Search for rare baryonic b decays with the LHCb experiment at the LHC

James McCarthy

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

September 28, 2015

ABSTRACT

A search was performed for the $\Lambda_b^0 \to \Lambda^0 \eta'$ and $\Lambda_b^0 \to \Lambda^0 \eta$ decays with the LHCb experiment. The full dataset recorded by LHCb in 2011 and 2012 is used, corresponding to $1.0\,\text{fb}^{-1}$ of proton-proton collision data collected at a centre of mass energy of 7 TeV, and $2.0\,\text{fb}^{-1}$ of data collected at a centre of mass energy of 8 TeV, respectively. The $B^0 \to K_S^0 \eta'$ decay is used as a normalisation channel, and a selection is designed and optimised using this decay. By measuring the ratio of the branching fractions for the signal decay to the normalisation decay, many systematic uncertainties cancel out. No significant signal is observed for the $\Lambda_b^0 \to \Lambda^0 \eta'$ decay, and some evidence is observed for the $\Lambda_b^0 \to \Lambda^0 \eta$ decay with a significance of 3σ. The Feldman-Cousins method is used to make the first measurement of the limit on the branching fractions. The limits are

$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta') < 3.1 \times 10^{-6} \text{ at } 90\% \text{ CL.}$$
$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta) \in [2.5, 22.8] \times 10^{-6} \text{ at } 90\% \text{ CL.}$$

DECLARATION OF AUTHOR'S CONTRIBUTION

The analysis and technical work presented in this thesis was carried out by myself.

As a technical contribution, I was part of the simulation group in LHCb. I was responsible for adapting part of the xml description of the LHCb detector to model the new supports for the beam pipe. In addition, I helped to validate the new releases of GEANT4 by comparing the hadronic physics lists, as described in this thesis. Finally, the studies of the GEANT4 cuts applied described here were performed entirely by me.

The analysis is my own work. The selection was designed and optimised by myself, and I used the ROOT and TMVA frameworks to train the BDTs. The trigger selection was designed and implemented by the LHCb collaboration. The fit model was also designed by me and the code was written using the RooFit framework. At regular intervals the analysis was presented to the Charmless B working group, and suggestions were made by various members of the group to help improve the analysis. The analysis was written up as an internal LHCb analysis note and reviewed by the group before being written up as a paper. The results will soon be published in JHEP, and are available on the arXiv [1].

My final contribution to the LHCb experiment was as the trigger liaison for the Charmless B working group, keeping the group informed of the current status of the trigger for Run II of the LHC, and testing out the planned trigger configurations.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my wife, Kate. She has been extremely patient and supportive over the last four years, especially when I was going through busy periods working all hours of the day. Without her, I wouldn't have been able to get to the end, and I hope it was all worth it when I finally get awarded my PhD.

Secondly, thanks to Nigel Watson, my supervisor. He has been a great supervisor, from allowing me the freedom to choose an analysis for myself to spending hours helping me try to get GEANT4 code to work in the LHCb framework.

I would also like to acknowledge the members of the LHCb collaboration who have helped with useful suggestions for improving the analysis. In particular, I am grateful for the help from the various conveners of the Charmless B decays working group, Stephane Monteil, Eduardo Rodrigues and Thomas Latham.

Special thanks also goes to Mum and Dad, who's support over the years guided me to university and onto a PhD.

Finally, this thesis is dedicated to Granddad Jim, who unfortunately passed away before I finished my PhD. He taught me the value of knowledge and the fact that you can't know where you are going unless you know where you have come from.

Dedicated to Granddad Jim

Contents

List of Tables

List of Figures

DEFINITIONS OF ACRONYMS

LHC Large Hadron Collider

CERN European Organisation for Nuclear Research

RF Radio-Frequency

PSB Proton Synchrotron Booster

PS Proton Synchrotron

SPS Super Proton Synchrotron

SLAC the Stanford Linear Accelerator Center

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

SM Standard Model of particle physics

EW ElectroWeak

CC "Charged Current"

NC "Neutral Current"

FCNC Flavour Changing Neutral Current

GUT Grand Unified Theory

NP New Physics

GPD General Purpose Detector

GPDs General Purpose Detectors

PDF probability density function

CP combined symmetry of Charge conjugation and Parity transformation

CPV *CP* violation

CKM Cabibbo-Kobayashi-Maskawa

HQET Heavy Quark Expansion Theory

MC Monte Carlo

HEP High Energy Physics

pp proton-proton

QGS Quark Gluon String

LEPar Low Energy Parameterisation

HEPar High Energy Parameterisation

EM Elecromagnetic

HAD Hadronic

CL Confidence Level

FC Feldman-Cousins

PV Primary Vertex

LL Long-Long

DD Downstream-Downstream

BDT Boosted Decision Tree

NN Neural Network

DOCA Distance of Closest Approach

DIRA Direction Angle

PID Particle Identification

HPDs Hybrid Photo-Detectors

MWPCs Multi-Wire Proportional Chambers

VELO VErtex LOcator

TT Tracker Turicensis

ST Silicon Tracker

RICH Ring Imaging CHerenkov

IT Inner Tracker

OT Outer Tracker

PRS Pre-Shower detector

SPD Scintillating Pad Detector

ECAL Electromagnetic Calorimeter

HCAL Hadronic Calorimeter

CHAPTER

1

INTRODUCTION

This chapter provides a general introduction to particle physics and to the subject of this thesis. After briefly describing the Standard Model of particle physics (SM), the Large Hadron Collider (LHC) and the LHCb experiment will be introduced, before the main objectives of the thesis are outlined.

1.1 The standard model of particle physics

Since the time of the Ancient Greek philosophers, attempts have been made to describe and understand the world around us. Scientists try to answer questions such as "How did the universe come into being?" and "What is everything made of?". Particle physics attempts to understand the world on the very small scale, by describing the fundamental "building blocks" (or particles) of which all matter is made, and how these particles interact with each other.

The particles that form all of the matter we are familiar with are protons, neutrons and electrons. The number of protons and neutrons in an atom determines the properties of the atom, and combinations of these three basic particles produce the entire range of chemical elements we observe in the periodic table. As well as the proton, neutron and electron, there is a vast array of similar particles that behave in different ways. These do not exist under typical conditions on Earth, but can be produced, for example, in high energy collisions of cosmic rays in the atmosphere and in particle physics experiments. These particles typically have a short lifetime ($\ll 10^{-6}$s) and will decay into the lighter matter particles.

In addition, we also find that every particle has a corresponding anti-particle, which has the same mass and opposite quantum numbers. In 1928, Dirac was attempting to reconcile Schroedinger's equation in quantum mechanics with special relativity. The solution to his equation involved the introduction of 4-component spinors. These components are interpreted as two particles with spin-up and spin-down, with the second particle being the anti-particle [2,3]. In 1932, Anderson carried out an experiment to study tracks from cosmic ray showers in a cloud chamber, and discovered a particle that had the same mass as the electron but with a positive charge [4]. This is interpreted as the positron, or the anti-electron, and provided experimental confirmation of Dirac's prediction and the existence of anti-matter.

For a long time, these particles were thought to be the fundamental constituents of all matter. However, puzzles started to arise when a group of particles with similar properties (referred to generically as hadrons) were discovered by studying the interactions of cosmic rays in the atmosphere. In the early 1960s, two physicists, Gell-Man and Ne-eman, independently formulated a theory that grouped these hadrons together and described them in terms of constituent components called "quarks" [5]. In this paper, there were three flavours of quarks described, up (u), down (d) and strange (s), each of which had a fractional charge of $-1/3|e|$ for the d and s quarks, and $+2/3|e|$ for the u quark. Quarks were discovered in an experiment at the Stanford Linear Accelerator Center (SLAC) in 1968 [6] by measuring the scattering angle of electrons in electron-proton collisions. When the energy of the electron is low, the

scattering angles are more likely to be small, which is indicative of scattering from a fundamental charged particle. However, when the electron energy is high, there is more scattering at larger angles than is expected indicating the presence of some substructure of the proton. This was the first experimental evidence to support the quark model of hadrons.

We now know that there are six flavours of quarks, and their corresponding anti-quarks. The quarks are grouped into two types, labelled the up-type quarks with charge $+2/3|e|$, and down-type quarks with charge $-1/3|e|$. Properties of each of the quark flavours are shown in Table 1.1.

Table 1.1: Properties of each quark flavour from [7].

Flavour	**Charge** ($|e|$)	**Mass** (GeV/c^2)
down (d)	$-1/3$	$0.0048^{+0.0005}_{-0.0003}$
up (u)	$+2/3$	$0.0023^{+0.0007}_{-0.0005}$
strange (s)	$-1/3$	0.095 ± 0.005
charm (c)	$+2/3$	1.275 ± 0025
bottom (b)	$-1/3$	4.18 ± 0.03
top (t)	$+2/3$	173.21 ± 0.87

In addition to the quarks, there are three generations of "leptons": the charged leptons, e^-, μ^-, and τ^-; and the neutrinos, ν_e, ν_μ and ν_τ.

The interactions of these particles are described in the SM, and their defining characteristic is whether they are subject to the strong interaction (quarks) or not (leptons). The SM is a quantum field theory formed from the product of SU(3)⊗SU(2)⊗U(1) symmetry groups. The SM is a gauge theory, in which the fundamental forces are described by an exchange model of vector gauge bosons.

The strong force is mediated by the gluon, g, and is described by Quantum Chromodynamics (QCD), which is an SU(3) group. The strong interaction acts on the colour quantum number of quarks and gluons. This quantum number can take one of three values for quarks, labelled RGB, and hence the SU(3) group is made of eight 3×3 matrices, known as the Gell-Mann matrices, which describe the colour quantum

numbers of gluons involved in the strong interaction. The interesting point here is that the gluon itself also has colour quantum numbers, and hence self-interactions of gluons are possible. These are absent from Quantum Electrodynamics (QED) since the photon does not have an electric charge, and lead to interesting properties of QCD including confinement and asymptotic freedom. The strength of the strong interaction becomes very high at long distances or low energies. For this reason, quarks and gluons can never be observed individually and are confined within hadrons. These could be mesons, which consist of a q-$\overline{\text{q}}$ pair, or baryons, which consist of three quarks or three anti-quarks. The strength of the interaction becomes weaker at shorter distances, which allows us to study individual quarks in high energy collisions.

The final group is an SU(2)⊗U(1) group that describes the ElectroWeak (EW) interaction which combines the electromagnetic and the weak interaction. The weak interaction is mediated by two vector bosons. Interactions that involve the exchange of $\text{W}^{\pm}$ bosons are called "Charged Current" (CC) interactions, and those mediated by a Z^0 boson are called "Neutral Current" (NC) interactions. Quarks and leptons form doublets of this group, with a new quantum number called weak isospin, I. The charged leptons form doublets with their corresponding neutrinos, and the up-type quarks form doublets with the down-type quarks, such that we have the following doublets of the SU(2) group:

$$\begin{pmatrix} \nu_e \\ \text{e}^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}, \begin{pmatrix} \text{u} \\ \text{d} \end{pmatrix}, \begin{pmatrix} \text{c} \\ \text{s} \end{pmatrix}, \begin{pmatrix} \text{t} \\ \text{b} \end{pmatrix}$$

with the corresponding anti-particles. Each particle in a doublet has the same value of $I = 1/2$, but different values of the the third component of isospin, $I_3 = +(-)1/2$ for upper (lower) component of the doublet.

The U(1) group describes the electromagnetic interaction of charged particles, which is mediated by the exchange of photons,γ, and is known as QED.

The final piece of the puzzle in the SM is the origin of the mass of particles, particu-

larly the $W^{\pm}$ and Z^0 bosons, and in the late 1960s various solutions were proposed. The solution consists of the introduction of a scalar field that permeates the entire space-time [8–13]. Particles obtain mass through Yukawa interactions, which describe an interaction between a scalar field and a Dirac (particle) field, with particles of higher mass having a stronger Yukawa coupling. The consequence of this scalar field is the breaking of the SU(2) symmetry of the EW sector, and the appearance of a new spin-0 particle, usually called the Higgs Boson. In 2012, a new particle was discovered by the ATLAS and CMS experiments at the LHC, with properties consistent with the SM Higgs boson, and with a mass of $\approx 125\,\mathrm{GeV}/c^2$ [14, 15].

More details of the SM and the impact on heavy flavour physics are discussed in Chapter 2.

1.1.1 Beyond the standard model

The SM is a beautiful theory that has been tested with extremely high precision [16]. However, there are some shortcomings of the model. In particular, there are 19 free parameters which describe the model. From a scientific point of view, this is somewhat unsatisfactory as it involves a large amount of fine tuning of the model. Other physical phenomena that are not described by the SM include the presence of dark matter/energy in the universe [17] and the observed weakness of the gravitational force. Reconciling the gravitational force with the other fundamental forces in the SM is something which requires a new approach to the model, and this is not trivial. The force could be mediated by spin-2 particles called gravitons [18], but this does not explain why the strength of the gravitational force is so small.

There are numerous new models which have been theorised to explain these problems. It is possible that the SM described above is a low energy approximation of some Grand Unified Theory (GUT) [19]. Although many such theories have been proposed, there is currently no compelling experimental evidence for any physics beyond the standard model.

One of the most interesting puzzles not explained by the SM is the difference between matter and anti-matter. Particles and anti-particles are assumed to have been produced from energy around the time of the Big Bang. Due to the conservation laws of the SM, equal numbers of particles and anti-particles are produced in these processes. Since in every direction we look in the universe we observe only matter, and no anti-matter, there must be some small difference in the behaviour of particles and anti-particles that has caused the currently observed imbalance. This difference can be described in the SM through the process of *CP* violation, described in more detail in Chapter 2. Since particles and antiparticles annihilate to produce photons, the total asymmetry between matter and anti-matter can be found be comparing the number of photons with the number of baryons in the universe. With all of the known sources of *CP* violation, there is still 10^9 times more matter than can be explained [20]. Searches for New Physics (NP) therefore include searches for new sources of *CP* violation, which is the focus of the LHCb experiment.

1.2 The Large Hadron Collider and the LHCb experiment

The LHC is a large particle physics facility, based at the European Organisation for Nuclear Research (CERN) in Geneva. Bunches of protons are accelerated around a 27 km ring, and collide with protons travelling in the opposite direction at four points around the ring. The high centre of mass energy of the proton-proton (pp) collisions allows new particles to be produced and studied. The LHC started running in 2010, with the proton beams colliding at a centre of mass energy of $\sqrt{s}$ =7 TeV, until the energy was increased to $\sqrt{s}$ =8 TeV in 2012. The LHC was then shut down to upgrade the magnets so that the beams could be accelerated to collide at the design energy of $\sqrt{s}$ =13–14 TeV. The higher centre of mass energy will extend the physics programmes of the LHC experiments.

There are four LHC experiments, with their detectors situated at the four collision points along the LHC ring. ATLAS and CMS are General Purpose Detectors (GPDs)

that surround their interaction regions. They are designed to study particles produced in pp collisions with high transverse momentum, p_{T}, and search directly for new particles being produced. ALICE is a heavy ion experiment, which investigates the collisions of lead nuclei, in addition to pp collisions. The extremely high energy density of the lead collisions produces conditions similar to those predicted to have existed in the universe in the first seconds after the Big Bang, and so studying these collisions may provide information about how the universe evolved. The fourth experiment is the LHCb experiment, which is a dedicated heavy flavour experiment. The work described in this thesis is based on the data collected by the LHCb experiment, and so this is discussed in more detail.

The LHCb detector is optimised for precision measurements of hadrons containing b- and c- quarks. In contrast with the GPDs, the acceptance of the LHCb detector covers only the forward direction. In addition the instantaneous luminosty of pp collisions in the LHCb detector is lower than that of the GPDs. This is because the average number of pp collisions per bunch crossing– the so-called pile up, ν– needs to be kept below a value of around 1.5 in LHCb. This provides a cleaner environment in which to study b-hadron decays and allows extremely precise measurements to be made. The LHCb detector can identify b-hadron decays in an event by looking for displaced vertices. Hadrons containing b-quarks have a relatively long lifetime ($\approx 10^{-12}\,\mathrm{s}$), so that particles originating from b-hadron decays come from a vertex that is spatially separated from the pp collision. By exploiting this, the decays of b-hadrons can be studied in detail, and new decays can be searched for. More details of the LHC and the LHCb detector are given in Chapter 3.

