymmetry is fixed to zero in the nominal fit but will be varied within $\pm 20\%$ for the evaluation of systematic uncertainty.

Selection efficiency maps: I use as default separate efficiency maps for Ξ_b^- and Ξ_b^+ samples in order to account directly for possible detection asymmetry. The shapes of these efficiency maps will be altered to evaluate systematic uncertainties, as done in the previous chapter in Sec. 4.6. After considering all sources of systematic uncertainty on the Ξ_b^- and Ξ_b^+ efficiency maps, I generate a set of such maps where each one is created by sampling, for each bin in the map, from a normal distribution with a mean equal to the bin content and standard deviation equal to the associated uncertainty in that bin. Fits will be conducted using each of these efficiency maps and the RMS of the distribution of the change in each fitted parameter is taken as the systematic uncertainty.

Production asymmetry: The Ξ_b^- production asymmetry has not yet been measured, and there is a lack of a convenient control mode that could be used to remove this source of asymmetry in the present analysis. Therefore, as a guide, the LHCb measurement of the Λ_b^0 production asymmetry [164] is taken, which is consistent with being zero:

$$\begin{aligned} A_P(\Lambda_b^0)_{\sqrt{s}=7\,\text{TeV}} &= 0.0011 \pm 0.0253\,(\text{stat}) \pm 0.0108\,(\text{syst})\,, \\ A_P(\Lambda_b^0)_{\sqrt{s}=8\,\text{TeV}} &= 0.0344 \pm 0.0161\,(\text{stat}) \pm 0.0076\,(\text{syst})\,. \end{aligned}$$

No measurement exists for Run II (2015–2016). Since the magnitude of the Ξ_b^- production asymmetry should be similar to that for the Λ_b^0 baryon, I take the production asymmetry to be zero in the baseline, for both Run I (2011–2012) and Run II (2015–2016). I vary the production asymmetry by $\pm 5\%$, which is considered conservative, to evaluate the systematic uncertainty due to possible production asymmetry. This is implemented by introducing a global asymmetry between the Ξ_b^- and Ξ_b^+ efficiency maps, which is indistinguishable from a production asymmetry in the analysis.

Polarisation: I have assumed that the transverse polarisation of the Ξ_b^- baryons produced in pp collisions is consistent with zero as observed in the case for Λ_b^0 baryons. To assign a systematic due to this assumption, I perform pseudo-experiments in which signal is generated with non-zero polarisation, and then passed through the fit procedure neglecting this effect. The resulting biases on the fitted parameters will be assigned as the associated systematic uncertainties.

Breit–Wigner parameterisation: Each resonant contribution has fixed parameters in the amplitude fit. These include masses and widths and Blatt–Weisskopf radial parameters. For the $\Lambda(1520)$ and $\Lambda(1670)$ resonances, the fit is repeated many times varying their masses and widths within the range of values quoted by the PDG (see Table 5.10). The Blatt–Weisskopf radius parameters $d_{\Xi_b^-}$ and d_{Res} are varied between 0–10 GeV^{-1} and 0–3 GeV^{-1} respectively. The RMS of the distribution of the change in each fitted parameter is taken as the systematic uncertainty. For the $\Lambda(1405)$ resonance, the lineshape is replaced with a Flatté parameterisation as used in Ref. [21].

Modelling of nonresonant shapes: Alternative descriptions of the NR shapes will be used to assign systematic uncertainties.

Alternative fit model: The effect of including additional signal components in the fit model will be examined to assign systematic uncertainty due to the composition of the baseline model.

In particular the $\Lambda(1600)$ and $\Lambda(1890)$ resonances, which are preferred in the LASSO approach to building the model, will be included.

In summary, I have presented in this chapter details of an amplitude analysis of the $\Xi_b^- \to pK^-K^-$ decay mode using a data sample on pp collision accumulated by LHCb over 2011–2016, comprising of $1\,\text{fb}^{-1}$ collected at $\sqrt{s} = 7\,\text{TeV}$ (2011), $2\,\text{fb}^{-1}$ collected at $\sqrt{s} = 8\,\text{TeV}$ (2012) and $2\,\text{fb}^{-1}$ collected at $\sqrt{s} = 13\,\text{TeV}$ (2015–16). I have highlighted the details of the procedure employed to improve the efficiency of signal selection of $\Xi_b^- \to pK^-K^-$ decays in Run I (2011–2012) by a factor 3.4 as compared to the branching fraction analysis presented in the previous chapter. I have also updated in this analysis the ratio of production fractions and branching fractions $\left(f_{\Omega_b^-}/f_{\Xi_b^-}\right) \times \left(\mathcal{B}\left(\Omega_b^- \to pK^-K^-\right)/\mathcal{B}\left(\Xi_b^- \to pK^-K^-\right)\right)$, which is consistent with, and more precise than, the measurement in the previous chapter. To establish a baseline amplitude model, I have investigated three separate methods, "top-down", "bottom-up" and LASSO regularisation. The model obtained from the "bottom-up" approach was chosen as my baseline model which includes the $\Lambda(1520)$, $\Lambda(1405)$ and $\Lambda(1670)$ resonances as well as $\text{NR}(\frac{1}{2}^+)$ and $\text{NR}(\frac{3}{2}^-)$ components. Using this model, the sensitivity that can be achieved on *CP* violation parameters was investigated. It is concluded from this study that, with the present LHCb data sample, only sensitivity to the fit fractions, their asymmetry and interference fit fractions can be achieved. After unblinding, it is possible that a significant *CP* violation will be observed, if the true physical parameters correspond to large *CP* violation in, *e.g.* $\Xi_b^- \to \Lambda(1520)K^-$ decays. We have also shown that analyses with larger data samples will allow *CP* violation in the helicity couplings to be probed.

CHAPTER 6

Summary

“ An end is only a beginning in disguise. ”

Craig D. Lounsbrough, *An Intimate Collision*

In the present thesis, I have investigated the data sample collected by the LHCb experiment to study three-body decays of the Ξ_b^- and Ω_b^- baryons to charmless final states $ph^-h'^-$, where $h^{(\prime)}$ denotes a kaon or pion.

Using the $3\,\text{fb}^{-1}$ of data collected by LHCb during the years 2011 and 2012, the decay $\Xi_b^- \to pK^-K^-$ has been observed for the first time at significance level of 8.7 Gaussian standard deviations. This data sample yields around 80 $\Xi_b^- \to pK^-K^-$ signal candidates. I also have first evidence of the decay of $\Xi_b^- \to pK^-\pi^-$ at a significance level of 3.4 standard deviations. No evidence of the $\Xi_b^- \to p\pi^-\pi^-$ decay, or of decays of the heavier b-baryon Ω_b^- to any $ph^-h'^-$ final state, are found. I have conducted the measurement of the products of branching fractions and b-hadron production fractions relative to the $B^- \to K^+K^-K^-$ normalisation channel. These measurements are reported in Table 6.1. Since the signals for $\Xi_b^- \to p\pi^-\pi^-$, $\Omega_b^- \to pK^-K^-$, $\Omega_b^- \to pK^-\pi^-$ and $\Omega_b^- \to p\pi^-\pi^-$ are not significant, I also show in Table 6.1 upper limits for these decay modes. The branching fractions of $\Xi_b^- \to pK^-\pi^-$ relative to $\Xi_b^- \to pK^-K^-$ decays is also measured.

$$\frac{\mathcal{B}(\Xi_b^- \to pK^-\pi^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} \;=\; (\,977 \pm 274\,(\text{stat}) \pm 85\,(\text{syst})\,) \times 10^{-3}\,.$$

I also quote the upper limits at 90 (95) % confidence level for the relative branching fraction of

Table 6.1: Relative branching fractions multiplied by fragmentation fractions ($R_{ph^-h'^-}$). The two uncertainties quoted on $R_{ph^-h'^-}$ are statistical and systematic. Upper limits are quoted at 90 (95) % confidence level for modes with signal significance less than $3\,\sigma$.

Mode	$\frac{\Xi_b^-}{f_u} \times \frac{\mathcal{B}(\Xi_b^-(\Omega_b^-)\to phh'^-)}{\mathcal{B}(B^+\to K^+K^-K^+)}$ (10^{-5})
$\Xi_b^- \to pK^-K^-$	$265 \pm 35 \pm 47$
$\Xi_b^- \to pK^-\pi^-$	$259 \pm 64 \pm 49$
$\Xi_b^- \to p\pi^-\pi^-$	$74 \pm 40 \pm 36 \quad < 147 \ (166)$
$\Omega_b^- \to pK^-K^-$	$-9 \pm 9 \pm 6 \quad < 18 \ (22)$
$\Omega_b^- \to pK^-\pi^-$	$-23 \pm 28 \pm 23 \quad < 51 \ (62)$
$\Omega_b^- \to p\pi^-\pi^-$	$48 \pm 33 \pm 28 \quad < 109 \ (124)$

$\Xi_b^- \to p\pi^-\pi^-$ with respect to $\Xi_b^- \to pK^-K^-$

$$\frac{\mathcal{B}(\Xi_b^- \to p\pi^-\pi^-)}{\mathcal{B}(\Xi_b^- \to pK^-K^-)} \quad < \quad 554 \ (626) \times 10^{-3}\,.$$

Improving the signal selection using the Run I (2011–2012) data collected by LHCb and also adding the Run II (2015–2016) data, I have conducted an amplitude analysis of $\Xi_b^- \to pK^-K^-$ decay. The amplitude analysis led to a model which includes $\Lambda(1520)$, $\Lambda(1405)$ and $\Lambda(1670)$ resonances as well as $\mathrm{NR}(\frac{1}{2}^+)$ and $\mathrm{NR}(\frac{3}{2}^-)$ nonresonant components. As this analysis is currently blind to the *CP* violation parameters and is under review by the LHCb collaboration, I have used this baseline model, accounting for efficiency variation and background contribution, to investigate the sensitivity to the *CP* violation parameters of the model. It is concluded from this study that, with the present LHCb data sample, only sensitivity to the fit fractions, their asymmetry and the interference fit fractions can be achieved.