1.3 The $\Lambda_{\mathrm{b}}^{0} \to \Lambda^{0}\eta^{(\prime)}$ decays

The LHCb experiment is testing the SM, and searching for NP beyond the SM, by investigating the properties of decays of b-hadrons. This can be achieved in two ways: by precisely measuring the properties of decays that have large branching

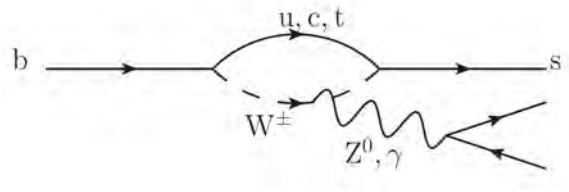

Figure 1.1: Example of a Feynman Diagram for a b → s transition.

fractions, are therefore are not limited by statistical uncertainties; and by searching for "rare" decays and measuring their branching fractions. This is particularly interesting for decays that involve a b-quark decaying into an s-quark, known as a b → s transition. This involves a down-type quark decaying into another down-type quark; an interaction which would take place via an exchange of a neutral particle. In the SM, this is known as a Flavour Changing Neutral Current (FCNC) and is forbidden to occur when only considering tree-level decays. They can therefore only occur through so-called "Penguin" diagrams, an example of which is shown in Figure 1.1, or other higher order processes. The transition can occur through a loop process, and the uncertainty principle allows the loop to gain a large amount of energy, since it only exists for a short amount of time. This means that heavy particles can be produced as virtual particles in the loop, and provides an indirect method of observing particles which are not produced directly in pp collisions. If new heavy particles can be produced in this loop, this will change the properties of the decay, such as the branching fraction, or the angular distribution of the decay products. Therefore, by measuring these properties, and comparing with the predictions from the SM, hints of NP can be observed.

Two examples of baryon decays involving these b → s transitions are $\Lambda_b^0 \to \Lambda^0 \eta'$ and $\Lambda_b^0 \to \Lambda^0 \eta$. Studies of the decays of b baryons are particularly interesting since the LHCb experiment is the first experiment to observe a very large number of Λ_b^0 decays. This means that it is a relatively unexplored field of particle physics, and there are many decays which have not yet been observed or studied in any detail. The $\Lambda_b^0 \to \Lambda^0 \eta^{(\prime)}$ decays[1] have not been observed, but the branching fractions

[1]In this document, the notation $\eta^{(\prime)}$will be used to denote either η or η' mesons.

have been predicted using SM calculations. In addition, these b-baryon decays are particularly interesting since they involve η and η' decays. Studies of these decays can therefore be used to investigate η–η' mixing, a peculiar phenomenon described in more detail in Section 2.2.

This thesis presents a search for these decays and a first measurement of their branching fractions. Chapter 2 describes the theoretical background and the physics motivation for studying this decay. Chapter 3 describes the LHCb detector and how it is used in this analysis, while Chapter 4 describes studies of the simulations used by the LHCb experiment and their impact on the analysis. Finally, full details of the analysis and the results are presented in Chapter 5 and Chapter 6, with the conclusions given in Chapter 7.

CHAPTER

2

PHYSICS

This chapter describes in more detail the theory of heavy flavour physics, and some of the physics motivations for running the LHCb experiment. After describing quarks in terms of the EW interaction of the SM, and how this leads to *CP* violation, the physics motivations for studying the $\Lambda_b^0 \to \Lambda^0 \eta^{(\prime)}$ decays are described, including the phenomenon of η–η' mixing.

2.1 Heavy flavour physics

The study of interactions involving heavy quarks, i.e. t and b quarks, is particularly interesting for investigations into the differences between matter and anti-matter. It is known in the SM that the combined symmetry of Charge conjugation and Parity transformation (*CP*) operation is not conserved in the EW interaction. This means that decays involving the EW interaction of quarks may not be the same

as decays involving the corresponding anti-quarks, and so are referred to as *CP*-violating processes. The extent of *CP* violation (*CPV*) in the SM has been measured, but studies of *CPV* in different decays is necessary to improve the precision of the measurement or to observe new sources of *CPV*.

In order to understand how *CPV* arises in heavy flavour physics, we can look at the Lagrangian for the EW interaction of quarks in the SM:

$$\mathcal{L}_{\mathrm{q}} = \sum_{m=1}^{3} \left(\overline{\mathrm{q}}^{0}_{mL} i \not{D} \mathrm{q}^{0}_{mL} + \overline{\mathrm{u}}^{0}_{mR} i \not{D} \mathrm{u}^{0}_{mL} + \overline{\mathrm{d}}^{0}_{mR} i \not{D} \mathrm{d}^{0}_{mL} \right), \tag{2.1}$$

where the sum is over the three quark generations, $m = 1, 2, 3$. $\overset{(-)}{\mathrm{q}}_{mL}$ are the left-handed quark doublets of the EW SU(2) group,

$$\mathrm{q}_{mL} = \begin{pmatrix} \mathrm{u}_{mL} \\ \mathrm{d}_{mL} \end{pmatrix}, \tag{2.2}$$

where u_{mL} represent the up-type quarks, $\mathrm{u}_1 = \mathrm{u}, \mathrm{u}_2 = \mathrm{c}, \mathrm{u}_3 = \mathrm{t}$; d_{mL} are the down-type quarks, $\mathrm{d}_1 = \mathrm{d}, \mathrm{d}_2 = \mathrm{s}, \mathrm{d}_3 = \mathrm{b}$; and u_{mR} and d_{mR} are the right-handed quark singlets of the EW interaction. The $\not{D}$ matrices are the covariant derivatives using Feynman slash notation,

$$\not{D} = \gamma^{\mu} D_{\mu} \tag{2.3}$$

$$= \gamma^{\mu} \left(\partial_{\mu} + \frac{ig}{2} \vec{\tau} \cdot \vec{W}_{\mu} + \frac{ig'}{2} Y B_{\mu} \right), \tag{2.4}$$

where Y is the hypercharge quantum number, $\vec{\tau}$ are the Pauli matrices, γ^{μ} are the "gamma matrices", and $\vec{W}_{\mu}$ and B_{μ} are the EW gauge fields. These are related to the physical charged $\mathrm{W}^{\pm}$ bosons

$$\mathrm{W}^{\pm}{}_{\mu} = \frac{W^{1}_{\mu} + i W^{2}_{\mu}}{\sqrt{2}}. \tag{2.5}$$

The neutral Z^0 boson and the photon are then a mixture of the other two fields,

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\mathrm{W} & \sin\theta_\mathrm{W} \\ -\sin\theta_\mathrm{W} & \cos\theta_\mathrm{W} \end{pmatrix} \times \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix}, \tag{2.6}$$

where θ_W is the weak mixing angle.

The superscript ^{0}, e.g. on q_{mL}^0, is used to denote the fact that we are considering the flavour eigenstates of quarks, that is, the states that are involved in the weak interaction. This distinguishes from the mass eigenstates, which are the states that propagate through space-time and gain their mass through the Yukawa interaction [21],

$$\mathcal{L}_\mathrm{Yuk} = -\sum_{m,n=1}^{F} \left(\overline{\mathrm{q}}_{mL} G_{mn}^D \Phi \mathrm{d}_{nR} + \overline{\mathrm{q}}_{mL} G_{mn}^U \Phi^c \mathrm{u}_{nR} + h.c. \right), \tag{2.7}$$

where Φ is the Higgs field, and $G^{D,U}$ are matrices of Yukawa couplings for each fermion.

One can see from Equation 2.1 that there is no mixing between the different generations in the EW interaction. Therefore if the flavour eigenstates and mass eigenstates are the same, it is not possible to have interactions involving a transition from a second generation to a first generation quark. However, experiments have observed the $\mathrm{n} \to \mathrm{pe}^-\overline{\nu}_e$ decay, which involves a $\mathrm{d} \to \mathrm{u}$ transition, and also the $\Lambda^0 \to \mathrm{pe}^-\overline{\nu}_e$ decay, which involves an $\mathrm{s} \to \mathrm{u}$ transition. This implies that the mass eigenstates and the flavour eigenstates are misaligned in the EW interaction. This can be accounted for by describing the mass eigenstates of the down-type quarks as a mixture of flavour eigenstates, given by the equation

$$\mathrm{d}_m = \sum_{n=1}^{F} V_{mn} \mathrm{d}_n^0, \tag{2.8}$$

where V is an $F \times F$ matrix representing the mixing between the states. For 3 generations of quarks known in the SM, V is the Cabibbo-Kobayashi-Maskawa (CKM)

matrix, $V_{\rm CKM}$, [22, 23] which is discussed in more detail in the following section.

2.1.1 Measuring the CKM matrix and the unitarity triangle

The values of the elements of $V_{\rm CKM}$ can tell us the extent to which the eigenstates are mixed. In addition, we find that irreducible phases are present in the matrix, which leads to *CPV* in decays involving the third generation quark flavours. By measuring the elements of $V_{\rm CKM}$, one can find the extent of *CPV* introduced by the misalignment of eigenstates in the EW interaction, and predict the branching fractions for different quark transitions.

Many representations of the CKM matrix have been proposed, and the standard representation uses mixing angles and phases,[24]

$$V_{\rm CKM} = \begin{pmatrix} V_{\rm ud} & V_{\rm us} & V_{\rm ub} \\ V_{\rm cd} & V_{\rm cs} & V_{\rm cb} \\ V_{\rm td} & V_{\rm ts} & V_{\rm tb} \end{pmatrix} \tag{2.9}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.10}$$

where $c_{ij} = \cos(\theta_{ij})$, $s_{ij} = \sin(\theta_{ij})$, θ_{ij} is the mixing angle between generations i and j, and δ is the CKM phase which gives rise to *CPV* in quark mixing.

The current status of the measurements of the $V_{\rm CKM}$ elements is presented in Ref. [7] and can be summarised as:

$$|V_{\rm CKM}| = \begin{pmatrix} 0.9743 \pm 0.0002 & 0.2253 \pm 0.0008 & (4.13 \pm 0.49) \times 10^{-3} \\ 0.225 \pm 0.008 & 0.986 \pm 0.016 & (41.1 \pm 1.3) \times 10^{-3} \\ (8.4 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.021 \pm 0.032 \end{pmatrix}. \tag{2.11}$$

Some striking features of this matrix are evident and give an insight into the physics

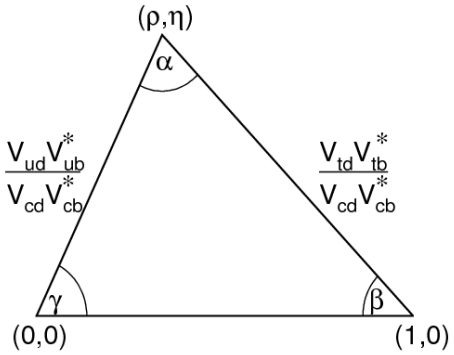

Figure 2.1: The unitarity triangle.

impact of this mixing. The mixing is dominated by the diagonal terms in the matrix, with $|V_{\rm ud}| \approx |V_{\rm cs}| \approx |V_{\rm tb}| \approx 1$. This implies that transitions between quark flavours will be dominated by transitions between quarks in the same generation ($1 \to 1$, $2 \to 2$ and $3 \to 3$ transitions). $V_{\rm CKM}$ then follows a hierarchical structure, with each of $2 \to 1$, $3 \to 2$ and $3 \to 1$ transitions being more suppressed than the previous. This hierarchy can be more conveniently visualised by expressing the CKM matrix using the Wolfenstein parameterisation [25],

$$V_{\rm CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ \lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.12}$$

The parameters are related to the representation in equation 2.11, with

$$\lambda = s_{12} = \frac{|V_{\rm us}|}{\sqrt{|V_{\rm ud}|^2 + |V_{\rm us}|^2}}, \tag{2.13}$$

$$A\lambda^2 = s_{23} = \lambda \left| \frac{V_{\rm cb}}{V_{\rm us}} \right|, \tag{2.14}$$

$$A\lambda^3(\rho + i\eta) = s_{13}e^{i\delta} = V_{\rm ub}. \tag{2.15}$$

The corrections at order $\mathcal{O}(\lambda^4)$ can be added to the matrix by expanding the $V_{\rm CKM}$

elements as follows:

$$
\begin{aligned}
V_{\mathrm{ud}} &= 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 + \mathcal{O}(\lambda^6) \\
V_{\mathrm{us}} &= \lambda + \mathcal{O}(\lambda^7) \\
V_{\mathrm{ub}} &= A\lambda^3(\rho - i\eta) \\
V_{\mathrm{cd}} &= -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] + \mathcal{O}(\lambda^7) \\
V_{\mathrm{cs}} &= 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) + \mathcal{O}(\lambda^6) \\
V_{\mathrm{cb}} &= A\lambda^2 + \mathcal{O}(\lambda^8) \\
V_{\mathrm{td}} &= A\lambda^3[1 - (\rho + i\eta)(1 - \frac{1}{2}\lambda^2)] + \mathcal{O}(\lambda^7) \\
&= A\lambda^3[1 - \overline{\rho} - i\overline{\eta}] + \mathcal{O}(\lambda^7) \\
V_{\mathrm{ts}} &= -A\lambda^2 + \frac{1}{2}A(1 - 2\rho)\lambda^4 - i\eta A\lambda^4 + \mathcal{O}(\lambda^6) \\
V_{\mathrm{tb}} &= 1 - \frac{1}{2}A^2\lambda^4 + \mathcal{O}(\lambda^6),
\end{aligned}
\tag{2.16}
$$

where we have introduced the notation

$$
\overline{\rho} \equiv \rho(1 - \frac{1}{2}\lambda^2), \quad \overline{\eta} \equiv \eta(1 - \frac{1}{2}\lambda^2), \tag{2.17}
$$

which simplifies the expression for V_{td}.

The CKM matrix is unitary, such that ${V_{\mathrm{CKM}}}^{\dagger}V_{\mathrm{CKM}} = \mathbb{1}$. This imposes some conditions on the elements of V_{CKM}, namely $\sum_k V_{ik}V_{jk}^* = \delta_{ij}$. One of these unitarity relations which is used to study *CPV* in the CKM matrix is the fact that the off-diagonal elements should sum to zero, and

$$
V_{\mathrm{ud}}V_{\mathrm{ub}}^* + V_{\mathrm{cd}}V_{\mathrm{cb}}^* + V_{\mathrm{td}}V_{\mathrm{tb}}^* = 0. \tag{2.18}
$$

Dividing by $V_{\mathrm{cd}}V_{\mathrm{cb}}^*$, which is the most precisely known term, we find

$$
\frac{V_{\mathrm{ud}}V_{\mathrm{ub}}^*}{V_{\mathrm{cd}}V_{\mathrm{cb}}^*} + 1 + \frac{V_{\mathrm{td}}V_{\mathrm{tb}}^*}{V_{\mathrm{cd}}V_{\mathrm{cb}}^*} = 0. \tag{2.19}
$$

This condition can be represented as a triangle in complex space, as shown in Figure 2.1, with vertices (0,0), (0,1) and $(\overline{\rho},\overline{\eta})$, with sides of length $|V_{\rm ud}V_{\rm ub}^*|/|V_{\rm cd}V_{\rm cb}^*|$, $|V_{\rm td}V_{\rm tb}^*|/|V_{\rm cd}V_{\rm cb}^*|$ and one. Since this is in complex space, the area of the unitarity triangles give a measure of the amount of *CP* violation present in the CKM matrix [26]. In addition, measurements of these sides, along with the three angles, α, β and γ, is an important test of the SM. Any deviation from this triangle, for example if the three angles do not sum to 180° or the sides do not meet at the expected vertices, would be an indication of new physics not described by the SM. Such conditions could arise if the CKM matrix is not unitary as expected, or if there are more than three generations of quarks. These effects could be extremely small compared with SM behaviour, and so precise measurements of all sides and angles are required.

The measurements of the sides of the triangle simply correspond to the measurement of the individual matrix elements. The most precisely measured matrix element is $V_{\rm ud}$. This is determined from decay rates of super-allowed $0^+ \to 0^+$ nuclear beta decays [27]. These are pure vector transitions, and so are theoretically very clean, and as they proceed through $\mathrm{u} \to \mathrm{d}$ transitions, studies of these decays are used to determine $V_{\rm ud}$. The normalisation of the triangle is determined by $|V_{\rm cd}V_{\rm cb}^*|$, and these elements have been measured precisely using semileptonic decays of b- and c- hadrons. $V_{\rm cd}$ is determined from a measurement of the branching fraction of the semileptonic $\mathrm{D}^{\pm(0)} \to \pi^{0(\pm)}\ell\nu$ decays [28, 29]. The precision of this measurement is limited by the knowledge of the hadronic form factors for the decays, which have been calculated using lattice QCD [30]. Similarly, $V_{\rm cb}$ is measured by studying the decay rates of $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu$ decays [7]. Other measurements of inclusive semileptonic b-hadron to c-hadron decays use theoretical models to eliminate the need to calculate the form factors, and hence have smaller uncertainties. The values presented in Equation 2.11 is an average of these two approaches. The last well known matrix element is $V_{\rm tb}$, which is determined from a measurement of top quark decays [31–37]. Assuming the CKM matrix is unitary, the rate of t-quarks decaying into b-quarks is given by

$$R = \frac{\mathcal{B}(\mathrm{t} \to \mathrm{Wb})}{\mathcal{B}(\mathrm{t} \to \mathrm{Wq})}. \tag{2.20}$$

In contrast, $V_{\rm tb}$ can also be determined without making any assumptions about the unitarity of $V_{\rm CKM}$, using the production cross-section for single top quarks in pp collisions. This is an easier measurement of the top quark to make, and so leads to a more precise value of $V_{\rm tb}$.

The two matrix elements which are not as precisely measured, and therefore limit our knowledge of the sides of the unitarity triangle, are $V_{\rm ub}$ and $V_{\rm td}$. These are the smallest matrix elements, of order $\mathcal{O}(\lambda^3)$, and so are difficult to measure. $V_{\rm ub}$ is determined from inclusive semileptonic decay rates of b-hadrons to u-hadrons, but due to the hierarchy of the CKM matrix, these decays suffer from a large background of semileptonic decays to c-hadrons [38, 39]. To reduce this background, the decays are required to have high lepton momentum, above the kinematic threshold available in decays to charm hadrons. Most recently, LHCb has made a measurement of this using the exclusive baryonic $\Lambda_{\rm b}^0 \to {\rm p}\mu\nu_\mu$ decay, which has a much lower background contribution of $\Lambda_{\rm b}^0 \to \Lambda_{\rm c}^+\mu\nu_\mu$ decays [40]. The CKM element $V_{\rm td}$ can be measured from B^0–$\overline{\mathrm{B}}^0$ mixing, which is discussed in more detail in Section 2.1.3.

The three angles of the unitarity triangle are determined by the phases in the CKM matrix, and given by the relationships

$$\begin{aligned}
\alpha &= \arg\left(-\frac{V_{\rm td}V_{\rm tb}^*}{V_{\rm ud}V_{\rm ub}^*}\right), \\
\beta &= \arg\left(-\frac{V_{\rm cd}V_{\rm cb}^*}{V_{\rm td}V_{\rm tb}^*}\right), \\
\gamma &= \arg\left(-\frac{V_{\rm ud}V_{\rm ub}^*}{V_{\rm cd}V_{\rm cb}^*}\right).
\end{aligned} \tag{2.21}$$

These phases, and hence the angles α, β and γ, can be determined by studying *CPV* in b-hadron decays, which is the principle focus of the LHCb experiment. The following sections detail how these measurements can be made at LHCb.

2.1.2 Direct CP violation

The simplest way of measuring the extent of *CPV* in the CKM matrix is to measure the direct *CPV* in different hadron decays, which manifests itself as a difference in the decay rates of the hadrons compared with the anti-hadron decays. *CPV* was first observed in 1964 with the observation of the $K^0_L \to \pi^+\pi^-$ decays [41]. In order for *CPV* to be present in a decay, specific criteria need to be satisfied. *CPV* arises due to the complex phases in the CKM matrix, which only appear when interactions with the third generation quarks are present; there is no *CPV* in the 2×2 matrix that describes the interactions of the first and second generation quarks. If there is a coupling to a b-quark or t-quark in the decay, then the amplitude for the decay will contain a weak phase, ϕ. The general form of the decay amplitude for the final state f, A_f and corresponding amplitude for the anti-particle, $\bar{A}_{\bar{f}}$, are

$$A_f = |a_f|e^{i(\delta+\phi)}, \tag{2.22}$$

$$\bar{A}_{\bar{f}} = |a_f|e^{i(\delta-\phi)}, \tag{2.23}$$

where $|a_f|$ is the real part of the amplitude which is the same for particle and anti-particle, and δ is the strong phase which arises from QCD interactions of the final state fermions. These amplitudes contribute to the decay rates in the form $|A_f|^2$ and $|\bar{A}_{\bar{f}}|^2$, and so the weak phases will disappear and no difference in the decay rates will be observed. In order to observe *CPV*, the presence of more than one interfering diagram with different strong and weak phases is necessary, so that

$$A_f = |a_1|e^{i(\delta_1+\phi_1)} + |a_2|e^{i(\delta_2+\phi_2)} + ... \tag{2.24}$$

$$\bar{A}_{\bar{f}} = |a_1|e^{i(\delta_1-\phi_1)} + |a_2|e^{i(\delta_2-\phi_2)} + ... \tag{2.25}$$

One interesting decay to look at is the $B^\pm \to {}^{(}\overline{D}{}^{)}K^\pm$ decay. Two possible Feynman diagrams for the amplitudes of these decays are shown in Figure 2.2. The product of the CKM elements involved in each diagram is of order $\mathcal{O}(\lambda^3)$, and so the amplitude of both diagrams will be similar, leading to interference and *CPV*. This can be

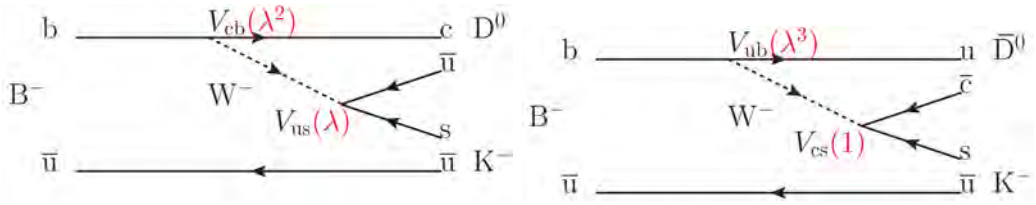

Figure 2.2: Two possible Feynman diagrams for the $B^{\pm} \to {}^{(}\overline{D}{}^{)}K^{\pm}$ decays.

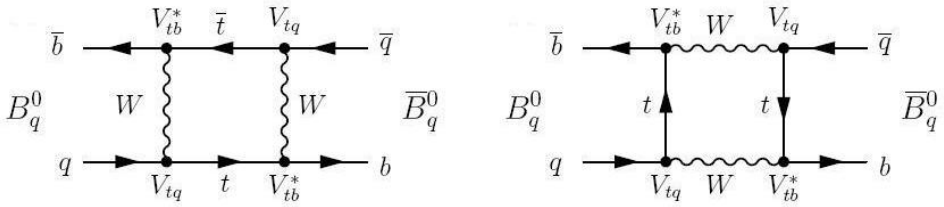

Figure 2.3: Feynman diagrams for $B^0_{(s)}$ mesons to oscillate into $\overline{B}{}^0_{(s)}$ mesons.

measured as an asymmetry in the decay width of B^+ decays, $\Gamma(B^+)$ compared with the decay width of B^- decays, $\Gamma(B^-)$, defined as

$$\mathcal{A}^{CP} = \frac{\Gamma(B^+) - \Gamma(B^-)}{\Gamma(B^+) + \Gamma(B^-)}. \tag{2.26}$$

Due to the presence of the V_{ub} matrix element in the Feynman diagram, the weak phase of these decays is related to the unitarity angle, γ. Therefore, measurements of $\mathcal{A}^{CP}$ in these decays will lead to a measurement of γ.

2.1.3 Mixing induced *CP* violation

Measurements of direct *CPV* in $B^{\pm}$ decays is straightforward as the charge of the final state particles must correspond to the charge of the initial b-hadron. The situation is not so straightforward when investigating the decays of neutral B^0 mesons, as it is experimentally more difficult to determine whether the decaying particle is a B^0 or $\overline{B}{}^0$ meson. However, there are interesting effects in B^0 mesons which make them useful to study, most notably the phenomenon of B^0–$\overline{B}{}^0$ mixing, where a B^0 meson can oscillate into a $\overline{B}{}^0$ meson and vice versa [42]. This happens through a box diagram,

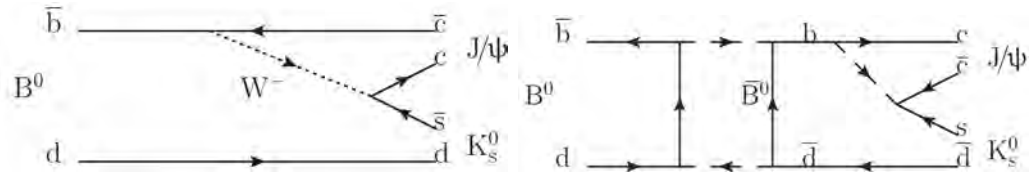

Figure 2.4: Two possible Feynman diagrams for the $B^0 \to J/\psi K^0_S$ decay.

shown in Figure 2.3. This mixing leads to a difference in the *CP* eigenstates of the B^0 mesons, and the eigenstates of the weak interaction, in an analogous process to the K^0 system. The *CP* eigenstates, $|B^0\rangle$ and $|\overline{B}^0\rangle$, are described by a mixture of the weak eigenstates, $|B_H\rangle$ and $|B_L\rangle$, where

$$|B_H\rangle = p|B^0\rangle + q|\overline{B}^0\rangle, \quad \text{and} \quad |B_L\rangle = p|B^0\rangle - q|\overline{B}^0\rangle. \tag{2.27}$$

Looking at the time evolution of these states, one finds that the system oscillates between B^0 and $\overline{B}^0$ mesons with a frequency of Δm, the difference between the masses of the $|B_H\rangle$ and $|B_L\rangle$ states. This oscillation frequency is determined by the size of the CKM matrix elements present in the box diagrams. For the B^0–$\overline{B}^0$ system, this frequency is proportional to $|V^*_{td}V_{tb}|^2$, and a measurement of this frequency, along with theoretical inputs of the constants of proportionality, can provide a measurement of one side of the unitarity triangle.

In addition, for decays of B^0 mesons into *CP* eigenstates, the mixing provides an extra Feynman diagram for the decay which introduces some *CPV*. This is shown in Figure 2.4 for the $B^0 \to J/\psi K^0_S$ decay. The diagram on the left is the dominant tree diagram for this decay. The diagram on the right shows the B^0 meson oscillating into a $\overline{B}^0$ before decaying in the $J/\psi K^0_S$ final state. Because the J/ψ and K^0_S are *CP* eigenstates, the final state is indistinguishable for the two diagrams. However, the oscillation involves the extra CKM matrix elements V_{td} and V_{tb}, and so there is a difference in the weak phase between the two diagrams. Interference between these decays introduces an $\mathcal{A}^{CP}$ which varies with time as the B^0 oscillates to $\overline{B}^0$ and back again, and is called a time dependent *CP* asymmetry. For a decay into a

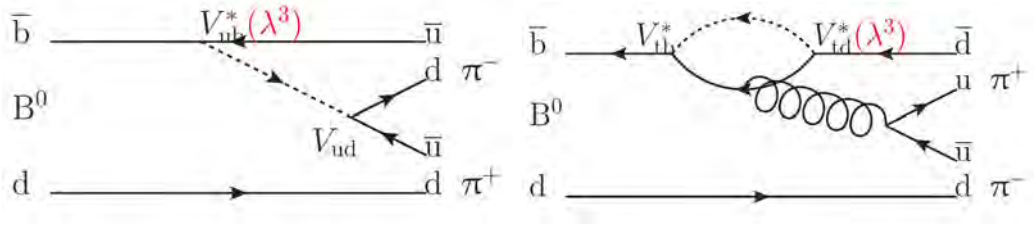

Figure 2.5: Two possible Feynman diagrams for the $B^0 \to \pi^+\pi^-$ decay.

CP eigenstate this asymmetry has the form

$$\mathcal{A}^{CP}(t) = \frac{\Gamma(B^0)(t) - \Gamma(\bar{B}^0)(t)}{\Gamma(B^0)(t) + \Gamma(\bar{B}^0)(t)} \tag{2.28}$$

$$= \mathcal{C}_f \cos(\Delta m t) - \mathcal{S}_f \sin(\Delta m t) \tag{2.29}$$

The coefficient $\mathcal{C}_f$ describes the extent of direct *CPV* in the decay, and the coefficient $\mathcal{S}_f$ describes the extent of mixing induced *CPV* in the decay. For the $B^0 \to J/\psi K^0_S$ decays, the penguin diagrams do not contribute any direct *CPV*, and so in the SM we have $\mathcal{C}_f = 0$. Any *CPV* arises from the interference of the two diagrams in Figure 2.4, which have the same final state and so any strong interaction effects will cancel. This is therefore a clean channel which is sensitive to the angle β, with $\mathcal{S}_f = \sin(2\beta)$. These coefficients have recently been measured in LHCb as [43]

$$\mathcal{S}_f = 0.731 \pm 0.035\,(\text{stat}) \pm 0.02\,(\text{syst}) \tag{2.30}$$

$$\mathcal{C}_f = -0.038 \pm 0.032\,(\text{stat}) \pm 0.005\,(\text{syst}). \tag{2.31}$$

These results can be compared with the world average which is dominated by the measurements from the B-factories [44]

$$\mathcal{S}_f = 0.687 \pm 0.028\,(\text{stat}) \pm 0.012\,(\text{syst}) \tag{2.32}$$

$$\mathcal{C}_f = 0.024 \pm 0.020\,(\text{stat}) \pm 0.016\,(\text{syst}). \tag{2.33}$$

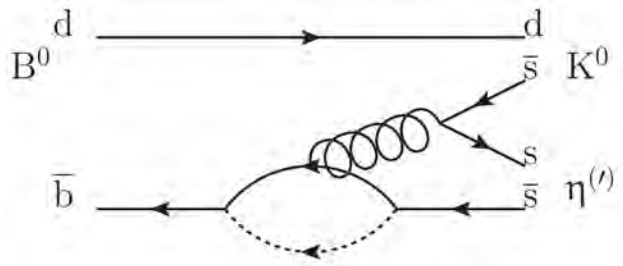

Figure 2.6: Feynman diagram for the $B^0 \to K^0_S \eta'$ decay.