The LHCb experiment is planned to undergo an upgrade during long shutdown 2 (2019–2020) that will allow operation at luminosities of $2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. After the upgrade, LHCb plans to continue data taking and hopes to achieve an integrated luminosity of at least $50\,\mathrm{fb}^{-1}$ by 2029. As a result, the signal yield of $\Xi_b^- \to pK^-K^-$ might well increase by a factor of 10. Therefore, by 2029 the statistical error on all the observables could reduce by a factor $1/\sqrt{10}$, which would allow *CP* violation in the helicity couplings to be probed.

To date there is no experimental observation of *CP* violation in b baryon decays. Significant progress has however been achieved towards such an observation. The observation of a Ξ_b^- decay to a charmless three-body final state (pK^-K^-) and the development of an amplitude analysis tools that can be used to investigate *CP* violation in this and similar channels are

important steps towards a better understanding of matter-antimatter asymmetries in the baryon sector. As more data is accumulated by the LHCb experiment, it is expected that discoveries will be possible.

APPENDIX A

VELO module and its support system

During the LHCb phase I upgrade over the years 2019 to 2020, the VELO sub-detector (currently a silicon strip detector with R-ϕ sensors) will be replaced by a hybrid pixel detector having pixels of size $55\,\mu\mathrm{m} \times 55\,\mu\mathrm{m}$. Pixels are an attractive choice for the upgraded VELO due to the high granularity and fast pattern recognition, allowing excellent vertex resolution and tracks separation that are crucial for the trigger performance. In the upgraded VELO sub-detector, there will be 26 pairs of VELO modules which will be arranged around the Interaction Point, with the help of a support system such that they cover the entire LHCb pseudo-rapidity range ($1.9 < \eta < 4.9$) to provide excellent tracking and vertexing measurements. The Oxford and Nikhef LHCb groups have put forward their respective designs for the support system for the VELO module, which are referred to throughout this note as the "Oxford" and "Nikhef" designs. This appendix describes the detector geometry of the VELO module and its support system. The details presented here have been documented in a LHCb internal note [165]. The work presented in this note is a combined effort of all authors of the note.

In Sec. A.1, I introduce the details and the methods used in the implementation of the geometry descriptions of the VELO module and its support system. In Sec. A.2, I discuss the actual geometry descriptions of different components that form the VELO module. In Secs. A.3 & A.4, I discuss respectively the designs of the VELO module support system proposed by the Nikhef and Oxford groups. A study of the amount of simulated material in the VELO is presented in Sec. A.5. I also highlight in this section, the impact of the two designs on the performance of several key physics quantities, studied using fully simulated events that were generated with the respective VELO detector descriptions. Finally, a summary of the appendix is presented in Sec. A.6.

A.1 Introduction

Information on the detector structure, geometry and materials is stored in the detector description database (DDDB) in XML (eXtendable Markup Language) files. The DDDB forms part of the more general conditions database (CondDB) [166], which contains additional information on detector conditions, *e.g.* temperature and pressure around the detector during the data taking period. The DDDB is used by `Geant4` [93, 167] to simulate the passage of particles through the LHCb detector. The LHCb tracking code also uses the DDDB information to account for multiple scattering from particle interactions in the material.

The full description of the VELO geometry is split into four parts: geometry, structure, parameters and materials. The geometry part describes the physical aspects of the detector and is the main concern for this appendix. The geometry part of the description involves the use of two types of volumes viz. logical volumes and physical volumes. Logical volumes are the basis for the whole geometry description and are created from seven primitive types of solid lattices namely boxes, simple trapezoids, general trapezoids, tube segments, sphere segments, conical tube segments and polyconical tube segments. These primitive shapes may be combined using union, subtraction and intersection Boolean operations. The logical volumes can also hold daughter volumes within them. The daughter volumes are implemented as physical volumes, which are specific instantiations of a logical volume. These physical volumes can be positioned and rotated to a particular location and orientation inside the mother logical volume. All the volumes are visualised using Panoramix [168], the LHCb event display. This allows individual elements of the detector description to be viewed in three dimensions.

The logical volume that contains the geometry description of the VELO module mounted on the support structure is labelled `lvModule{X}WithSupport`, where `{X} ∈ {0, 51}`. Here `{X}` represents the number of VELO modules in the upgraded VELO sub-detector. The volumes describing the VELO module and its support system are labelled `lvModule{X}` and `lvSupport`, respectively. Figure A.1 shows the two `lvModule{X}WithSupport` for Oxford and Nikhef designs. In the following sections, I will discuss the components of the logical volume `lvModule{X}WithSupport` in detail.

A.2 VELO module

The VELO comprises 52 modules, each mounted on the support structure. All the components of the VELO module are included in the mother volume, `lvModule{X}`, which consists of an `lvHybrid` and four `lvLadder{X}_{0,1,2,3}` logical volumes. The `lvHybrid` is a hybrid circuit that consists of data connectors, layers of kapton–copper and cooling substrate. The

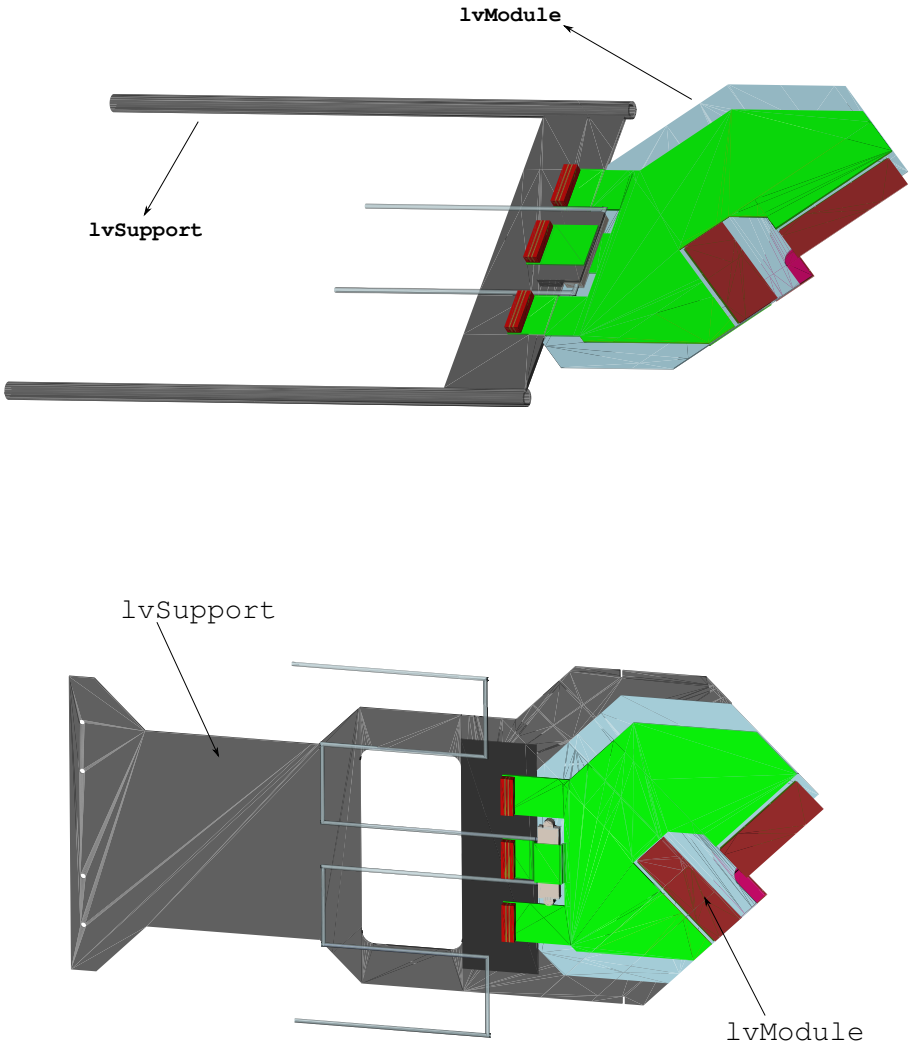

Figure A.1: The VELO module with the (top) Nikhef and (bottom) Oxford support structures.

lvLadder{X}_{0,1,2,3} is a tile that consists of three silicon VeloPix ASICs and a silicon sensor. The volumes included in lvModule{X} are illustrated in Figure A.2 and are discussed in detail in this section.

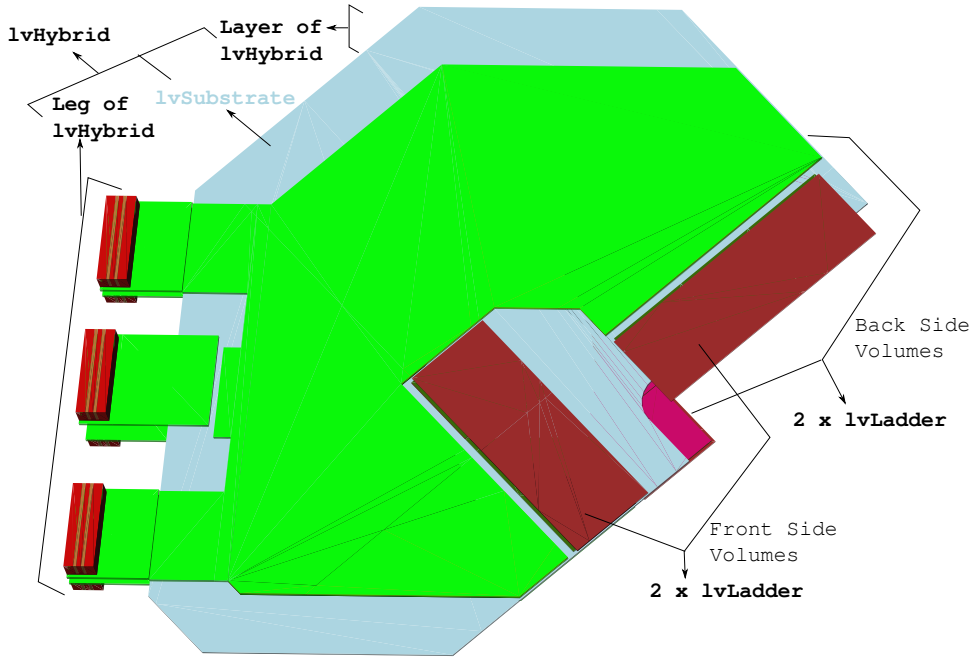

Figure A.2: The VELO Module 'lvModule{X}'.