2.1.4 *CP* violation in charmless B meson decays

In order to measure the final angle of the unitarity angle, α, decays that involve the interference between amplitudes with the CKM elements $V^*_{tb}V_{td}$ and $V^*_{ub}V_{ud}$ are studied. This can be accessed in decays of B^0 mesons into final states with no charm mesons, for example $B^0 \to \pi^+\pi^-$. V_{ub} is accessible through the $b \to u$ transition shown in the tree diagram of Figure 2.5. The phase α can in principle be measured through the *CPV* introduced by the diagram resulting from B^0–$\overline{B}^0$ mixing which allows access to the $V^*_{tb}V_{td}$ matrix elements. Studies of these decays are, however, very challenging. One reason for this is that the amplitude for the $b \to u$ transition is very low ($\mathcal{O}(\lambda^3)$), and so any observations of this decay will be limited by available statistics. In addition, theoretical treatment of the results is complicated by the presence of a diagram for the $b \to d$ transition resulting in the same final state. This is shown in the penguin diagram in Figure 2.5, which has a similar amplitude to the tree diagram and introduces some direct *CPV* with a different phase. In order to extract α from these decays it is necessary to measure the *CP* asymmetries in $B^0 \to \pi^0\pi^0$ and $B^\pm \to \pi^\pm\pi^0$ to isolate the contributions from the different strong and weak phases in these decays. This has been done by the B-factories, and the coefficients of the mixing induced *CPV* in this case are found to be [44]

$$\mathcal{S}_{\pi^+\pi^-} = -0.66 \pm 0.07 \tag{2.34}$$

$$\mathcal{C}_{\pi^+\pi^-} = -0.30 \pm 0.05. \tag{2.35}$$

Measurements of *CPV* in charmless b-meson decays are also very useful for providing complimentary measurements of the angle β. Similar to the analysis of $B^0 \to J/\psi K^0_S$, one can look at b-meson decays to charmless *CP* eigenstates, such as $B^0 \to K^0_S \eta'$. The dominant diagram for this decay is shown in Figure 2.6, and *CPV* can arise due to the weak phase β in the mixing diagram. The interest in measuring β using these decays is that the diagram for the decay is a penguin diagram, which includes a loop dominated by the top or charm quark in the SM. However, heavy particles are predicted to exist in new physics models, and these could enter into the loop and change the amplitude for the process. If these new particles have a different *CPV* structure to SM particles, this could introduce a different weak phase to the diagram. This would lead to a measurement of $\sin(2\beta)$ which is inconsistent with the measurement from charmonium modes, and so would indicate the presence of new physics.

The current status of measurements of the unitarity triangle is shown in Figure 2.7, using the frequentist approach of the CKMFitter collaboration [45]. The angles have been measured by fitting the triangle using inputs from the measured CKM matrix elements and constraints from direct measurements of the angles, and the best fit values are found to be

$$\begin{aligned} \beta &= (21.89^{+0.74}_{-0.77})^\circ \\ \gamma &= (67.08^{+0.97}_{-2.17})^\circ \\ \alpha &= (91.0^{+2.3}_{-1.2})^\circ. \end{aligned} \tag{2.36}$$

These can be compared with the world averages for the direct measurements of these angles [46],

$$\begin{aligned} \beta_{\rm dir} &= (21.50^{+0.75}_{-0.74})^\circ \\ \gamma_{\rm dir} &= (73.2^{+6.3}_{-7.0})^\circ \\ \alpha_{\rm dir} &= (87.7^{+3.5}_{-3.3})^\circ. \end{aligned} \tag{2.37}$$

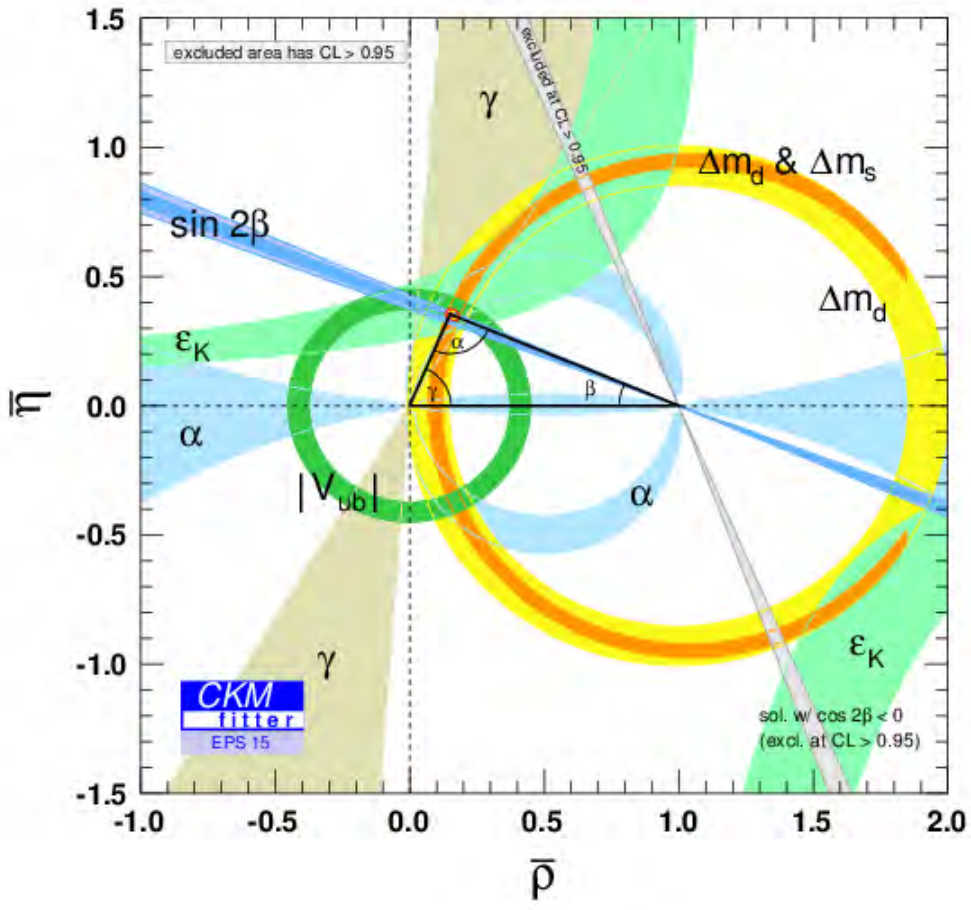

Figure 2.7: The current status of measurements of the unitarity triangle from the CKM-Fitter collaboration [45].

These measurements provide stringent tests of the CKM matrix in the SM. The value of $\sin(2\beta)$ measured using B meson decays into charmonium final states ($\mathcal{S}_{b\to c\bar{c}s} = 0.682 \pm 0.019$) [46] is compatible with the measurement from charmless final states ($\mathcal{S}_{b\to q\bar{q}s} = 0.655 \pm 0.032$) [46]. Unitarity of the CKM matrix has been measured from the CKM elements [7],

$$\begin{aligned} |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 &= 0.9999 \pm 0.0006 \\ |V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 &= 1.024 \pm 0.032 \\ |V_{ud}|^2 + |V_{cd}|^2 + |V_{td}|^2 &= 1.000 \pm 0.004 \\ |V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 &= 1.025 \pm 0.032. \end{aligned} \tag{2.38}$$

The sum of three angles in the unitarity triangle is $\alpha + \beta + \gamma = (175 \pm 9)^\circ$. Improvements to the measurements of the three angles will provide a better test of whether these angles sum to 180°. So far, no deviations from the SM expectations have been observed, but there is still scope for discrepancies to arise as the uncertainties are improved. Finally, the total amount of *CPV* in the CKM matrix can be calculated using the Jarlskog invariant, J, given by

$$J = \sum_{m,n} \epsilon_{ikm}\epsilon_{jln} Im[V_{ij}V_{kl}V_{il}^*V_{kj}^*], \tag{2.39}$$

which is twice the area of the corresponding unitarity triangle. This has been measured for the triangle in Figure 2.7 to be $J = (2.96^{+0.19}_{-0.17}) \times 10^{-5}$. With all the sources of *CPV* known in the SM, this is still a factor 10^9 too small to explain the observed difference in the amount of matter compared with anti-matter in the universe.

2.2 η–η' mixing

One of the exciting surprises to come from the study of light, neutral mesons is the concept of a mixing of singlet and octet states and the breaking of SU(3) flavour symmetry, most notably with the η and η' mesons. The physical η and η' particles

are defined by the mixing of SU(3) singlet ($|\eta_0\rangle$) and octet ($|\eta_8\rangle$) states, characterised by the mixing parameter θ_P:

$$\begin{pmatrix} |\eta\rangle \\ |\eta'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta_P & -\sin\theta_P \\ \sin\theta_P & \cos\theta_P \end{pmatrix} \begin{pmatrix} |\eta_8\rangle \\ |\eta_0\rangle \end{pmatrix}, \tag{2.40}$$

where

$$|\eta_0\rangle = \frac{1}{\sqrt{3}}|u\bar{u} + d\bar{d} + s\bar{s}\rangle \tag{2.41}$$

and

$$|\eta_8\rangle = \frac{1}{\sqrt{6}}|u\bar{u} + d\bar{d} - 2s\bar{s}\rangle. \tag{2.42}$$

Due to the symmetry of SU(3), it is possible, and often more convenient, to represent the particles using a different basis. Here we use the quark flavour basis, defined in Ref. [47], with the two flavour states:

$$|\eta_q\rangle = \frac{1}{\sqrt{2}}|u\bar{u} + d\bar{d}\rangle \tag{2.43}$$

$$|\eta_s\rangle = |s\bar{s}\rangle. \tag{2.44}$$

The $\eta^{(\prime)}$mesons are then a mixture of a light quark state and a strange quark state, defined with a different mixing angle ϕ_P, by:

$$\begin{pmatrix} |\eta\rangle \\ |\eta'\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi_P & -\sin\phi_P \\ \sin\phi_P & \cos\phi_P \end{pmatrix} \begin{pmatrix} |\eta_q\rangle \\ |\eta_s\rangle \end{pmatrix}. \tag{2.45}$$

This representation shows how the different proportions of the strange quark state, $|\eta_s\rangle$ gives rise to the different masses of the two mesons, since the mass of the strange quark is much larger than the mass of the up or down quarks. Transformations between these two bases are simple due to a relation between the mixing angles [47]

$$\theta_P = \phi_P - \tan^{-1}(\sqrt{2}). \tag{2.46}$$

Many phenomenological and experimental studies have been carried out to determine the values of the mixing angles, and measurements are in the range $\phi_{\mathrm{p}} \approx 38 - 46^{\circ}$ [48–60]. This amount of mixing of the quark flavours is surprisingly large when compared to the mixing in equivalent particles in other SU(3) nonets. For example, in the nonet of vector mesons, the mixing of light and heavy flavours is small, with $|\omega\rangle \approx |\eta_{\mathrm{q}}\rangle$ and $|\phi\rangle \approx |\eta_{\mathrm{s}}\rangle$.

Using this quark flavour basis, one must in principle consider all possible SU(3) flavour singlet states to contribute to the mixing, including the gluonic wavefunction, $|\mathrm{gg}\rangle$, and the heavier quarkonia wavefunctions, $|\mathrm{c\bar{c}}\rangle$ and $|\mathrm{b\bar{b}}\rangle$. Since the mass scale of the gluonic wavefunction is much smaller than the $\mathrm{c\bar{c}}$ mass scale, only the former is considered to contribute to the mixing of the mass eigenstates. However, due to the much smaller mass of the η meson, it is assumed that the gluonic contribution to the η mass eigenstate is negligible [47]. The gluonic component is then introduced into the η' wavefunction through a new mixing angle ϕ_{G}:

$$|\eta\rangle \approx \cos\phi_{\mathrm{P}}|\eta_{\mathrm{q}}\rangle - \sin\phi_{\mathrm{P}}|\eta_{\mathrm{s}}\rangle \tag{2.47}$$

$$|\eta'\rangle \approx \cos\phi_{\mathrm{G}}\sin\phi_{\mathrm{P}}|\eta_{\mathrm{q}}\rangle + \cos\phi_{\mathrm{G}}\cos\phi_{\mathrm{P}}|\eta_{\mathrm{s}}\rangle + \sin\phi_{\mathrm{G}}|\mathrm{gg}\rangle \tag{2.48}$$

The gluonic component is small, has been measured in Refs. [58–60], and found to be consistent with zero, albeit with large uncertainties. Most recently, the mixing angles have been measured by the LHCb experiment using $\mathrm{B}^0_{(\mathrm{s})} \to \mathrm{J/\psi}\,\eta^{(\prime)}$ decays, and are found to be $\phi_{\mathrm{p}} = (43.5^{+1.5}_{-2.8})^{\circ}$ and $\phi_{\mathrm{G}} = (0 \pm 25)^{\circ}$.

The main consequence of this mixing is the difference in branching fractions for B decays to η and η'. This gluonic mixing leads to an enhanced branching fraction for the decays to η' compared with the equivalent decay to η. For example, the branching fraction for the decay $\mathrm{B}^0 \to \mathrm{K}^0\eta'$ has been measured by the BaBar collaboration to be $(6.85 \pm 3.8) \times 10^{-5}$ [61] compared with the branching fraction for $\mathrm{B}^0 \to \mathrm{K}^0\eta$, which has been measured to be $(1.15^{+1.00}_{-0.98}) \times 10^{-6}$ (over 50 times smaller!). The same trend has been seen in many other B decays to $\eta^{(\prime)}$, for example

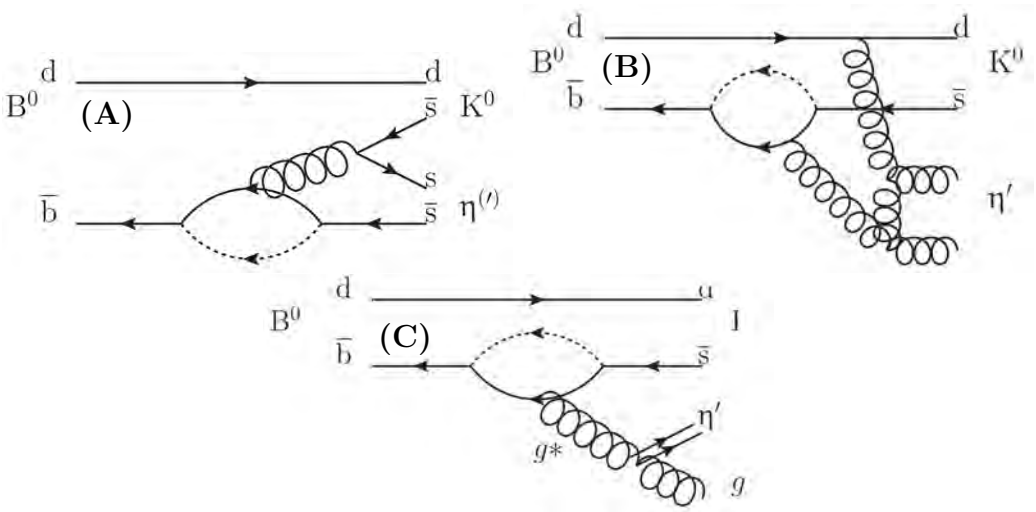

Figure 2.8: Feynman diagrams for B decays into $\eta^{(\prime)}$mesons, showing the enhanced branching fraction to decays to the η' due to the gluonic contribution to the wavefunction.

in $B^+ \to K^+\eta'$ (branching fraction $= (71.5 \pm 1.3 \pm 1.2) \times 10^{-6}$) compared with $B^+ \to K^+\eta$ (branching fraction $= (2.94^{+0.39}_{-0.34} \pm 0.21) \times 10^{-6}$) [61]. The reason for this is due to the gluonic contribution of the η' wavefunction, shown in Figure 2.8 for the $B^0 \to K^0\eta^{(\prime)}$ decays. Figure 2.8 A) shows the Feynman diagram for the $B^0 \to K^0\eta^{(\prime)}$ decays through the $b \to s$ loop transition. Figure 2.8 B) shows the non-spectator contribution, where a gluon is radiated from the spectator quark and forms the η' through the gluonic wavefunction. Figure 2.8 C) shows the η' being produced via the so-called "anomalous" coupling between the η' and a gluon [62]. Since diagrams 2.8 B) and 2.8 C) are only available through the gluonic component of the wavefunction, the interference between these diagrams leads to different branching fractions for decays to η' and η. If the interference is constructive, then the branching fraction will be enhanced for the η' decay, relative to the η. If it is destructive then this can lead to comparable, or lower, branching fractions for the η'.