VELO Tile: The lvLadder{X}_{0,1,2,3} is a mother volume that forms a tile and consists of an lvChips and an lvSensor{X}_{0,1,2,3} logical volumes. The lvChips is itself another mother volume that contains three lvChip and three lvGlue logical volumes, forming a chip array. The lvChip represents a single silicon VeloPix ASIC and the lvGlue represents a Stycast (a mixture of carbon, hydrogen and oxygen) glue layer between each ASIC and the substrate. The chip array is mounted to lvSensor{X}_{0,1,2,3}, which comprises a single logical volume called lvDet. This lvDet consits of silicon and is a subtraction of 4 boxes from a parent box that results in an active area surrounded by the inactive edges. The volumes included in lvLadder{X}_{0,1,2,3} are illustrated in Figure A.3.

Hybrid: The lvHybrid is a complex mother volume containing the four silicon tiles. This element consists of three different parts: the lvSubstrate, the main part of the lvHybrid and the "legs" of the lvHybrid. In addition, the hybrid also includes six data connector volumes,

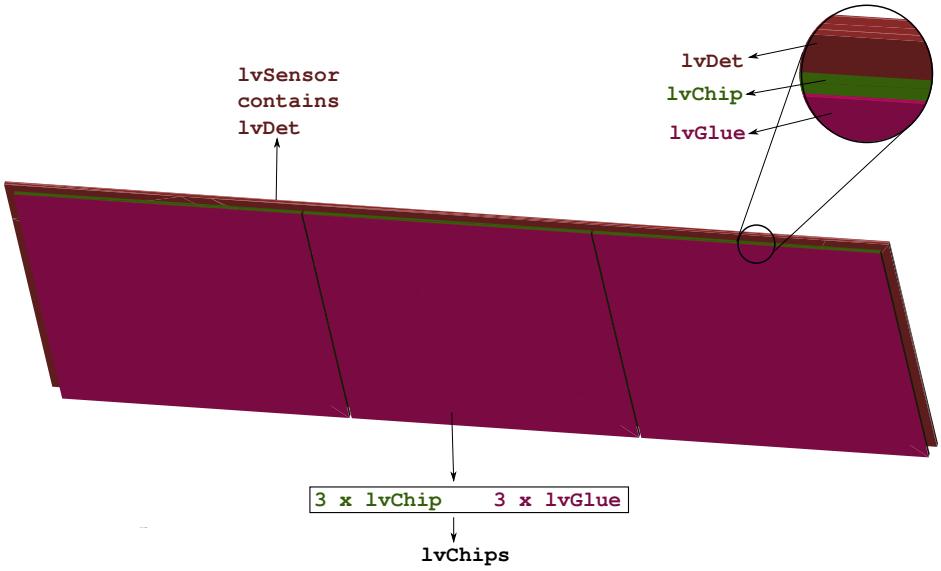

Figure A.3: Components of the `lvLadder{X}_{0,1,2,3}` logical volume which describes the VELO tile.

`lvDataConnector`, three on the front and three on the back of the mid-plate of the VELO support module, sitting on-top of the kapton-copper-kapton layers.

An `lvDataConnector` is represented as a box-shaped mother volume that defines a data connector, placed on all legs of the hybrid. Each data connector, `lvDataConnector`, is a mother volume built out of a `lvDataConnectorLCP` and two `lvDataConnectorContact` logical volumes. The `lvDataConnectorContact` is consists of Liquid Crystal Polymer (LCP) material and described as the subtraction of 2 boxes from the parent box to accommodate the two `lvDataConnectorLCP` volumes. The `lvDataConnectorLCP` logical volume is consists of beryllium-copper and is just a primitive type box. All the volumes that form `lvDataConnector` are illustrated in Figure A.4.

A.3 Oxford design

The full Oxford support design is contained within the logical volume `lvSupport` and consists of four components: `lvOxfSupport`, which is the main body of the support; `lvManifold`, which is an interface between the module section and the cooling pipes; `lvReinforce`, which is a block used to hold the manifold in place and attached on the opposite side of the module

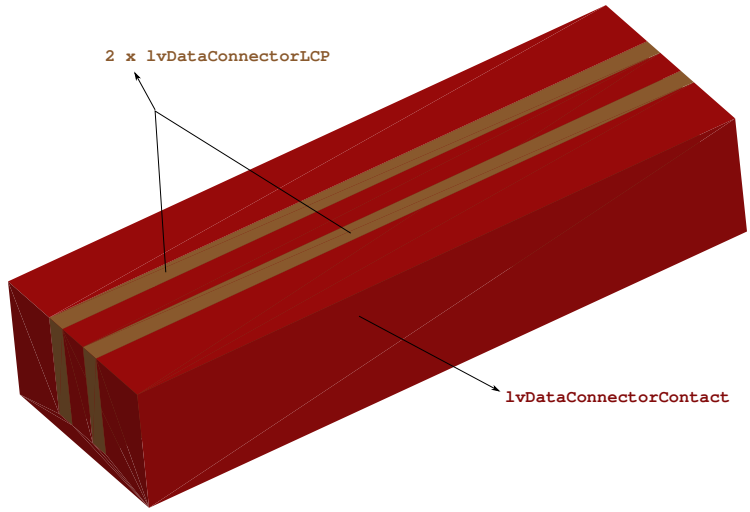

Figure A.4: The logical volumes that constitute `lvDataConnector`.

and lvDeliveryPipe, which is a logical volume describing the two cooling pipes that are connected to the manifold. I will discuss in this section the logical volumes lvManifold, lvReinforce and lvDeliveryPipe in detail.

Cooling connector and reinforcement block: The cooling connector 'lvManifold' and reinforcement block 'lvReinforce' are two separate pieces which are of the same general shape but of different thicknesses. Additionally, the cooling connector has a hollow middle section with cylindrical holes of slightly larger diameter in the side than the cooling pipes. The cooling connector is formed of Invar whereas the reinforcement block is formed of ceramic. These two pieces are mounted either side of the main module and the cooling pipes connect into the side of the cooling connector.

The lvManifold volume consists of three separate pieces, lvManifoldMiddle and two instances of lvManifoldSide. Figure A.5 shows the full lvManifold volume. The lvReinforce is a single logical volume and is shown in Figure A.6.

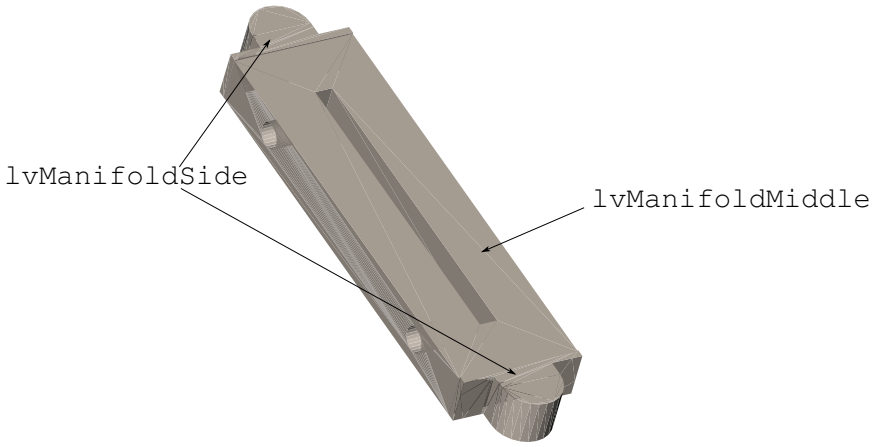

Figure A.5: The cooling connector joining the cooling pipes to the hybrid 'lvManifold'.

Cooling pipes: The two cooling pipes, that carry CO_2 coolant, in the support design are both described by a single logical volume, lvDeliveryPipe. The logical volume as implemented is shown in Figure A.7.

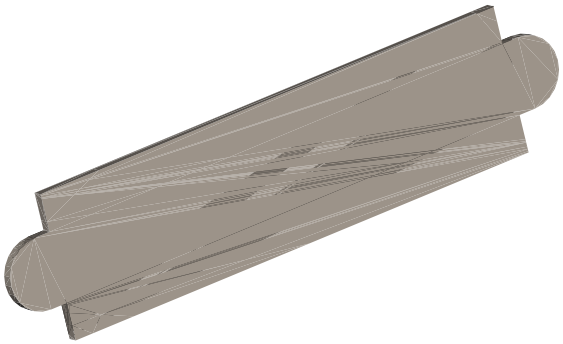

Figure A.6: Reinforcing block used on the opposite side of the hybrid to the cooling connector ‘lvReinforce’.