2.3 The $\Lambda_b^0 \to \Lambda^0 \eta^{(\prime)}$ decays

No baryonic decays to η' or η mesons have been observed to date. Measurements of decays such as these would be particularly interesting in understanding the compo-

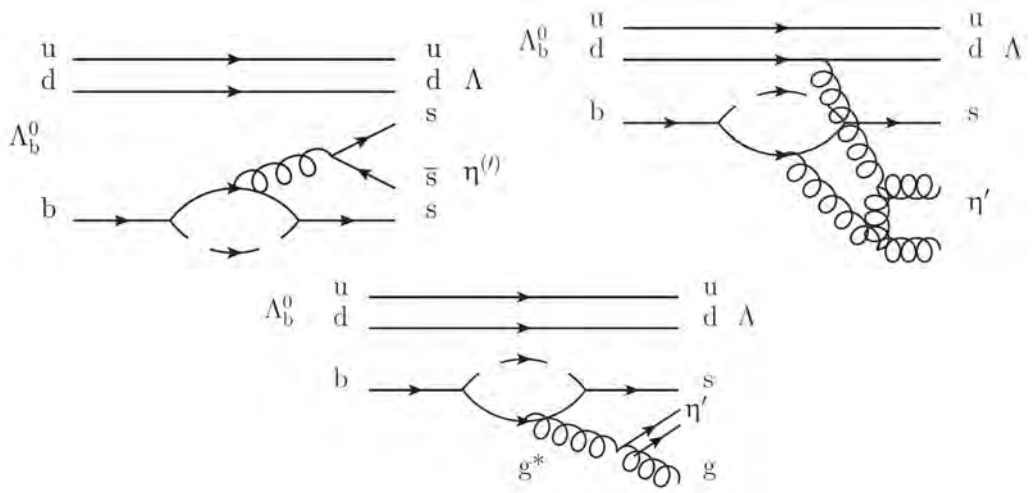

Figure 2.9: Feynman diagrams for Λ_b^0 decays into $\eta^{(\prime)}$mesons, showing the enhanced branching fraction to decays to the η' due to the gluonic contribution to the wavefunction.

sition of the $\eta^{(\prime)}$mesons, and would provide complementary information to analogous studies of b-mesons. One example of decays which can be used to study η–η' mixing are the charmless $\Lambda_b^0 \to \Lambda^0\eta'$ and $\Lambda_b^0 \to \Lambda^0\eta$ decays. Some possible Feynman diagrams for these decays are shown in Figure 2.9. The process is very similar to the corresponding b-meson decay shown in Figure 2.8, with an extra spectator quark from the Λ_b^0 baryon. Theoretical predictions have been made for the branching fractions of these decays, and the results of the predictions of Ref. [63] are presented in this section.

In order to calculate the branching fraction for these decays, one needs to measure the decay amplitude, $\mathcal{M}$, given by

$$\mathcal{M} = \langle \Lambda^0 \eta^{(\prime)} | H_{eff} | \Lambda_b^0 \rangle, \tag{2.49}$$

where H_{eff} is the effective hamiltonian for the decay. For a general decay involving a $b \to sq\overline{q}$ transition, the effective hamilitonian is given by

$$H_{eff} = \frac{4G_F}{\sqrt{2}} \left[V_{ub}V_{us}^*(c_1O_1^u + c_2O_2^u) + V_{cb}V_{cs}^*((c_1O_1^c + c_2O_2^c) - V_{tb}V_{ts}^* \sum_{i=3}^{11} c_iO_i \right] + h.c. \tag{2.50}$$

where c_i are the Wilson Coefficients for operators O_i, which are determined by the

short distance scale corrections to the operations. The operators O_1 and O_2 describe the tree level $b \to q$ decay, O_{3-6} are the operators which describe the strong penguin diagrams where a gluon is exchanged in the loop, O_{7-10} are the electroweak penguin operators which describe the exchange of a γ or Z^0 boson or the contribution from mixing diagrams. Finally, the operator O_{11} is the chromomagnetic operator which describes the gluon splitting into two quarks.

In order to calculate the matrix elements for the $\Lambda_b^0 \to \Lambda^0 \eta^{(\prime)}$ decays, the authors use the factorisation approach to parameterise each amplitude in the decay process. For the transitions to $\eta^{(\prime)}$mesons, the decay constants $f_{\eta^{(\prime)}}^q$ are needed, where

$$\langle 0|\overline{q}\gamma^\mu\gamma_5 q|\eta^{(\prime)}\rangle = i f_{\eta^{(\prime)}}^q p_{\eta^{(\prime)}}^\mu \quad (q = u, s). \tag{2.51}$$

Since the $\eta^{(\prime)}$mesons are described by a mixing of the η_0 and η_8 states, the physical decay constants are related to the decay constants of these states by [64, 65]

$$\begin{aligned} f_\eta^u &= \frac{f_8}{\sqrt{6}}\cos(\theta_8) - \frac{f_0}{\sqrt{3}}\sin(\theta_0) \\ f_\eta^s &= -2\frac{f_8}{\sqrt{6}}\cos(\theta_8) - \frac{f_0}{\sqrt{3}}\sin(\theta_0) \\ f_{\eta'}^u &= \frac{f_8}{\sqrt{6}}\sin(\theta_8) + \frac{f_0}{\sqrt{3}}\cos(\theta_0) \\ f_{\eta'}^u &= -2\frac{f_8}{\sqrt{6}}\sin(\theta_8) + \frac{f_0}{\sqrt{3}}\cos(\theta_0). \end{aligned} \tag{2.52}$$

This is useful since the mixing angles have been measured phenomenologically to be $\theta_8 = -21^\circ$ and $\theta_0 = -9.2^\circ$ [65] and the decay constants are found to be $f_8 = 166\,\text{MeV}$ and $f_0 = 154\,\text{MeV}$ [66], so the decay to the $\eta^{(\prime)}$mesons is completely described.

The $\Lambda_b^0 \to \Lambda^0$ transition can be factorised into the vector and axial-vector matrix

elements describing the b$\to$ s transition, which are parameterised as

$$\begin{aligned}\langle\Lambda^0|\bar{s}\gamma_\mu b|\Lambda_{\rm b}^0\rangle &= \bar{u}_{\Lambda^0}\left[f_1\gamma_\mu + i\frac{f_2}{m_{\Lambda_{\rm b}^0}}\sigma_{\mu\nu}q^\nu + \frac{f_3}{m_{\Lambda_{\rm b}^0}q_\mu}\right]u_{\Lambda_{\rm b}^0}\\ \langle\Lambda^0|\bar{s}\gamma_\mu\gamma_5 b|\Lambda_{\rm b}^0\rangle &= \bar{u}_{\Lambda^0}\left[g_1\gamma_\mu\gamma_5 + i\frac{g_2}{m_{\Lambda_{\rm b}^0}}\sigma_{\mu\nu}q^\nu\gamma_5 + \frac{g_3}{m_{\Lambda_{\rm b}^0}q_\mu\gamma_5}\right]u_{\Lambda_{\rm b}^0},\end{aligned} \tag{2.53}$$

where u_X are the Dirac spinors for the baryons, $q^\mu = p^\mu_{\Lambda_{\rm b}^0} - p^\mu_{\Lambda^0}$ is the difference in four-momenta of the $\Lambda_{\rm b}^0$ and Λ^0 baryons, and f_i and g_i are the hadronic form factors for the decay.

This transition can be factorised in a different way using the Heavy Quark Expansion Theory (HQET), where the matrix element for the $\Lambda_{\rm b}^0\to\Lambda^0$ transition can approximated using a parameterisation of the form [67]

$$\langle\Lambda^0|\bar{s}\Gamma b|\Lambda_{\rm b}^0\rangle = \bar{u}_{\Lambda^0}\left[F_1(q^2) + \not{v}F_2(q^2)\right]\Gamma u_{\Lambda_{\rm b}^0}, \tag{2.54}$$

where v is the four-velocity of the $\Lambda_{\rm b}^0$ baryon ($v = p_{\Lambda_{\rm b}^0}/m_{\Lambda_{\rm b}^0}$), and F_i are the form factors in this representation. They are related to the form factors in the factorisation approach by

$$\begin{aligned}F_2 = f_2 = f_3 = g_2 = g_3,\\ F_1 + rF_2 = f_1 = g_1,\end{aligned} \tag{2.55}$$

where $r = m_{\Lambda^0}/m_{\Lambda_{\rm b}^0}$. The form factors can be calculated in the HQET parameterisation, and used to calculate the branching fraction for these decays. The theoretical uncertainties on these hadronic form factors give rise to the dominant uncertainty on the predicted branching fraction. In Ref. [63] the form factors are calculated using two different approaches: the QCD sum rules method [68] and the pole model [67,69]. The results from each method are summarised in the following sections.

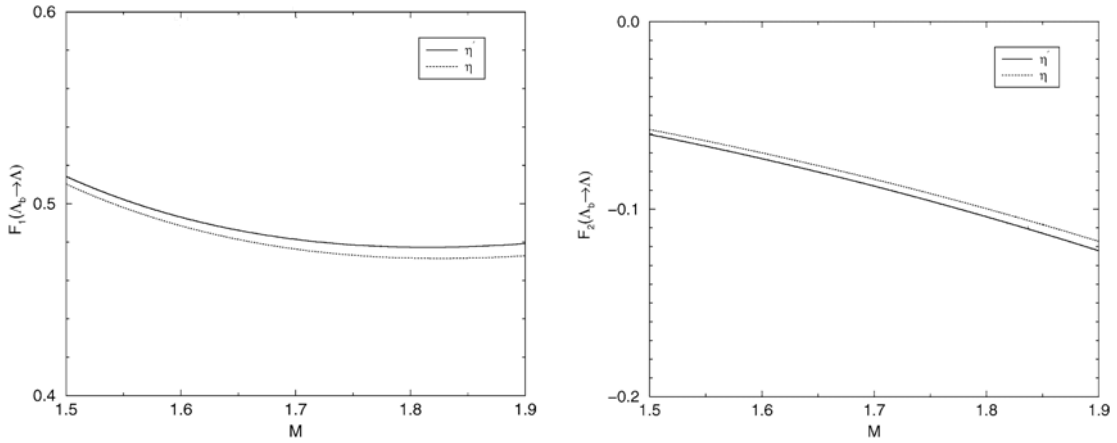

Figure 2.10: Hadronic form factors as a function of the Borel parameter for $\Lambda_b^0 \to \Lambda^0 \eta'$ and $\Lambda_b^0 \to \Lambda^0 \eta$ decays. [63]

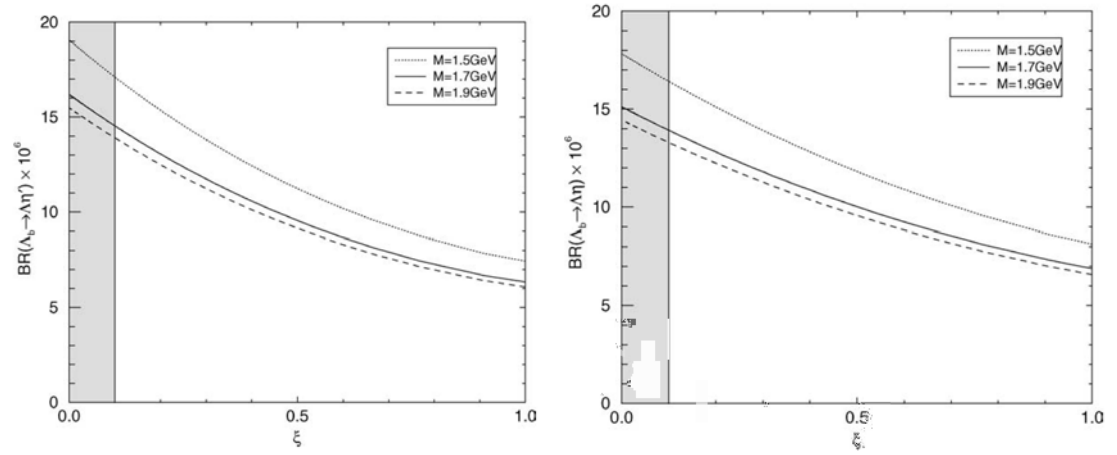

Figure 2.11: Predicted branching fraction for $\Lambda_b^0 \to \Lambda^0 \eta'$ (left) and $\Lambda_b^0 \to \Lambda^0 \eta$ (right) decays as a function of $\xi = 1/N_c$ for different values of the Borel parameter. [63]

2.3.1 Prediction from QCD sum rules

The form factors F_1 and F_2 are calculated for the $\Lambda_b^0 \to \Lambda^0\eta'$ and $\Lambda_b^0 \to \Lambda^0\eta$ decays using the QCD sum rules approach and are found to have a strong dependence on the Borel mass parameter, $M = 4T/m_b$, which is a parameter introduced in the QCD sum rules expansion. The dependence can be seen in Fig. 2.10. The effect of the uncertainty on the form factors is evident in Fig. 2.11 which shows the predicted branching fraction of each decay for three values of the Borel parameter, as a function of $\xi = 1/N_c$, where N_c is the effective number of colours in the interaction. This gives a range of predicted branching fractions from the QCD sum rules method of

$$\begin{aligned} \mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta') &= (6.0 - 19.0) \times 10^{-6} \\ \mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta) &= (6.5 - 17.9) \times 10^{-6} \end{aligned} \tag{2.56}$$

Analysis of many hadronic decays of B mesons favour values of $\xi \approx 0.1$ and $M \approx 1.7\,\text{GeV}$, which gives $\mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta') = 15 \times 10^{-6}$ and $\mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta) = 14 \times 10^{-6}$. It is interesting to note that the predictions for the $\Lambda_b^0 \to \Lambda^0\eta'$ and $\Lambda_b^0 \to \Lambda^0\eta$ decays are similar, in contrast with b-meson decays which have very different branching fractions to η and η'.

2.3.2 Prediction from pole model

In the pole model, the form factors are given by the simple expression

$$F_i(q^2) = N_i \left(\frac{\Lambda_{QCD}}{\Lambda_{QCD} + z} \right)^2, \tag{2.57}$$

where $\Lambda_{QCD} \approx 200\,\text{MeV}$, $z = p_{\Lambda^0} \cdot p_{\Lambda_b^0}/m_{\Lambda_b^0}$, $N_1 = 52.32$ and $N_2 = -13.08$. This expression gives form factors which are much smaller than those obtained using the QCD sum rules, and so the predictions for the branching fractions are smaller. They

are estimated to be

$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta') = (1.8 - 4.5) \times 10^{-6}$$
$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta) = (1.8 - 3.8) \times 10^{-6}. \quad (2.58)$$

With $\xi = 0.1$, the branching fractions would be $\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta') = 4.08 \times 10^{-6}$ and $\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta) = 3.55 \times 10^{-6}$.

2.3.3 Anomaly contribution

The anomaly contribution described in Section 2.2 has been taken into account in the prediction of the branching fractions by introducing a term in the matrix element proportional to

$$\langle \eta^{(\prime)} | \bar{s} \gamma_5 s | 0 \rangle = i \frac{(f^s_{\eta^{(\prime)}} - f^u_{\eta^{(\prime)}}) m^2_{\eta^{(\prime)}}}{2 m_s} \quad (2.59)$$

It is found that this contribution has a large effect on the predicted branching fractions for these decays, through interference with other terms. By using the hadronic form factors calculated using the QCD sum rules approach, the branching fractions are predicted to be $\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta') = 40 \times 10^{-6}$ and $\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta) = 8 \times 10^{-6}$. This shows a clear enhancement in the branching fraction for the $\Lambda_b^0 \to \Lambda^0 \eta'$ decay, and a reduction in the prediction for the $\Lambda_b^0 \to \Lambda^0 \eta$ decay. When including the anomaly term in the calculation of the $\Lambda_b^0 \to \Lambda^0 \eta'$ branching fraction, this term interferes destructively with the terms from the penguin decays, thus reducing the branching fraction. For the $\Lambda_b^0 \to \Lambda^0 \eta$ decay, the anomaly term enters with the opposite sign compared with the $\Lambda_b^0 \to \Lambda^0 \eta'$ decay. This has the effect of destructively interfering with the penguin amplitudes, increasing the predicted branching fraction. The anomaly term has a very large effect on the predicted branching fractions for these decays, and so an experimental measurement of these decays will provide insight into the validity of the anomolous contribution to $\eta^{(\prime)}$decays.