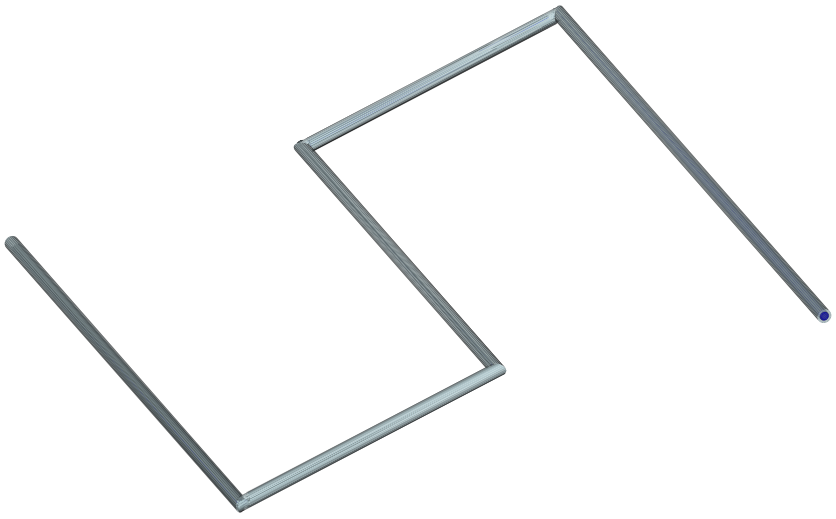

Figure A.7: The coolant delivery pipes for the modules ‘lvDeliveryPipe’.

A.4 Nikhef design

The Nikhef support design consists of a basic support structure (`lvNikhefSupport`) formed of carbon fibre, two CO_2 delivery pipes (`lvDeliveryPipe`) which help in the transport of the CO_2 coolant, and a CO_2 manifold (`lvManifold`) to carry the coolant from delivery pipes to the VELO module. The volumes included in `lvSupport` are shown in Figure A.8. In this section, I describe all the logical volumes that make up the Nikhef version of `lvSupport` in detail.

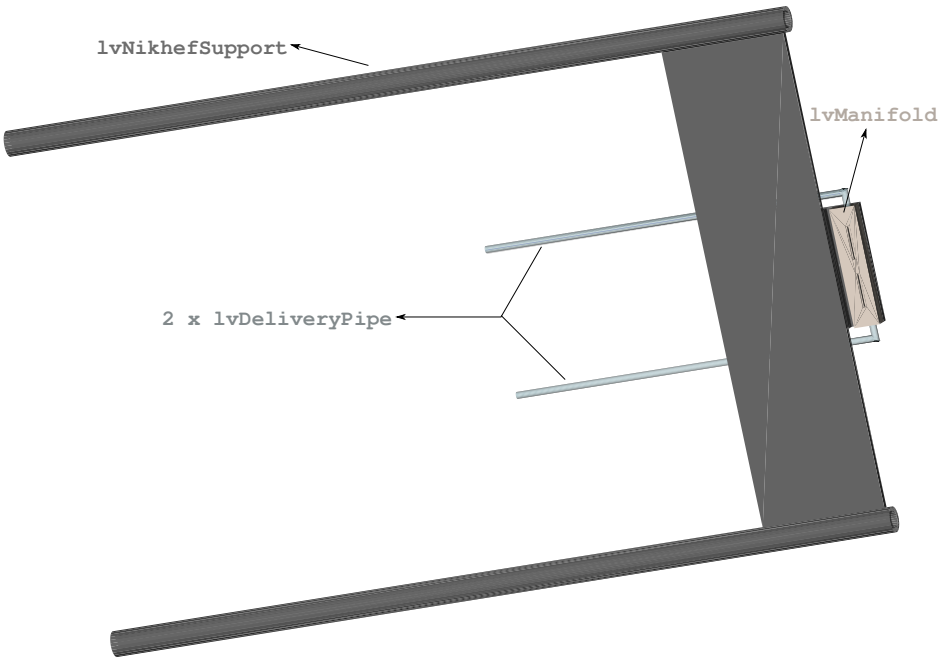

Figure A.8: The Nikhef module support '`lvSupport`'.

Nikhef support: The `lvNikhefSupport` logical volume consists of a mid-plate (`lvSupportMidPlate`), which acts as a floor for data connectors to sit on, two cylindrical legs (`lvSupportPipe`) on either side of the mid-plate, a small plate (`lvSupportPlate1`) and a slightly larger plate (`lvSupportPlate3`). Together with the delivery pipe and hybrid, they hold the manifold in place. Figure A.9 shows the `lvNikhefSupport` and all of its components that are formed of carbon fibre.

Cooling connector: The `lvManifold`, illustrated in Figure A.10, is a logical volume constructed from a primitive box-type solid after carrying out subtractions on one of its faces of two

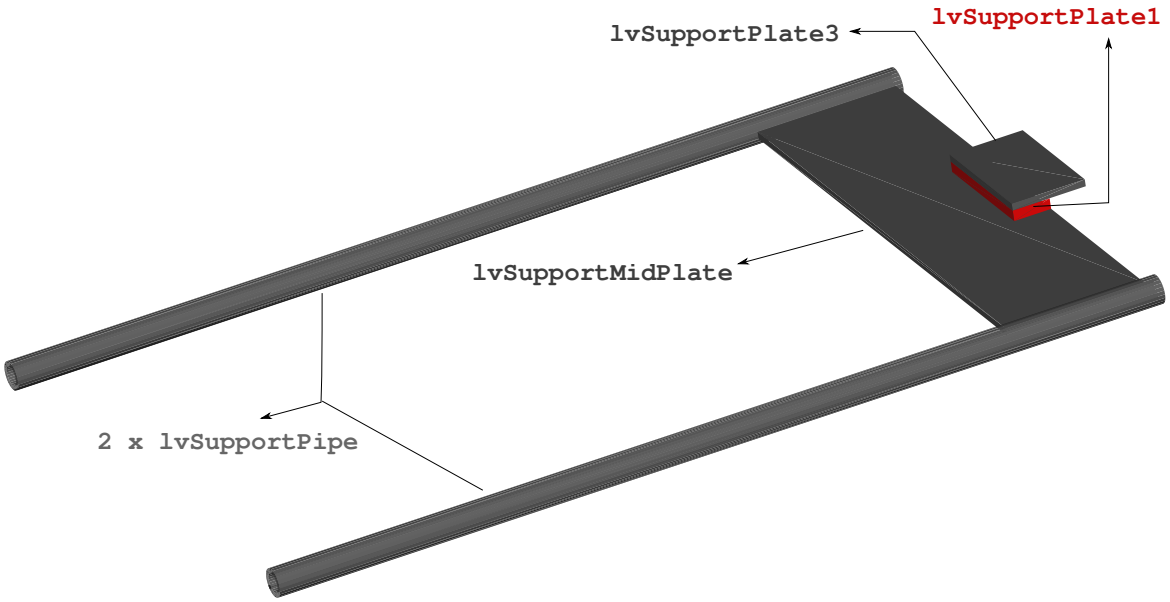

Figure A.9: The basic structure of the Nikhef module support '`lvNikhefSupport`'.

pairs of big and small boxes, to provide connections for the CO_2 coolant, and subtractions on both of its sides of big and small solid tubes, to provide inlets for the coolant. This volume is formed of Invar.

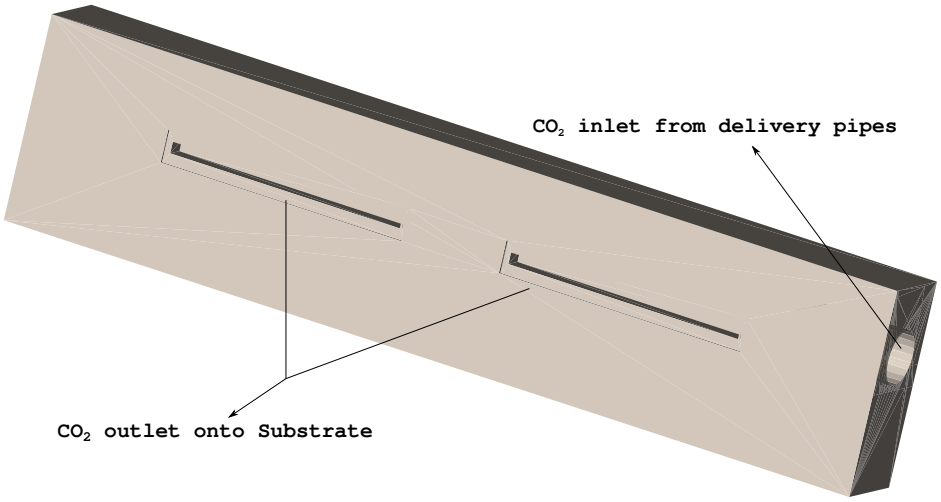

Figure A.10: The CO_2 manifold '`lvManifold`'.

Cooling pipes: The `lvDeliveryPipe` is a logical volume which is in an "L" shape consisting of shorter and longer hollow tubes filled with CO_2. The tubes ('`lvDeliveryPipeOuterLong`' and '`lvDeliveryPipeShortConnector`') are both formed of stainless steel enclosing the CO_2 tubes (`lvCO2DeliveryPipeOuterLong` and `lvCO2DeliveryPipeShortConnector`). These daughter volumes that constitute `lvDeliveryPipe` are illustrated in Figure A.11.

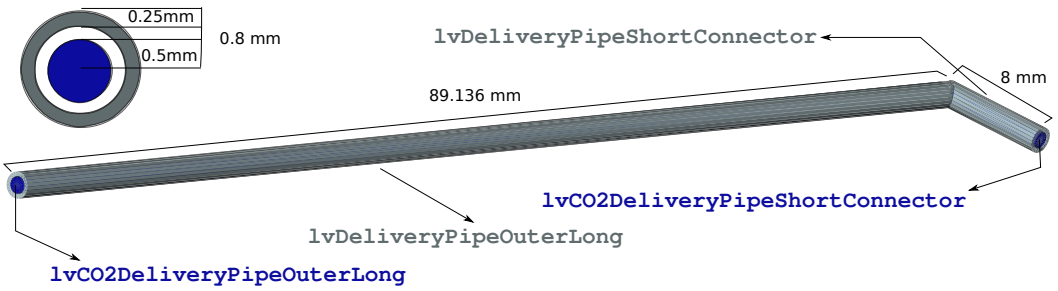

Figure A.11: The CO_2 coolant delivery pipe '`lvDeliveryPipe`', and its components; also shown are the dimensions of the different components.