2.3.4 Summary

Theoretical predictions have been made for the branching fractions of the $\Lambda_b^0 \to \Lambda^0\eta^{(\prime)}$ decays by factorising the effective Hamiltonian for the decays. Two methods have been used to calculate the hadronic form factors relevant for the $\Lambda_b^0 \to \Lambda^0$ decay, and these suffer from large theoretical uncertainties. The branching fractions are therefore predicted to be in the range

$$\begin{aligned} \mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta') &= (1.8 - 19.0) \times 10^{-6}, \\ \mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta) &= (1.8 - 17.9) \times 10^{-6}. \end{aligned} \tag{2.60}$$

CHAPTER

3

DETECTOR

This chapter describes in more detail the LHC accelerator complex as well as the LHCb detector and the various subdetectors of which it is composed. Each subdetector is optimised to measure a particular quantity or particle. The general physics requirements that motivated the design choices of the LHCb detector will be discussed, as well as the LHC complex. After describing the workings of each subdetector, the use of these detectors in tracking and triggering will be described. Full details of the LHCb detector and its performance can be found in Refs. [70, 71].

3.1 Requirements of the LHCb detector

The first requirement of the LHCb detector is to have a high acceptance for b-hadrons produced in pp collisions to maximise the statistics of the physics channels studied. At the collision energies of the LHC, b-quarks are produced at very small

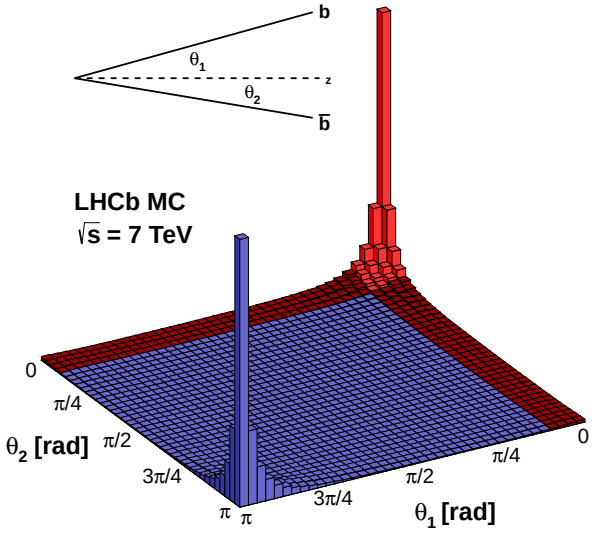

Figure 3.1: Correlation in production angle of b and $\overline{\text{b}}$ quarks in pp collisions. θ_1 and θ_2 are the angle between the momentum vector and the z-axis for the b-quark and $\overline{\text{b}}$-quark respectively.

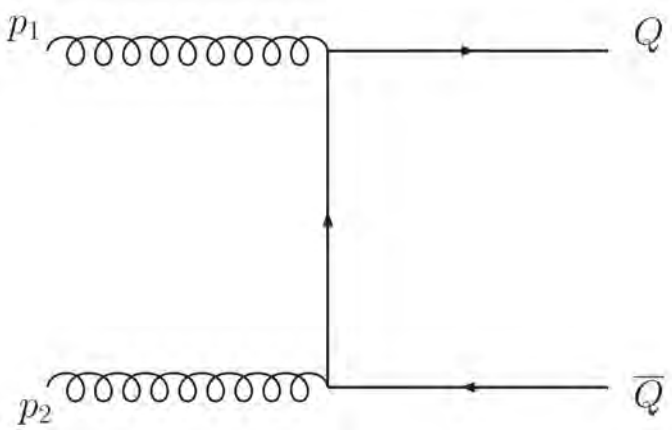

Figure 3.2: Production mechanism of b and $\overline{\text{b}}$ quarks in pp collisions.

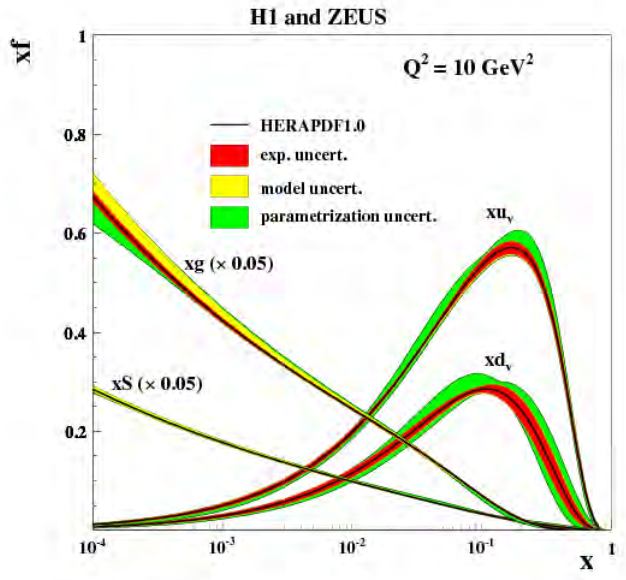

Figure 3.3: Parton distribution function of gluons in a proton [72].

angles to the beam direction and the b and $\overline{\text{b}}$ are generally produced in the same direction. This is shown in Figure 3.1 which plots the production angle of the b- and $\overline{\text{b}}$- quarks with respect to the beam direction from a Monte Carlo (MC) simulation of pp collisions. It can be seen that the production of the two quarks is highly correlated at low angles. This is due to the production mechanism of b-quark pairs in pp collisions, which is dominated by gluon-gluon fusion, shown in the Feynman diagram of Figure 3.2. For this diagram, the fraction of the proton energy carried by the gluon, also referred to as the Bjorken-x variable, x, is given by

$$x_1 = \frac{m_T}{\sqrt{s}}(e^{y_Q} + e^{y_{\bar{Q}}}), \tag{3.1}$$

$$x_2 = \frac{m_T}{\sqrt{s}}(e^{-y_Q} + e^{-y_{\bar{Q}}}),$$

for each of the two gluons, where m_T is the transverse mass, $m_T \equiv \sqrt{m_Q^2 + {p_\text{T}}^2}$, m_Q is the mass of the quark, p_T is the transverse momentum of the quark and y is the rapidity of the quark or antiquark [73]. Figure 3.3 shows the parton density function of gluons, i.e. the fraction of gluons containing a Bjorken-x value as a function of x. There is a high probability that the gluon has a low Bjorken-x value, and so from

Equation 3.1 and the large mass of the b-quark, it can be seen that heavy quarks are produced with low values of p_{T} and y. These quarks will then form hadrons which travel in approximately the same direction as the b-quark, and the LHCb detector can be only instrumented in the forward region to exploit these production properties to reduce the cost of the experiment.

Another requirement is to identify hadrons containing b-quarks amongst the background of light hadrons. One way this is done by LHCb is to look for vertices which are displaced from the Primary Vertex (PV), the position of the initial pp collision. There should be more than 5 tracks originating from this vertex, and so this is how it is identified. Compared to most particles produced in collisions, b-hadrons have, on average, longer lifetimes. This means that they can travel $\approx$3 mm before decaying, and can be identified by looking for tracks which originate from a displaced vertex, some distance from the PV. In order to distinguish displaced vertices, tracks need to be reconstructed with very high precision.

Lots of decays which are of interest to the LHCb collaboration contain hadrons. In order to study these in detail, it is important to be able to identify the hadron species, and in particular to distinguish between pions, kaons and protons, and the detector needs to have good Particle Identification (PID) capabilities.

Finally, the detector must be able to measure the energy and momentum of a range of particles with high precision.

3.2 Overview of the Large Hadron Collider complex

The beams of protons used in the LHC are accelerated to their collision energies through a series of smaller accelerators. Figure 3.4 shows the overview of the accelerator complex at CERN, including the intermediate accelerators used for the LHC beams. The first stage accelerates protons from a hydrogen source using a linear accelerator, Linac2. The accelerator consists of a series of Radio-Frequency (RF)

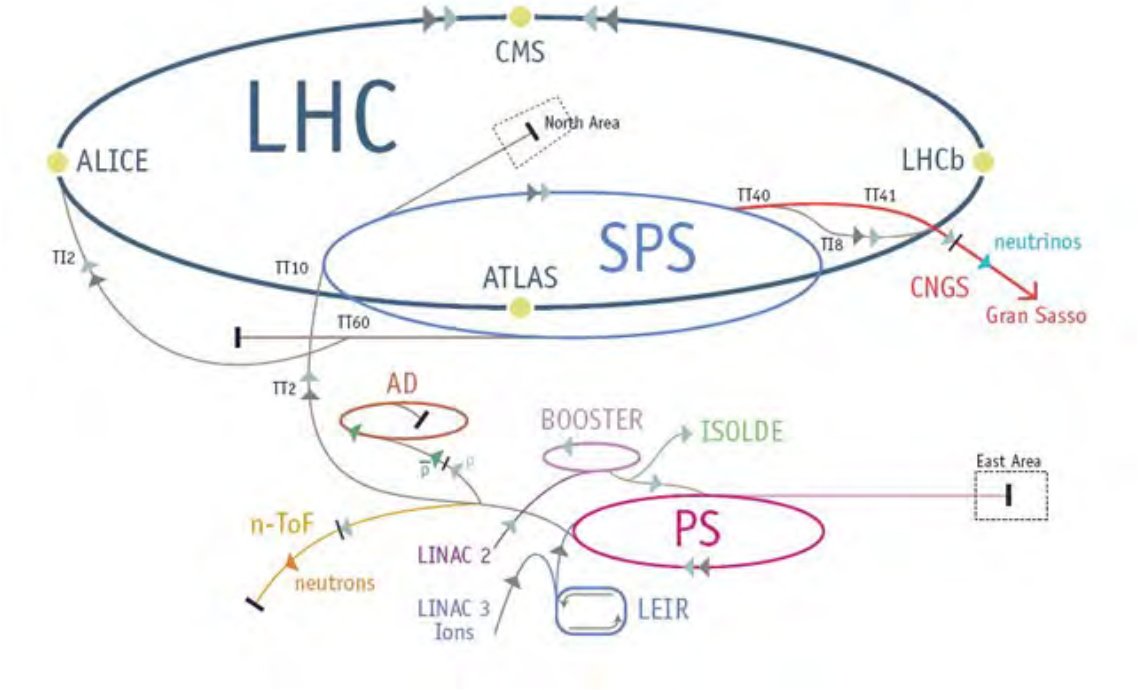

Figure 3.4: The LHC complex.

cavities that use alternating electric fields to attract the protons as they approach the cavity, and repel them as they leave the cavity. This accelerates the protons and leaves them in bunches rather than a continuous beam. At the end of Linac2, the protons have an energy of 50 MeV and are injected into a series of synchroton accelerators. These are circular accelerators which use magnetic fields both to guide the proton bunches in a circular orbit and to focus them, and RF cavities to increase their energy. The protons gain energy in each cycle of the accelerator, and so the magnetic field must be synchronised with the protons' cycle to keep the path circular. The 50 MeV protons from Linac2 are injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV. They are then accelerated in the Proton Synchrotron (PS) to 25 GeV and in the Super Proton Synchrotron (SPS) to 450 GeV before being injected into the LHC where they are accelerated to a maximum energy of 4 TeV.

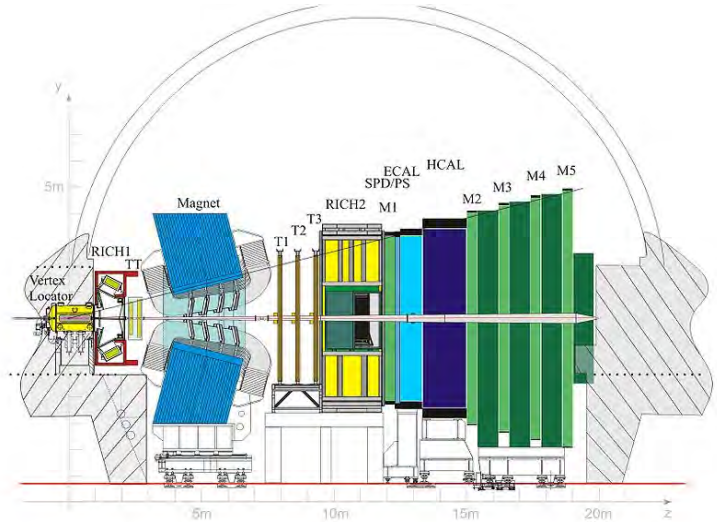

Figure 3.5: Schematic overview of the LHCb detector.

3.3 Overview of the LHCb detector

Figure 3.5 shows a schematic overview of the LHCb detector. The detector is a forward arm spectrometer, meaning it measures particles travelling in the very forward directions. To this end, the LHCb detector is arranged in layers, and covers an angular range of 10 mrad to 300 mrad in the plane defined by the x-z axes in Figure 3.5, and 250 mrad in the orthogonal longitudinal plane. This is very effective for studying the decays of particles containing b-quarks, since in pp collisions b-quarks are produced at very small angles to the beam direction, as discussed in Section 3.1.

The LHCb detector contains a tracking system, which is designed to reconstruct the trajectories of charged particles, and hence to measure their momenta and points of origin. This system consists of the VErtex LOcator (VELO), the Silicon Tracker (ST) and the Outer Tracker (OT), which are separated by a dipole magnet that produces an integrated magnetic field of 4 Tm. For PID, the LHCb detector makes use of: the Ring Imaging CHerenkov (RICH) detectors to identify protons, kaons, and pions; the calorimeters to identify electrons, photons and π^0 mesons; and the muon detector to identify muons. In addition, the calorimeters provide measurements of the energies of particles, and the muon stations measure the momenta of muons. The proton beams circulate in the LHC in a vacuum to avoid beam-gas interactions,

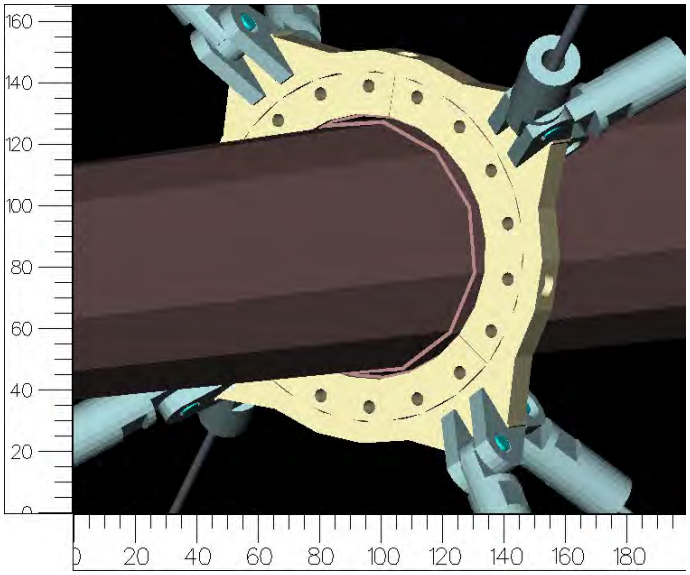

Figure 3.6: Beam pipe supports.

and this is maintained within the LHCb beam pipe. The requirements of the beam pipe are that it is strong enough to contain the vacuum at a pressure of 10^{-8} to 10^{-9} mbar, while keeping the material budget to a minimum. As primary tracks traverse material, energy loss and scattering processes take place, which may produce secondary particles. This has the effect of deteriorating the performance of the LHCb detector, and so the total material seen by the primary particles must be taken into account. For these reasons, beryllium has been chosen for the beam pipe material, since it is strong, but has a large radiation length of 35 cm, meaning particles are less likely to interact as they pass through. The thickness of the beam pipe ranges from 1 mm close to the interaction point, to 2.4 mm further downstream. For stability, the beam pipe is held in various places by the supports shown in Figure 3.6. These are made from carbon fibre, and are designed to be strong enough to hold the beam pipe whilst keeping the material budget to a minimum.

3.4 The vertex locator

The first piece of the detector seen by particles produced in collisions is the VELO [74], which is part of the LHCb tracking system. The aim of the VELO is to track par-

ticles coming from the collision in order to reconstruct the vertex from which they originated.

The VELO is made of layers of silicon strip sensors with a pitch of about 40 μm. The silicon is doped such that a conduction band and valence band are set up. As a charged particle traverses the silicon, electrons are excited from the valence band into the conduction band, leaving a hole in the valence band. The number of electron-hole pairs is proportional to the amount of energy deposited by the ionising radiation. The advantage of using semiconductors over gaseous ionisation detectors is that more electrons are produced per unit energy deposited. The sensors are biased, which creates an electric field to separate the electrons and holes before they recombine, and the electrons are collected at an anode which detects a pulse [75].

Each layer contains strips of silicon orientated to measure either the radial position or the angular position of hits, referred to as r-sensors and ϕ-sensors respectively. The VELO detector is arranged in two halves, arranged to the left and right of the beam line, with each half consisting of alternating layers of r- and ϕ-sensors. By following the hits in each layer of the VELO, the tracks of charged particles can be reconstructed and extrapolated back towards the pp collision point to calculate their origin vertex. In order to do this accurately, it is better to extrapolate the tracks as little as possible, and so the VELO needs to be close to the interaction region. This must be balanced by the amount of radiation the VELO is exposed to: closer to the beam there is a much larger dose of radiation, which damages the sensitive silicon of the VELO. In addition, when the beam is injected into the LHC ring, the width of the beam is relatively large, and would overlap with the space occupied by the VELO sensors. For this reason, the VELO has been designed to be movable, so that when the beams are being accelerated around the ring, the two halves of the VELO sit 3 cm away from the beam. When stable beam has been declared by the LHC, the VELO is moved in so that the first layer is just 8 mm away from the collisions.

In order to keep the occupancy in the VELO to a minimum, there is no beam pipe in the VELO region. Instead, the sensors are mounted in an aluminium box which

contains the VELO within the vacuum. The two halves of the detector are separated by an aluminium foil, which is corrugated to allow the sensors to overlap when brought into the closed position, and whose purpose is to protect the VELO from the intense electromagnetic field, the "wakefield", associated with the LHC beams.

Since the VELO can reconstruct the displacement of vertices from the PV with a resolution of $(15+29/p_{\mathrm{T}})$ μm it is ideally suited to identifying the displaced vertices of b-hadrons, and reducing systematic uncertainties for precision measurements.

3.5 The tracking system

The tracking detectors in LHCb are designed to measure the momentum of charged particles. The system consists of the ST and the OT [76]. The ST is composed of four stations of two different trackers: one station of the Tracker Turicensis (TT) and three stations of the Inner Tracker (IT). Each of the ST stations is composed of four layers of silicon microstrips, which track charged particles as they travel through the stations.

The ST is designed to cope with the high occupancy of particles coming from the pp collision. Due to the high density of charged particles in a typical event, the silicon strips are separated by 200 μm in order to distinguish hits from individual particles. Using this pitch gives a spacial resolution of 50 μm for each track.

The TT is located just downstream of the VELO, and the IT is located downstream of the dipole magnet. By measuring the curvature of charged particles in the magnetic field, the momentum of the particles can be determined.

The OT consists of three straw tube drift chambers, which are positioned around the outside of the three IT stations. These chambers contain a gas which is a mixture of Argon (70%) and CO_2 (30%), which becomes ionised when charged particles travel through. The ionisation then drifts towards two layers of drift tubes where it creates an electric signal. By measuring the drift time of the charge, and combining with

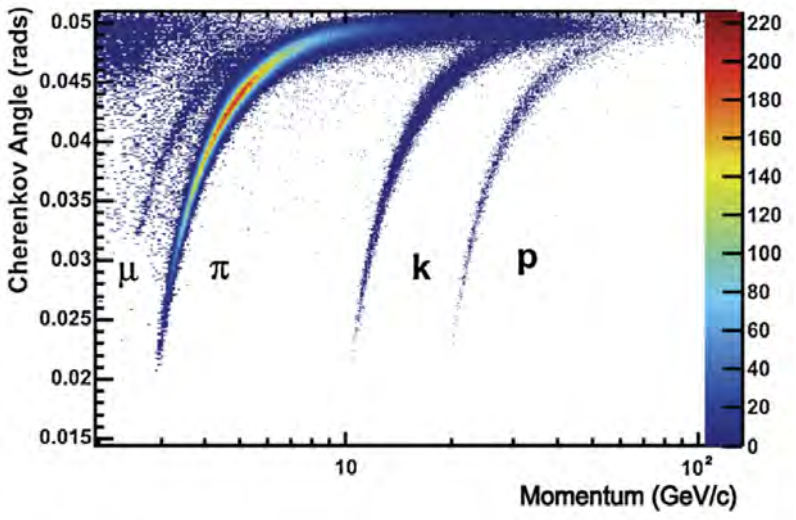

Figure 3.7: Cherenkov angle vs. particle momentum using C_4F_{10} as a RICH radiator [77].

the wire which measured the charge, the coordinates of the charged particle can be calculated, and then tracked through the stations to measure the momentum [75]. The mixture of gases in the chamber is chosen to ensure a short drift time, typically less than 50 ns, and a precise coordinate resolution of 200 μm.

Combining all parts of the LHCb tracking system, the momentum of charged particles can be measured with a resolution of $\sigma p/p = 0.4\%$ at low momentum, increasing to 0.6% at 100 GeV/c.

3.6 The RICH detectors

The PID in the LHCb detector is largely performed with two RICH detectors, labelled RICH1 and RICH2 [77].

RICH1 is located upstream of the dipole magnet, between the VELO and the TT and covers an angular range of ±25 mrad to ±300 mrad (horizontal) or ±250 mrad (vertical), which is the full angular acceptance of the LHCb detector. It uses Cherenkov light from aerogel and C_4F_{10} radiators. When a charged particle travels faster than the speed of light in the medium through which it is travelling, it emits a cone of

light. The angle of the cone is given by the equation

$$\cos(\theta_c) = \frac{1}{n\beta} \tag{3.2}$$

where $\beta = v/c$ is the speed of the particle relative to the speed of light, and n is the refractive index of the radiator. The relationship between the Cherenkov angle and the momentum of the particle is therefore different depending on the mass of the particle and so plotting this relationship gives some discrimination between protons, kaons and pions. Radiators with different refractive indexes show good separation in different regions of particle momentum. In order to measure the Cherekov angle, the cones of light need to be converted into rings. This is done by placing convex mirrors at the end of the RICH detectors. The Cherenkov light is detected using Hybrid Photo-Detectors (HPDs) in the focal plane of the mirror, and so the angle of the cone of light can be measured from the radius of the observed ring of light. Knowing the momentum of the particle from the tracking stations, the particle can be identified according to its position on the Cherenkov angle versus momentum curve. These relationships are shown for different particles using C_4F_{10} as a radiator in Figure 3.7. For the radiators used in RICH1, there is good separation in the momentum range $1 - 60\,\mathrm{GeV}/c$.

RICH2 is located downstream of the magnet, and has an angular coverage of ± 15 mrad to ± 120 mrad (horizontal) or ± 100 mrad (vertical). It has a lower angular coverage than RICH1, since it is designed to provide PID for particles in the high momentum range from $\approx 15\,\mathrm{GeV}/c$ to above $100\,\mathrm{GeV}/c$, and high momentum particles are produced at lower angles. RICH2 uses a CF_4 radiator, which provides good separation of pions and kaons at higher momenta.

3.7 The calorimeters

There are two main jobs of the calorimeter system in LHCb: the first is to identify electrons, photons and hadrons and measure their energy, and the second is to select

high energy particles for the trigger system (described in Section 3.10). To do this, the calorimeter is split into two sections, the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL) [78].

The ECAL consists of alternating layers of lead and plastic scintillator. The lead is designed to cause electrons and photons to deposit their energy and produce an electromagnetic shower. The shower is sampled in each layer of scintillator so that the energy of the shower is measured at each stage of the shower. This allows the original energy of the electron or photon to be measured. The energy resolution of the ECAL is [79]

$$\frac{\sigma(E)}{E} \approx \frac{10\%}{\sqrt{E(GeV)}} \oplus 1\%. \tag{3.3}$$

The other important job of the ECAL is to separate electrons from neutral pions for the trigger. In order to do this, a small section of the ECAL is placed before the first layer of lead, as a Pre-Shower detector (PRS) detector. In front of the PRS is a Scintillating Pad Detector (SPD), which only detects charged particles. So electrons will be detected by the SPD, but photons will not. Since π^0 mesons decay into 2 photons, the PRS allows neutral pions to be separated from 2 photons, by seeing how many particles were present before the shower.

The HCAL is designed to measure the energy of hadrons, which will not deposit much energy in the ECAL. Iron is used as the absorber material here, as it is more dense than lead, and will cause the hadrons to shower. Since hadronic showers are harder to measure, the energy resolution is not as good as the ECAL. The resolution of the HCAL is

$$\frac{\sigma(E)}{E} = \frac{(69 \pm 5)\%}{\sqrt{E(GeV)}} \oplus (9 \pm 2)\%. \tag{3.4}$$

3.8 The muon detectors

The final part of the LHCb detector is the muon detector [80]. This is another important part of the detector as muons play a major role in the key decays of

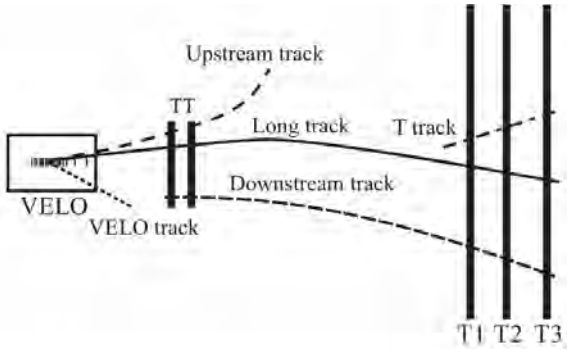

Figure 3.8: Summary of the different track types in the LHCb detector.

interest to LHCb, such as the search for $B^0{}_{(s)} \to \mu^+\mu^-$ and $B^0_s \to J/\psi\,\phi$ decays, and so this detector is used for both offline analysis, and to trigger on high energy muons. There are five muon chambers (labelled M1-M5). The first muon station (M1), is placed upstream of the calorimeters, and used as a momentum measurement for the trigger. Muons are minimum ionising particles, so they do not deposit much of their energy in the calorimeters. Four of the five muon chambers (M2-M5) are therefore placed downstream of the calorimeters, so that only muons are observed here. Each chamber is separated by 80 cm thick iron plates to select the most penetrating muons. The chambers themselves are Multi-Wire Proportional Chambers (MWPCs) [75], which track charged particles as they go through. Chambers M1-M3 have a good spacial resolution for tracking particles. Stations M4 and M5 have poorer spatial resolution, and are there to select the most penetrating muons.

3.9 Track reconstruction

Tracks are reconstructed in the LHCb detector using the tracking stations described in this chapter, by fitting the hits observed in each layer. The quality of the track is then defined by the χ^2 of this fit, and good tracks are defined as having a track χ^2 less than 3. This cut is chosen to have a high efficiency for reconstructing tracks (>95%), whilst removing ghost tracks [81]. Ghost tracks are tracks which result from random signals produced in the tracking stations, but which happen to line up in such a way as they can be reconstructed as real tracks. To further reduce the rate

of ghost tracks, a cut can be placed on the ghost probability of each track, which uses information from all tracking stations to produce a multivariate classifier which defines a probability that the track is the result of random noise in the detector.

Once reconstructed, the tracks are then classified according to the subdetectors used to reconstruct them. These different track types are shown in Figure 3.8, and can be summarised as:

1. **Long Tracks:** These are tracks which pass through the VELO and all tracking stations. These are useful because they contain the maximum amount of information about the track and so lead to better reconstructed resolutions.

2. **Downstream Tracks:** These are tracks which do not leave hits in the VELO, only in the ST or OT.

3. **Upstream Tracks:** These leave hits in the VELO and the TT but not in the other tracking stations. They are typically low momentum particles and so are not generally from signal decays.

4. **Velo Tracks:** These leave hits only in the VELO. These are not useful for momentum measurements.

5. **T tracks:** These leave hits only in the Tracking stations downstream of the magnet. These are typically not useful as they don't contain the detailed information obtained from the VELO and TT.

Clearly, the most useful track to use is the long track, since it uses information from all tracking detectors. However, if the tracks originate from long lived particles, they could travel some distance in the detector before decaying, and therefore will produce downstream tracks.

3.10 The trigger system

During 2011 and 2012, the LHCb experiment had an interaction rate of 20 MHz. However, the detector can only record events at a rate of 5 kHz. The job of the LHCb trigger is to reduce the rate of events to an acceptable level by selecting only interesting events [82]. Interesting events in LHCb typically have high energy leptons, photons or hadrons which originate from a displaced vertex, and this is the basis for how they are selected by the trigger. Trigger decisions also need to be made in a short amount of time, since collisions occur every 50 ns. This requirement has important consequences for which parts of the detector can be used for triggering, and indeed it had an impact on the choices for each sub detector. The trigger system in LHCb is split into two levels, Level 0 (L0) and a Higher Level Trigger (HLT). L0 is a hardware trigger, and HLT is a software trigger, which is again split into two levels. Each level will be described in more detail in this section. Full details of the performance of the LHCb trigger are described in [83].

3.10.1 The L0 trigger

The L0 trigger is a hardware trigger, meaning that is uses raw signals from the detector to make a trigger decision, and is designed to reduce the event rate from 20 MHz to 1 MHz. The L0 uses information from the calorimeters and muon chambers to reconstruct L0 objects to be filtered through to the first level of HLT, HLT1.

The calorimeters are used to select high energy particles by summing the energy deposit in clusters of 2×2 segments of the calorimeter. Information from the calorimeters is then used to classify the highest energy cluster into one of five L0 objects: L0Photon, L0Electron, L0LocalPi0, L0GlobalPi0 or L0Hadron. A further L0 object is found by looking for high momentum tracks in the muon chambers, and these are classified as L0Muon objects.

A cluster of energy in the ECAL can be classified as an electron, photon or π^0.

Only the cluster in the ECAL with the highest summed energy is considered as a L0 candidate. A cluster with a track in both the PRS and the SPD is classified as an electron, and a cluster with a track in the PRS but not the SPD is classified as a photon. π^0 mesons are identified by the presence of two photons in the ECAL, with only one hit in the PRS. They are classified as either "local π^0" or "global π^0". The ECAL sections are grouped into different front-end (FE) electronics boards, which handle the summing of the energy in the cells, and the trigger decision, with each FE board processing 32 ECAL cells. If both of the photons are handled by the same FE board, it is classified as a local π^0, and if the two photons are in ECAL cells corresponding to different FE boards, it is classified as a global π^0.

L0 Hadron objects are identified by energy deposits in 2×2 cells in the HCAL which are over threshold. All L0 candidates are then passed onto the high level trigger for further processing.

3.10.2 The HLT1 trigger

The L0 candidates are sent to the online computer farms to be processed by the high level trigger, which reduces the rate of events from 1 MHz to 5 kHz for recording. The HLT has access to all information from the LHCb detector, including tracking information from the VELO, ST and OT, and so particles and events can be fully reconstructed. The processing of the offline trigger is then split into two levels, HLT1 and HLT2. HLT1 consists of several lines which use information from the tracking stations to confirm the decisions made at level 0. As well as this, further requirements are placed on the quality of the track, the transverse momentum and the impact parameter of the track in order to reduce the number of background non $b\bar{b}$ events.