A.5 Material scan and performance plots

Based on the XML descriptions of the Nikhef and Oxford designs, the total material thickness of a single station comprising a left-right pair of modules, measured along the beam direction z, can be determined using a ray-tracing method. The thickness is determined in terms of length along z axis as a percentage of a radiation length X_0, and is presented in Figure A.12. In these scans the contribution from the RF foil (see Sec. 3.3.2 for more details) is not included. It is evident from the scans that the amount of material in the region close to the beam line is not affected by the choice of the support design, hence no significant differences in physics performance are expected.

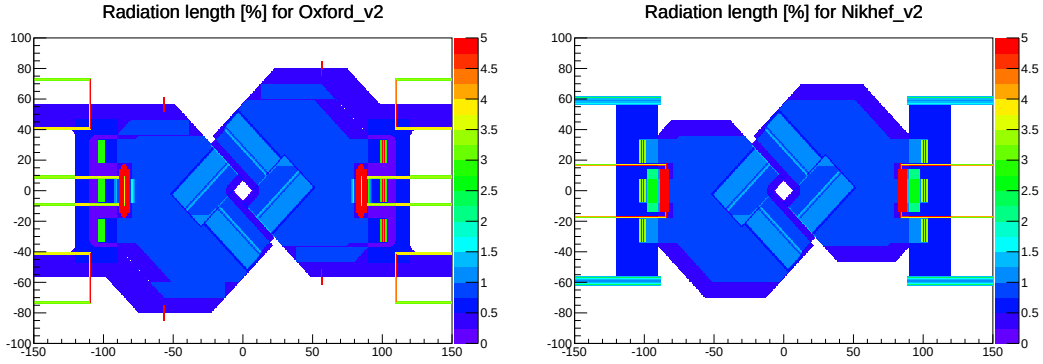

Figure A.12: Thickness along z (in percentage of a radiation length X_0) of a single module for (left) Oxford and (right) Nikhef support designs, based on a perpendicular ray-tracing method.

The overall material length traversed by a particle originating from the origin $(x, y, z) = (0, 0, 0)$ as a function of the angles (ϕ, η) is shown for both the designs in Figure A.13. For these plots the RF foil and all other simulated VELO components are included. Some slight differences are visible, notably a small amount of additional material for the Oxford design in the region $1 < \eta < 3$, and $\phi \approx 0, \pi$. Limiting to the region $2 < \eta < 5$, the average thickness of material traversed as a particle travels through the VELO is 23.1% (22.8%) for the Oxford (Nikhef) design. The material traversed prior to the first and second detector layer is indistinguishable for the two designs, at 1.7% and 3.0% respectively.

To compare the performance of the VELO under the two different designs, using the most important analysis-level quantities, samples of Monte Carlo simulation are generated (each with $\sim$ 40k minimum bias events) for each design and passed through the full simulation chain.

For both module designs, the uncertainty on the impact parameter as a function of $1/p_{\mathrm{T}}$ (where p_{T} is the transverse momentum of the track) and the resolution on the x component of the

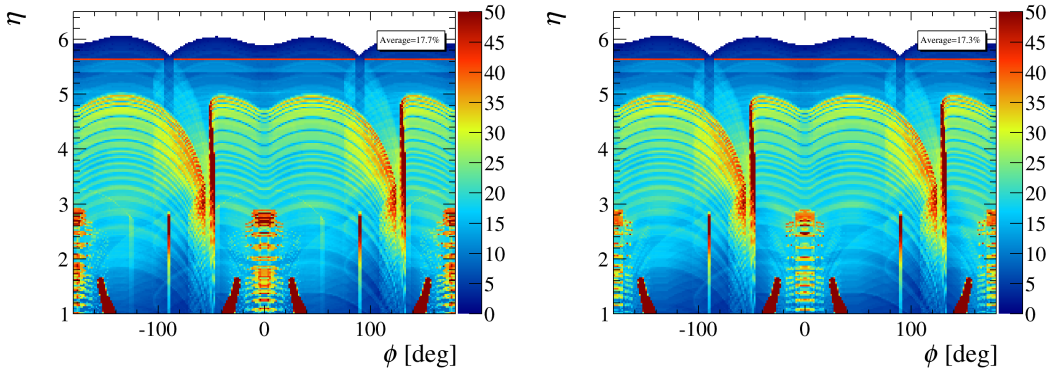

Figure A.13: Total material traversed (in % of a radiation length) by a particle originating at $(0, 0, 0)$ as it travels through the VELO, as a function of the 2D angular space in (ϕ, η) for (left) Oxford and (right) Nikhef VELO module support systems.

primary vertex (PV) are shown in Figure A.14. Similar results are observed for the other two coordinates of the PV. These plots indicate that there are no significant differences in performance driven by the design choice. Track reconstruction efficiencies have also been inspected for both designs. No significant differences are observed, either for the integrated efficiencies, or for efficiency as a function of various kinematic and geometrical variables.

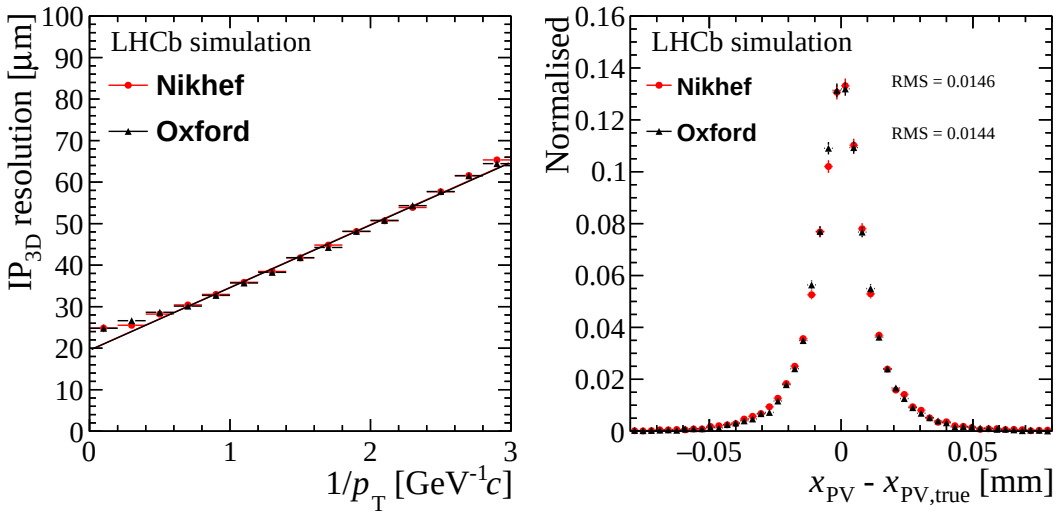

Figure A.14: Track IP and primary vertex spatial resolution for the two proposed module designs; No significant differences are observed.

A.6 Summary

This appendix provides a brief summary of the geometry description of the upgraded VELO module and the two designs proposed by the Oxford and Nikhef LHCb groups for its support system. I have discussed the components of the VELO module and support system that I have been mainly involved in building. A detailed description of the work has been documented in the LHCb internal note [165]. I have also presented the results of material scans and the performance plots of the VELO module for both of the designs. This study concluded that no significant differences in physics performance are exhibited by the two designs of the support structure. The work presented in this appendix formed part of the review process to choose one of the designs as the VELO module support system for the LHCb upgrade. In the end, it was decided that the Nikhef design exhibited higher mechanical stability than the Oxford design and was therefore chosen as the support structure for the upgraded VELO module.

APPENDIX B

Data and simulation consistency for MVA input variables

Possible MC/data discrepancy was investigated using the $B^- \to p\bar{p}K^-$ control channel (inclusion of charge-conjugate processes is implied here) which has high signal statistics and the same topology as our signal modes. For the comparison plots of the input variables that go into the training of the MVA classifier shown in Fig. B.1, we have applied on the MC samples the PID weights obtained from the `PIDCalib` package which corresponds to the respective PID cuts in the data sample.

We do not see a huge discrepancy in any of the variables shown in Fig. B.1 between MC and signal sWeighted data. There are two things to note here:

- In the MVA training, we use the p_{T} asymmetry defined for a cone with an angle of 1.7 radians in the η–ϕ space (`B_STRIP_PTASYM_1_7`). Unfortunately we are unable to investigate the data/MC discrepancy of the variable defined for this particular angle, since it is not stored in the micro-DST format of the `StrippingBu2hhh_pph_inclLine` stripping line. A similar variable with an angle of 1.5 radians is, however, stored, and we are satisfied that if this variable is consistent between data and MC, the corresponding variable with 1.7 radians is highly likely to also be so.
- We also do not show the distribution of the variable `Transf_B_STRIP_VTXISODCHI2ONETRACK`, since the `StrippingBu2hhh_pph_inclLine` micro-DST data files do not have this variable stored. However, it has been shown in the analyses of $\Lambda_b^0(\Xi_b^0) \to phhh$ that this variable does not show a discrepancy between background subtracted data and MC [105].

With regard the cross-check of whether the resampled PID variables match the signal sWeighted data or not, we additionally apply weights on the MC sample so that some of the kinematic variables in MC match the sWeighted $B^- \to p\bar{p}K^-$ data. The resampling is done using the track p, η and event multiplicity (`nTracks`) variables. Figures B.2 and B.3 show the uncorrected, corrected (resampled using `PIDGen` package) and sWeighted PID variable distributions for Run I (2011–2012) and Run II (2015–2016) respectively.

Generally there is a good agreement between sWeighted data and the resampled variables except for slight discrepancies in low p `PROBNN` and high p and K `PROBNN` regions. The h1 and h3 proton tracks show different level of agreements to the sWeighted data which can be attributed to different kinematics of these two final state particles. The low p PROBNN values is a known issue since the proton tracks have a cut of `DLLp`> -5.0 in the stripping. For this discrepancy to be resolved, the resampling must be done before any stripping cuts were applied, which is not possible. A discrepancy can also be seen in the region of high values of the transformed `PROBNN` variables for (anti)protons. Note here that, there will always be residual differences due to the fact that the kernel densities, used in the `PIDGen` package do not perfectly model the true distribution of `PROBNN`, p_{T}, η and `nTracks`. This is not expected to affect the analysis, since the transformation tends to zoom into this region, and therefore the discrepancy is confined in a small region of the `PROBNN` variable that is actually used in the selection. Other analyses have seen similar effects and reached the same conclusion. Therefore, we consider the agreement between corrected MC and sWeighted data to be satisfactory.

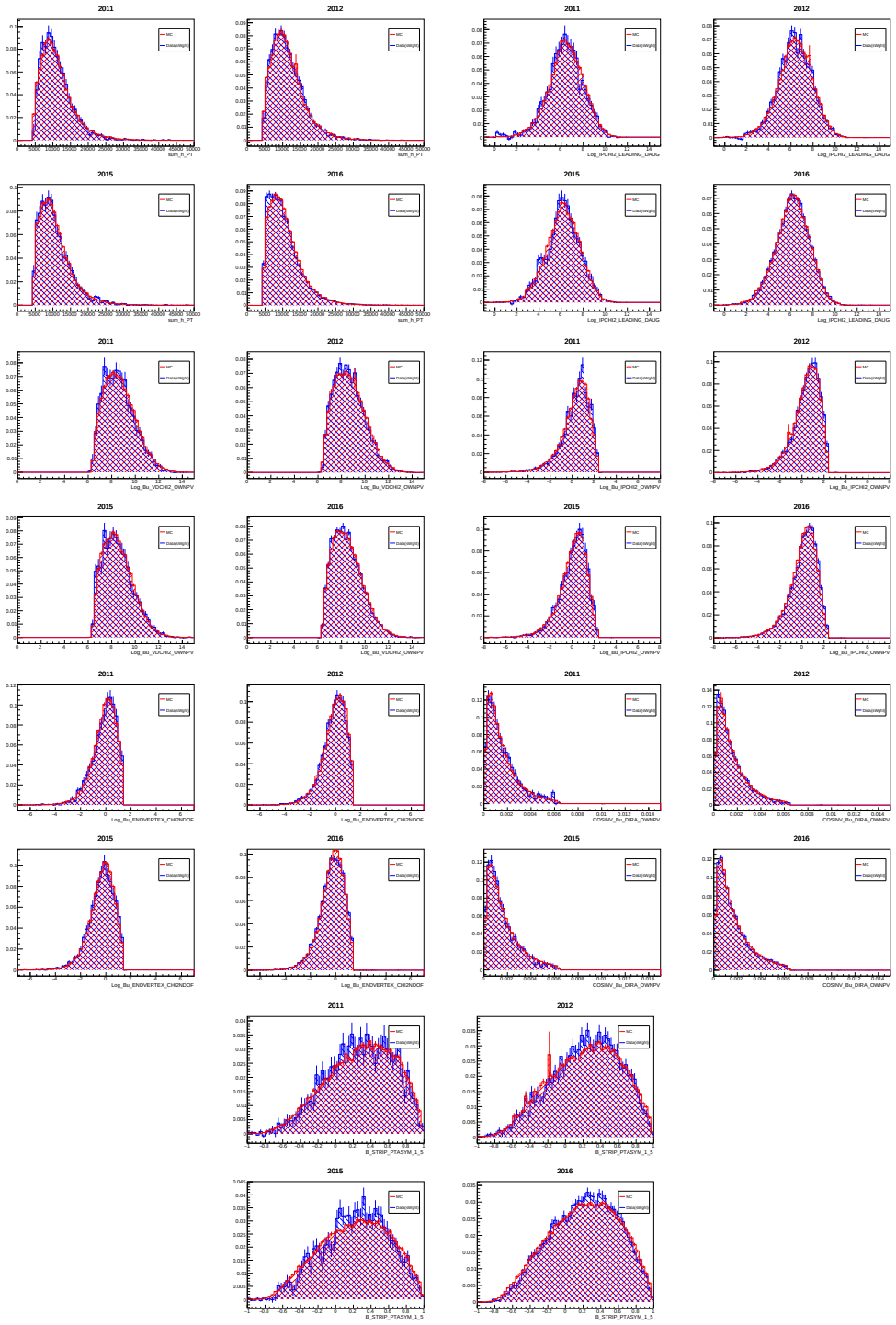

Figure B.1: The subset of MVA input variable distributions in MC and signal sWeighted data for $B^- \to p\bar{p}K^-$ mode. The distributions are grouped in a set of 4 figures each corresponding to the 4 data taking periods. The corresponding year of the distribution is shown in the title and the variable name is on the x-axis.

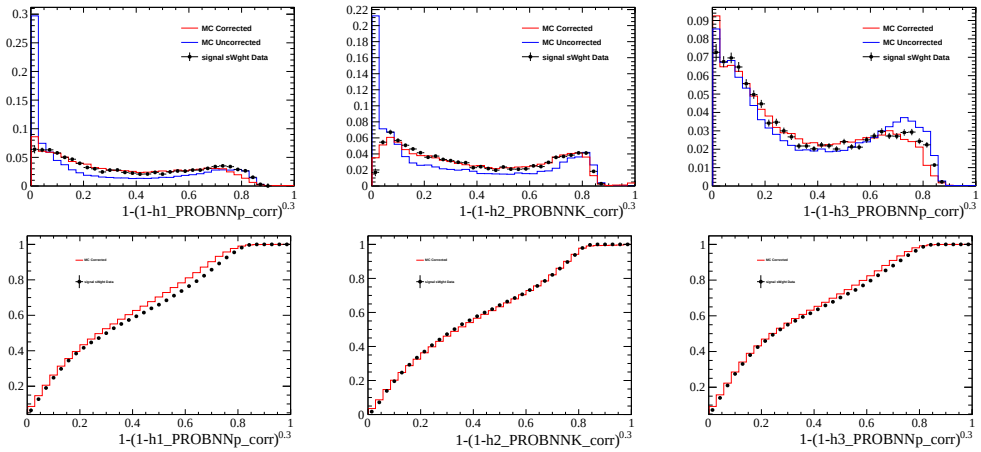

Figure B.2: Comparison of the resampled PID variables (using `PIDGen`) with signal sWeighted $B^- \to p\overline{p}K^-$ data in Run I (2011–2012). Here 'h1', 'h2' and 'h3' are $\overline{p}$, K^- and p particles respectively (and charge conjugate for B^+ decays). Bottom plots show the cumulative distributions of the top plot.

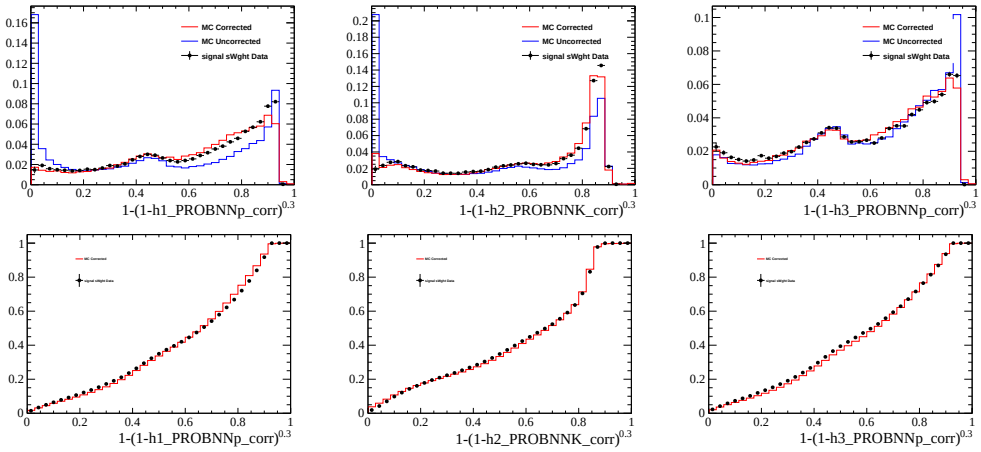

Figure B.3: Comparison of the resampled PID variables (using `PIDGen`) with signal sWeighted $B^- \to p\overline{p}K^-$ data in Run II (2015–2016). Here 'h1', 'h2' and 'h3' are $\overline{p}$, K^- and p particles respectively (and charge conjugate for B^+ decays). Bottom plots show the cumulative distributions of the top plot.

Bibliography

[1] A. A. Penzias and R. W. Wilson, *A measurement of excess antenna temperature at 4080-Mc/s*, Astrophys. J. **142** (1965) 419.

[2] E. Hubble, *A relation between distance and radial velocity among extra-galactic nebulae*, Proc. Nat. Acad. Sci. **15** (1929) 168.

[3] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,61(1991)].

[4] M. Poincaré, *Sur la dynamique de l'électron*, Rendiconti del Circolo Matematico di Palermo (1884-1940) **21** (1906) 129.

[5] S. Weinberg, *The quantum theory of fields. vol. 1: Foundations*, Cambridge University Press, 2005.

[6] E. P. Wigner, *On unitary representations of the inhomogeneous Lorentz group*, Annals Math. **40** (1939) 149, [Reprint: Nucl. Phys. Proc. Suppl.6,9(1989)].

[7] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: A practical introduction*, CRC Press, 1989.

[8] M. Thomson, *Modern particle physics*, Cambridge University Press, New York, 2013.

[9] C.-N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Phys. Rev. **96** (1954) 191.

[10] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. **D2** (1970) 1285.

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

[12] W. E. Burcham, *The Rutherford memorial lecture, 1983: Rutherford and beta decay*, Proc. Roy. Soc. Lond. **A389** (1983) 215.

[13] T. D. Lee and C.-N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104** (1956) 254.

[14] J. Goldstone, A. Salam, and S. Weinberg, *Broken symmetries*, Phys. Rev. **127** (1962) 965.

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

[16] *Standard Model of particle physics*, `https://commons.wikimedia.org/wiki/File:Standard_Model_Of_Particle_Physics--Most_Complete_Diagram.png`. Accessed: 2017-11-1.

[17] J. C. Romao and J. P. Silva, *A resource for signs and Feynman diagrams of the Standard Model*, Int. J. Mod. Phys. **A27** (2012) 1230025, `arXiv:1209.6213`.

[18] D. J. Gross and F. Wilczek, *Ultraviolet behavior of nonabelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343.

[19] V. D. Barger and R. J. N. Phillips, *Collider physics: 1993*, in *7th summer school Jorge Andre Swieca: Particles and fields Sao Paulo, Brazil, January 10-23, 1993*, pp. 0064–168, 1993. `arXiv:hep-ph/9309250`.

[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] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.

[23] M. E. Peskin, *Chiral symmetry and chiral symmetry breaking*, in *Les Houches summer school in theoretical physics: Recent Advances in field theory and statistical mechanics Les Houches, France, August 2-September 10, 1982*, 1982.

[24] J. C. Ward, *An identity in quantum electrodynamics*, Phys. Rev. **78** (1950) 182.

[25] Y. Takahashi, *On the generalized Ward identity*, Nuovo Cim. **6** (1957) 371.

[26] E. Noether, *Invariante variationsprobleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235.

[27] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 1 and 2*, Developments in the quark theory of hadrons, volume 1. Edited by D. Lichtenberg and S. Rosen. pp. 22-101 (1964) 22.

[28] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Lett. **8** (1964) 214.

[29] *CDF collaboration finds new baryons that contain b quarks*, `http://cerncourier.com/cws/article/cern/29766`. Accessed: 2017-11-1.

[30] A. Ali, C. Hambrock, A. Ya. Parkhomenko, and W. Wang, *Light-cone distribution amplitudes of the ground state bottom baryons in HQET*, Eur. Phys. J. **C73** (2013) 2302, `arXiv:1212.3280`.

[31] H. Greaves and T. Thomas, *On the CPT theorem*, Stud. Hist. Phil. Sci. **B45** (2014) 46, `arXiv:1204.4674`.

[32] CPLEAR collaboration, A. Angelopoulos *et al.*, *First direct observation of time reversal noninvariance in the neutral kaon system*, Phys. Lett. **B444** (1998) 43.

[33] BaBar collaboration, J. P. Lees *et al.*, *Observation of time reversal violation in the B^0 meson system*, Phys. Rev. Lett. **109** (2012) 211801, `arXiv:1207.5832`.

[34] BaBar, B. Aubert *et al.*, *Limits on the decay-rate difference of neutral B mesons and on CP, T, and CPT violation in $B^0\overline{B}^0$ oscillations*, Phys. Rev. Lett. **92** (2004) 181801, `arXiv:hep-ex/0311037`.

[35] BaBar, B. Aubert *et al.*, *Limits on the decay rate difference of neutral$-B$ mesons and on CP, T, and CPT violation in $B^0\overline{B}^0$ oscillations*, Phys. Rev. **D70** (2004) 012007, `arXiv:hep-ex/0403002`.

[36] LHCb collaboration, R. Aaij *et al.*, *Search for violations of Lorentz invariance and CPT symmetry in $B^0(B^0_s)$ mixing*, Phys. Rev. Lett. **116** (2016) 241601, `arXiv:1603.04804`.

[37] ALPHA, M. Ahmadi *et al.*, *Observation of the 1S-2P Lyman-α transition in antihydrogen*, Nature **561** (2018), no. 7722 211, [Nature(2018)].

[38] ALPHA, M. Ahmadi *et al.*, *Characterization of the 1S-2S transition in antihydrogen*, Nature **557** (2018), no. 7703 71.

[39] J. D. Lykken, *Beyond the Standard Model*, in *CERN Yellow Report CERN-2010-002, 101-109*, 2010. `arXiv:1005.1676`.

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

[41] M. Trodden, *Baryogenesis and leptogenesis*, eConf **C040802** (2004) L018, arXiv:hep-ph/0411301.

[42] G. Steigman and R. J. Scherrer, *Is the universal matter - antimatter asymmetry fine tuned?*, 2018. arXiv:1801.10059.

[43] WMAP collaboration, C. L. Bennett *et al.*, *First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Preliminary maps and basic results*, Astrophys. J. Suppl. **148** (2003) 1, arXiv:astro-ph/0302207.

[44] A. G. Cohen, *CP violation and the origins of matter*, in *Proceedings, 29th SLAC summer institute on particle physics: Exploring electroweak symmetry breaking (SSI 2001): Menlo Park, California, August 13-24, 2001*, p. ch01, 1999.

[45] N. Terasawa and K. Sato, *Lepton and baryon number asymmetry of the universe and primordial nucleosynthesis*, Prog. Theor. Phys. **80** (1988) 468.

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

[47] D. V. Perepelitsa, *Sakharov conditions for baryogenesis*, Recuperado de http://phys. columbia. edu/˜ dvp/dvp-sakharov. pdf, el **3** (2008).

[48] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, NATO Sci. Ser. B **59** (1980) 135.

[49] S. L. Adler, *Axial vector vertex in spinor electrodynamics*, Phys. Rev. **177** (1969) 2426.

[50] 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.

[51] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.

[52] T. Mannel, *Theory and phenomenology of CP violation*, Nuclear Physics B-Proceedings Supplements **167** (2007) 170.

[53] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

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

[55] L.-L. Chau and W.-Y. Keung, *Comments on the parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[56] L. Wolfenstein, *Violation of CP invariance and the possibility of very weak interactions*, Phys. Rev. Lett. **13** (1964) 562.

[57] 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.

[58] 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.

[59] UTfit, M. Bona *et al.*, *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Delta m(s) and BR(B —¿ tau nu(tau))*, JHEP **10** (2006) 081, arXiv:hep-ph/0606167.

[60] Y. Nir, *CP violation in meson decays*, in *High-energy physics. Proceedings, 3rd Latin american CERN-CLAF school, Malargue, Argentina, February 27-March 12, 2005*, pp. 79–145, 2006. arXiv:hep-ph/0510413.

[61] T. Gershon and V. V. Gligorov, *CP violation in the B system*, Rept. Prog. Phys. **80** (2017) 046201, arXiv:1607.06746.

[62] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetries in two-body B^0 or B_s^0-meson decays to charged pions and kaons*, Phys. Rev. **D98** (2018) 032004, arXiv:1805.06759.

[63] 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.

[64] J. D. Richman, *1997 Les Houches summer school: Heavy quark physics and CP violation*, http://courses.physics.ucsd.edu/2010/Winter/physics222/references/driver_houches12.pdf.

[65] LHCb collaboration, R. Aaij *et al.*, *A precise measurement of the B^0 meson oscillation frequency*, Eur. Phys. J. **C76** (2016) 412, arXiv:1604.03475.

[66] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to J/\psi K^0_S$ decays*, Phys. Rev. Lett. **115** (2015) 031601, arXiv:1503.07089.

[67] J. M. Pendlebury *et al.*, *Revised experimental upper limit on the electric dipole moment of the neutron*, Phys. Rev. **D92** (2015) 092003, arXiv:1509.04411.

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

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

[70] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137, arXiv:hep-ph/0005110.

[71] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, arXiv:0901.0002.

[72] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, arXiv:1009.2731.

[73] *LHCb material for presentations*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html. Accessed: 2018-01-18.

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

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

[76] J. R. Harrison, *Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays*, PhD thesis, Manchester U., 2014.

[77] *Measurement in a magnetic field*, https://cbooth.staff.shef.ac.uk/phy6040det/magfield.html. Accessed: 2018-01-18.

[78] V. V. Gligorov, *Conceptualization, implementation, and commissioning of real-time analysis in the High Level Trigger of the LHCb experiment*, arXiv:1806.10912.

[79] H. Yin, *Tracking and vertex reconstruction at LHCb for Run II*, https://cds.cern.ch/record/2265227, May, 2017.

[80] *Chris Jones presentation*, https://indico.cern.ch/event/226062/contributions/475644/attachments/371741/517276/ANNPIDRetuning-Reco14-06052013.pdf. Accessed: 2018-01-18.

[81] *Cherenkov radiation*, http://www.masteringphysicssolutions.net.

[82] LHCb collaboration, 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.

[83] LHCb collaboration, B. Sciascia, *LHCb Run 2 trigger performance*, PoS **BEAUTY2016** (2016) 029.

[84] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME–Journal of Basic Engineering **82** (1960), no. Series D 35.

[85] 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.

[86] 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.

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

[88] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13\,TeV$*, JHEP **03** (2016) 159, Erratum ibid. **09** (2016) 013, Erratum ibid. **05** (2017) 074, arXiv:1510.01707.

[89] LHCb collaboration, R. Antunes-Nobrega and other, *LHCb computing: Technical design report*, LHCC-2015-019. Submitted on 11 May 2005.

[90] *LHCb Gaudi web page*, http://lhcb-comp.web.cern.ch/lhcb-comp/Frameworks/Gaudi/.

[91] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[92] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics*, .

[93] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[94] *LHCb Boole web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[95] *LHCb Gauss web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/gauss/.

[96] *LHCb Moore web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/moore/.

[97] *LHCb Brunel web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/.

[98] *LHCb DaVinci web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[99] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the phase space of $B^\pm \to K^\pm\pi^+\pi^-$ and $B^\pm \to K^\pm K^+K^-$ decays*, Phys. Rev. Lett. **111** (2013) 101801, arXiv:1306.1246.

[100] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in the phase space of $B^\pm \to K^+K^-\pi^\pm$ and $B^\pm \to \pi^+\pi^-\pi^\pm$ decays*, Phys. Rev. Lett. **112** (2014) 011801, arXiv:1310.4740.

[101] CDF collaboration, T. Aaltonen *et al.*, *Observation of new charmless decays of bottom hadrons*, Phys. Rev. Lett. **103** (2009) 031801, arXiv:0812.4271.

[102] LHCb collaboration, R. Aaij *et al.*, *Searches for Λ_b^0 and Ξ_b^0 decays to $K^0_S p\pi^-$ and $K^0_S pK^-$ final states with first observation of the $\Lambda_b^0 \to K^0_S p\pi^-$ decay*, JHEP **04** (2014) 087, arXiv:1402.0770.

[103] LHCb collaboration, R. Aaij *et al.*, *Observation of the $\Lambda_b^0 \to \Lambda\phi$ decay*, Phys. Lett. **B759** (2016) 282, arXiv:1603.02870.

[104] LHCb collaboration, R. Aaij *et al.*, *Observations of $\Lambda_b^0 \to \Lambda K^+\pi^-$ and $\Lambda_b^0 \to \Lambda K^+K^-$ decays and searches for other Λ_b^0 and Ξ_b^0 decays to Λh^+h^- final states*, JHEP **05** (2016) 081, arXiv:1603.00413.

[105] LHCb collaboration, R. Aaij *et al.*, *Measurement of branching fractions of charmless four-body Λ_b^0 and Ξ_b^0 decays*, JHEP **02** (2018) 098, arXiv:1711.05490.

[106] LHCb, R. Aaij *et al.*, *Search for CP violation in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays*, Submitted to: Phys. Lett. (2018) arXiv:1807.06544.

[107] LHCb collaboration, R. Aaij *et al.*, *Measurement of matter-antimatter differences in beauty baryon decays*, Nature Physics **13** (2017) 391, arXiv:1609.05216.

[108] LHCb, R. Aaij *et al.*, *Search for CP violation using triple product asymmetries in $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$ and $\Xi_b^0 \to pK^-K^-\pi^+$ decays*, JHEP **08** (2018) 039, arXiv:1805.03941.

[109] BaBar collaboration, B. Aubert *et al.*, *Search for CP violation in neutral D meson Cabibbo-suppressed three-body decays*, Phys. Rev. **D78** (2008) 051102, arXiv:0802.4035.

[110] I. Bediaga *et al.*, *On a CP anisotropy measurement in the Dalitz plot*, Phys. Rev. **D80** (2009) 096006, arXiv:0905.4233.

[111] M. Williams, *Observing CP violation in many-body decays*, Phys. Rev. **D84** (2011) 054015, arXiv:1105.5338.

[112] I. Bediaga *et al.*, *Second generation of "Miranda procedure" for CP violation in Dalitz studies of B (& D & τ) decays*, Phys. Rev. **D86** (2012) 036005, arXiv:1205.3036.

[113] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the mass and lifetime of the Ξ_b^- baryon*, Phys. Rev. Lett. **113** (2014) 242002, arXiv:1409.8568.

[114] LHCb collaboration, R. Aaij *et al.*, *Measurements of the mass and lifetime of the Ω_b^- baryon*, Phys. Rev. **D93** (2016) 092007, arXiv:1604.01412.

[115] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.

[116] 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.

[117] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

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

[119] 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.

[120] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

[121] L. Anderlini and other, *The PIDCalib package*, LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021.

[122] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[123] BaBar, 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.

[124] BaBar collaboration, J. P. Lees *et al.*, *Measurement of CP-violating asymmetries in $B^0 \to (\rho\pi)^0$ decays using a time-dependent Dalitz plot analysis*, Phys. Rev. **D88** (2013) 012003, arXiv:1304.3503.

[125] LHCb collaboration, R. Aaij *et al.*, *Measurements of the $\Lambda_b^0 \to J/\psi\,\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7\,\text{TeV}$*, Phys. Lett. **B724** (2013) 27, arXiv:1302.5578.

[126] LHCb collaboration, R. Aaij *et al.*, *Study of the D^0p amplitude in $\Lambda_b^0 \to D^0p\pi^-$ decays*, JHEP **05** (2017) 030, arXiv:1701.07873.

[127] B. Efron, *Bootstrap methods: another look at the jackknife*, Springer, 1992.

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

[129] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[130] F. James and M. Roos, *Minuit: A system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun. **10** (1975) 343.

[131] A. Lazzaro and L. Moneta, *MINUIT package parallelization and applications using the RooFit package*, J. Phys. Conf. Ser. **219** (2010) 042044.

[132] S. Ruder, *An overview of gradient descent optimization algorithms*, arXiv:1609.04747.

[133] B. Wetton, *Finite difference methods for ordinary and partial differential equations: Steady-state and time-dependent problems*, JSTOR, 2009.

[134] G. Cowan, *Statistical data analysis*, Oxford University Press, 1998.

[135] LHCb collaboration, R. Aaij *et al.*, *Observation of the decay $\Xi_b^- \to pK^-K^-$*, Phys. Rev. Lett. **118** (2017) 071801, arXiv:1612.02244.

[136] M. Whitehead, T. Gershon, M. Kreps, and P. Harrison, *Observation of the decay $B^0 \to D^0K^+K^-$ with the LHCb detector at CERN*, Jun, 2012. Presented 16 Aug 2012.

[137] D. P. O'Hanlon, *Studies of CP-violation in charmless three-body b-hadron decays*, CERN-THESIS-2017-184.

[138] CDF collaboration, T. A. Aaltonen *et al.*, *Measurements of direct CP-violating asymmetries in charmless decays of bottom baryons*, Phys. Rev. Lett. **113** (2014) 242001, arXiv:1403.5586.

[139] LHCb collaboration, R. Aaij *et al.*, *Evidence for exotic hadron contributions to $\Lambda_b^0 \to J/\psi p\pi^-$ decays*, Phys. Rev. Lett. **117** (2016) 082003, arXiv:1606.06999.

[140] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and regression trees*, Wadsworth and Brooks, Monterey, CA, 1984.

[141] L. Breiman, *Random forests*, Machine Learning **45** (2001) 5.

[142] T. Chen and C. Guestrin, *XGBoost: A scalable tree boosting system*, in *Proceedings of the 22Nd ACM SIGKDD International conference on knowledge discovery and data mining*, KDD '16, p. 785, 2016. arXiv:1603.02754. doi: 10.1145/2939672.2939785.

[143] J. Stevens and M. Williams, *uBoost: A boosting method for producing uniform selection efficiencies from multivariate classifiers*, JINST **8** (2013) P12013, arXiv:1305.7248.

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

[145] T. Schaul *et al.*, *PyBrain*, Journal of Machine Learning Research **11** (2010) 743.

[146] *NeuroLab*, http://pythonhosted.org/neurolab/.

[147] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in python*, Journal of Machine Learning Research **12** (2011) 2825.

[148] *REP (Reproducible Experiment Platform)*, http://yandex.github.io/rep/.

[149] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

[150] LHCb collaboration, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **08** (2013) 117, arXiv:1306.3663.

[151] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.

[152] LHCb collaboration, R. Aaij *et al.*, *Observation of $B^0 \to \overline{D}^0 K^+ K^-$ and evidence for $B^0_s \to \overline{D}^0 K^+ K^-$*, Phys. Rev. Lett. **109** (2012) 131801, arXiv:1207.5991.

[153] J. D. Richman, *An Experimenter's Guide to the Helicity Formalism*, .

[154] E. Wigner, *Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra*, 1959, Academic Press.

[155] G. Breit and E. Wigner, *Capture of Slow Neutrons*, Phys. Rev. **49** (1936) 519.

[156] Belle, BaBar, A. J. Bevan *et al.*, *The Physics of the B Factories*, Eur. Phys. J. **C74** (2014) 3026, arXiv:1406.6311.

[157] J. Blatt and V. Weisskopf, *Theoretical Nuclear Physics*, 1991, p. 627.

[158] M. Ebden, *Gaussian processes: A quick introduction*, ArXiv e-prints (2015) arXiv:1505.02965.

[159] *GPy python package*, http://gpy.readthedocs.io/en/deploy/.

[160] LHCb collaboration, R. Aaij *et al.*, *Dalitz plot analysis of $B^0_s \to \overline{D}^0 K^- \pi^+$ decays*, Phys. Rev. **D90** (2014) 072003, arXiv:1407.7712.

[161] B. Guegan, J. Hardin, J. Stevens, and M. Williams, *Model selection for amplitude analysis*, JINST **10** (2015) P09002, arXiv:1505.05133.

[162] P. d'Argent *et al.*, *Amplitude analyses of $D^0 \to \pi^+\pi^-\pi^+\pi^-$ and $D^0 \to K^+K^-\pi^+\pi^-$ Decays*, JHEP **05** (2017) 143, arXiv:1703.08505.

[163] M. Williams, *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*, JINST **5** (2010) P09004, arXiv:1006.3019.

[164] LHCb collaboration, R. Aaij *et al.*, *Measurement of B^0, B^0_s, B^+ and Λ^0_b production asymmetries in 7 and 8 TeV pp collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.

[165] A. C. Crocombe, T. E. Latham, A. Mathad, and M. R. J. Williams, *Geometry descriptions and performance studies of the VELO upgrade module support candidate designs*, LHCb-INT-2018-025. CERN-LHCb-INT-2018-025.

[166] *LHCb conditions database pages*, https://twiki.cern.ch/twiki/bin/view/LHCb/CondDBPages.

[167] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[168] *The Panoramix project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/panoramix/.