The main HLT1 line used by this analysis is the HLT1TrackAllL0 line [84]. This line takes all charged L0 candidates and triggers if the decision is confirmed and the tracks pass the requirements described above. Other lines have also been studied for

this analysis. The HLT1TrackPhoton line is designed specifically for confirming the L0 photon decision. Since it is possible for photons to trigger as an electron at level 0, this line takes both L0Photon and L0Electron candidates. However, the photons present in this analysis are lower energy than the photons from radiative decays, and so do not pass the tight energy constraints of the trigger line. This means that, while this line is useful for radiative decays, it is not very efficient at selecting events for this analysis.

In addition, a HLT1 line exists to confirm the L0Muon decision for muon candidates.

3.10.3 The HLT2 trigger

The next stage of the HLT trigger involves reconstructing the entire decay in an event, and looking at the topology of the whole event, in order to select only events with interesting decays. There are many HLT2 lines in LHCb designed to trigger on particular interesting decays. The lines which are useful for this analysis are the general 2,3 and 4 body topological (topo) decay lines, known as HLT2Topo2,3,4BodyBBDT [85]. This uses a multivariate technique, a Boosted Decision Tree (BDT), which combines information using several variables to decide if an event is signal or background. The general working of a BDT is described in more detail in Chapter 5. This trigger line reduces the total rate of events to 2 kHz so that they can be recorded to tape, and analysed offline.

In addition, HLT2 lines exist to select radiative decays, decays of charmed mesons and muons, which have an output rate of 3 kHz, making the total output rate of the LHCb trigger 5 kHz.

CHAPTER

4

SIMULATION

This chapter describes the work performed towards the simulation of the LHCb detector using the GEANT4 framework. This includes the integration of basic GEANT4 code which is used to test and compare the physics lists, which describe particle interactions according to their energy, in different versions of GEANT4. In addition, a detailed study was performed to investigate the effect of the cuts applied in GEANT4 on the distribution of hits observed in the LHCb detector. The chapter includes an introduction to GEANT4 and how it fits into the LHCb simulation model, before presenting the results of the comparison of physics lists in Section 4.2 and the results of the studies of GEANT4 cuts in Section 4.3.

4.1 Introduction: Simulation of events in LHCb

One of the most useful and widely used tools in High Energy Physics (HEP) is the simulation of physics processes using Monte Carlo (MC) techniques. Since processes in particle physics are probabilistic in nature, it is not possible to calculate exactly what will happen in a given pp collision. However, the MC method is a way of predicting what will happen, on average, for many collisions. The method involves using pseudo-random numbers to model a given process by sampling a probability density function which defines the behaviour of the process. By repeating this sampling using many pseudo-random numbers, the overall behaviour of the system can be predicted.

In HEP, this technique is used at various stages to predict what is likely to happen in an event, from the initial pp collision through to the reconstructed events used in a physics analysis. Each stage of the event requires very different input parameters and calculations, and so the simulation is separated into three stages: the production of b-hadrons, the decay of the hadron of interest, and the simulation of the detector.

A program called Pythia [86] is used to simulate the initial pp collision, and hence the production of the b-hadron and particles produced in the soft QCD processes which occur in the collision, also known as the underlying event. In order to save computing resources, it is necessary to ensure that the specific type of b-hadron we are interested in is produced in each simulated event. If this is not the case, then we must repeat the simulation until the b hadron is produced, and to avoid wasting CPU time. In addition to simulating a pp collision, the settings in Pythia can be changed to simulate a set of one or more specified processes. In the case of the study of b-hadron decays in LHCb, it is necessary to simulate only the process $gg \to b\bar{b}$, which is the dominant process for the production of b-quarks in pp collisions. A process called "repeated hadronisation" is then used to ensure the b-quark forms the particular hadron of interest. For example, if we are interested in producing a B^0 meson, then the hadronisation of the b-quark is repeated until it hadronises with

a d-quark to form a B^0 meson. In this way, the underlying event needs only to be simulated once, and we are guaranteed to produce the hadron of interest.

LHCb is particularly interested in the decay of the b-hadron, and this is modelled by EvtGen [87], which will force the b-hadron to decay into the daughters of interest for a particular analysis and provide information on the daughter products. These daughters can be produced with kinematics which are distributed uniformly across the available phase space, or according to some expected kinematic distribution, and the event will be modelled through the whole decay chain. This process clearly depends on the kinematics of the b-hadron, and so Pythia and EvtGen must work together to produce the entire event, and information from both steps can be used to save CPU time by only producing events which result in all daughters being produced within the acceptance of the LHCb detector. Full details of how Pythia and EvtGen are set up and used in LHCb can be found in Ref. [88].

The output from the production stage of the simulation is usually referred to as generator level MC. This is MC for which the full decay chain has been produced, and the kinematics of the signal and the underlying event are known, but the resolution of the detector is not accounted for, and so the MC cannot be compared with data. Simulation of the detector is done using a dedicated program called Geant4, which uses a model of the LHCb detector, with an accurate description of the distribution of materials, to simulate the behaviour of a particle as it traverses each subdetector. By simulating the interactions of all particles in both the active and passive parts of the detector, one can predict how the particles will be detected by LHCb, including resolution effects, and trigger efficiencies. In this way, MC is produced in exactly the same form as data, and so a direct comparison of the two could be performed to look for any deviations in the data from the expected behaviour.

If any such deviations are observed, then it is important to know whether the differences are due to genuine new physics effects, or to problems in the modelling of the process. In order to understand the modelling as completely as possible, detailed studies are carried out comparing data and MC for processes which are well known.

This can include "benchmark" physics processes which have been extensively studied, such as mass distributions for well-established decays, or by comparing the distribution of hits or energies observed in each part of the LHCb detector. If these distributions do not match, then it is necessary to understand which part of the simulation chain– the production of particles or modelling of the detector– is responsible for the discrepancies. One issue is that the overall number of hits observed in the detector depends on the number of tracks in the underlying event. This is modelled by PYTHIA, and if it is incorrectly estimated then this will effect the distributions observed. This chapter focusses on investigations of the effect of the modelling of the detector on the agreement between data and MC. If the particles are simulated to interact in the detector in a way which is different from the real behaviour, then this will affect the reconstruction of particles, and hence the distribution of particles observed. The interaction of particles with materials in GEANT4 is handled by "physics lists", which describe the physics of various processes in different energy regimes. The physics lists used by LHCb are discussed and compared in Section 4.2. In order to save time and resources at the simulation stage, it is necessary to apply cuts to the tracks produced by PYTHIA and by interactions in the detector. In this way, the simulation does not model the progress and interactions of extremely low energy particles, which would not be observed in the real data due to trigger requirements or detection thresholds. The danger of this is that the GEANT4 cuts will be too tight, and so tracks that would be observed in real data will not be modelled, which would affect the distributions observed. Section 4.3 describes a detailed study of the effect of these GEANT4 cuts on the agreement of distributions between data and MC, and also on the time taken to simulate each event, which is of critical importance for mass production of simulated events within LHCb.

4.2 Physics Lists in GEANT4

4.2.1 Available lists

A number of physics lists are available in GEANT4, each of which is optimised for a different purpose. The electromagnetic physics lists describe the interactions of electrons and photons in the detector material, and studies have been performed by members of the LHCb collaboration to choose the most appropriate list suited to our needs. This section describes similar studies carried out for the hadronic physics lists, which describe the interactions of hadrons in materials. There are three physics lists designed to simulate hadronic interactions in HEP, and these are called LHEP, FTFP_BERT, and QGSP_BERT.

In previous versions of the LHCb simulation software, LHEP has been the hadronic physics list which has been used. This list models hadronic interactions based on the energy of the hadrons. Instead of detailing each interaction and enforcing energy conservation in each step, this list uses a parameterisation which is tuned to data to reproduce shower shapes and high energy behaviour of hadrons. Separate parameterisations are used for hadrons with energies above and below 25 GeV. These parameterisations are slightly less detailed than the other physics lists, which has the advantage that the simulation is faster to run. However, it has been found that the agreement between data and simulation is not perfect for hadrons described by these parameterised models. Therefore, other models have been made which are slower to simulate, but contain a more detailed description of hadronic interactions and so produce a better agreement between MC and data. In addition, the description of the interaction of particles in the LHEP physics list changes very quickly across the defined energy threshold. In reality, there should be a gradual change in behaviour when going from high energy to low energy hadrons. The new physics lists provide some overlap between physics models in each energy range, to provide this smooth transition. Here, two new physics lists are compared with the LHEP list. Both of these lists use a string model to simulate the interactions of high energy

hadrons in materials, where a string is defined as an object which can carry energy and momentum, and has two fixed endpoints with well defined quark content.

In the QGSP_BERT list, the high energy interactions are modelled using the Quark Gluon String (QGS) model [89–92], which describes the interaction of high energy hadrons with nuclei in the detector material as an exchange of pomerons between the interacting partons. The pomeron is treated as a pair of strings, with their endpoints fixed at the interacting hadrons, and the strings become excited in this exchange process. This excitation causes the string to fragment, that is to decay into a hadron and a new string. This process is repeated until the remaining string does not have sufficient energy to hadronise. This model is valid for hadron energies above $\approx 25\,\text{GeV}$, and the low energy parameterisation is used for hadronic energies in the intervals 9–25 GeV. Around the 25 GeV boundary, one of the two models is chosen randomly, to produce a smooth transition from one model to the other.

The FTFP_BERT list uses a diffractive scattering model to describe the high energy interaction of hadrons and nucleons, based on the Fritiof model of diffraction [93]. In this model, a string is formed for each of the two scattered particles and excited by momentum exchange in the interactions. Fragmentation is then modelled in the same way as the QGS model. The advantage of this model is that it is valid for hadron energies down to $\approx 5\,\text{GeV}$, and so the low energy parameterisation is not used by this physics list. The expectation is that this model will be more accurate at lower energies.

Both FTFP_BERT and QGSP_BERT physics lists employ a cascade model to describe the low energy interactions, including inelastic scattering of hadrons with energy $< 10\,\text{GeV}$ and the deexcitation of excited nuclei, specifically the Bertini cascade model [94]. This models the classical processes involved in the scattering of low energy hadrons, creating secondary particles until the kinetic energy of all particles is reduced to zero.

The breakdown of the three physics lists and their regions of validity in energy space

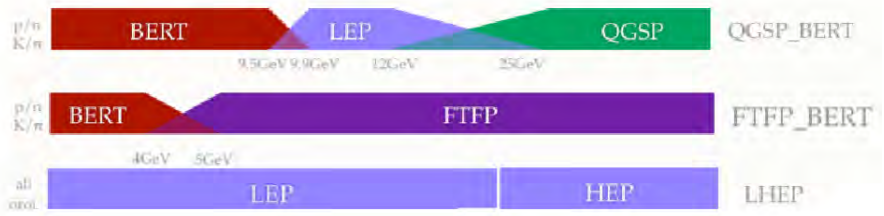

Figure 4.1: A diagrammatic representation of the composition of different physics lists and the energy range over which each part of the list is valid in.

are shown in Figure 4.1.

4.2.2 Comparison of Lists

In order to study these physics lists, and to determine their validity for use in LHCb, the interaction of particles with different target materials is investigated. GEANT4 is used to simulate a beam of particles travelling through a target material. For a range of incident particle energies, the total interaction cross section (elastic plus inelastic cross sections) is measured. This is repeated for different incident particles, comparing the cross sections for particles and anti-particles and for a range of target materials. The results are then compared for the three different physics lists, and compared with data from the COMPASS experiment [95].

In addition, it is important that these physics lists model well the differences in the interaction cross sections for particles and anti-particles. In LHCb many analyses are measuring the difference in the decay of particles and anti-particles, and to disentangle the effects of the *CP* violation from detection asymmetry, the interaction of particles and anti-particles in the LHCb detector must be determined.

Figure 4.2 shows the interaction cross section as a function of kinetic energy for protons and anti-protons interacting with aluminium, beryllium and silicon. These three materials are used extensively in the LHCb detector, particularly close to the interaction region, and so it is important to understand how GEANT4 will simulate

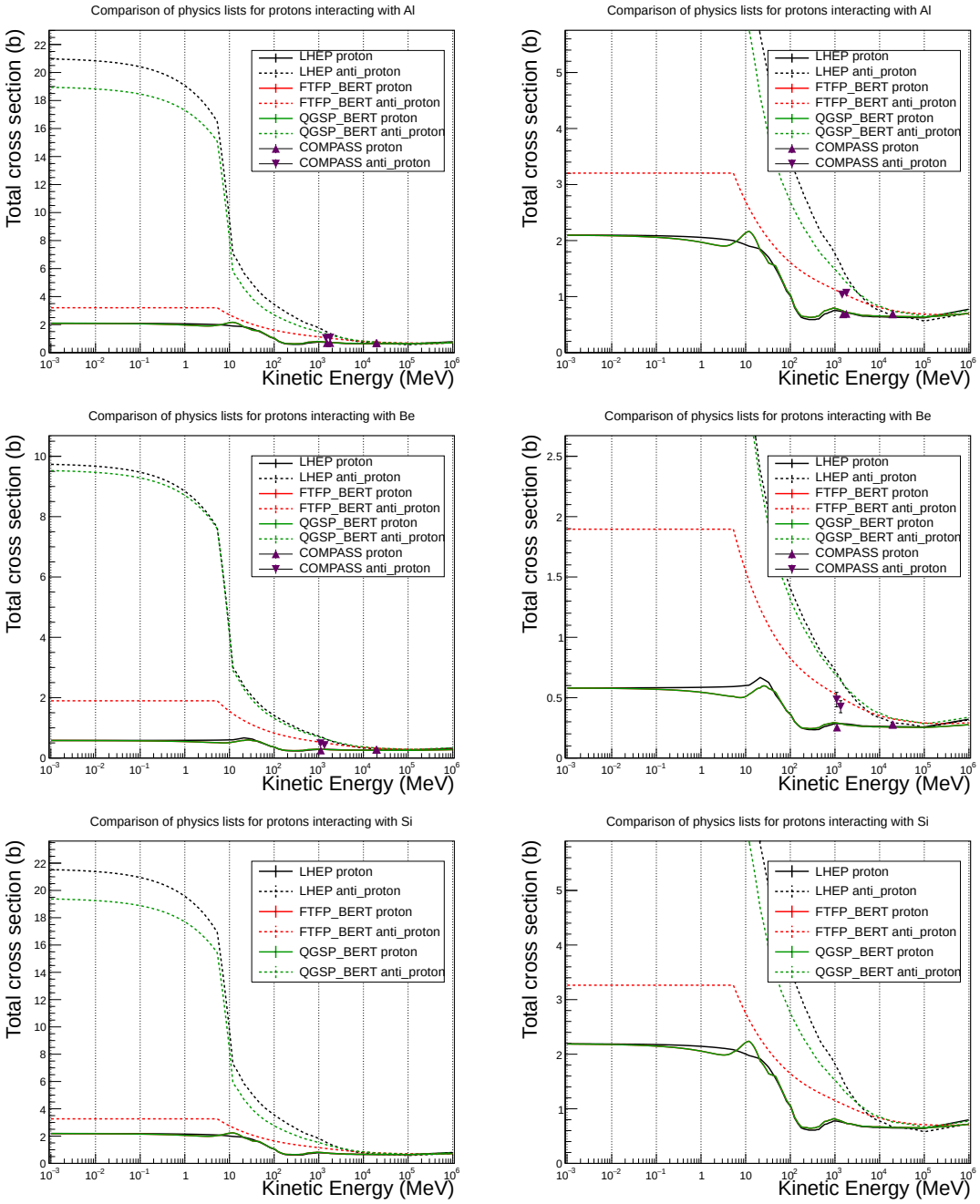

Figure 4.2: Comparison of the total interaction cross section for protons interacting in aluminium (top), beryllium (middle) and silicon (bottom). Right shows the same plot zoomed on the y-axis, to emphasise the regions of greater experimental importance.

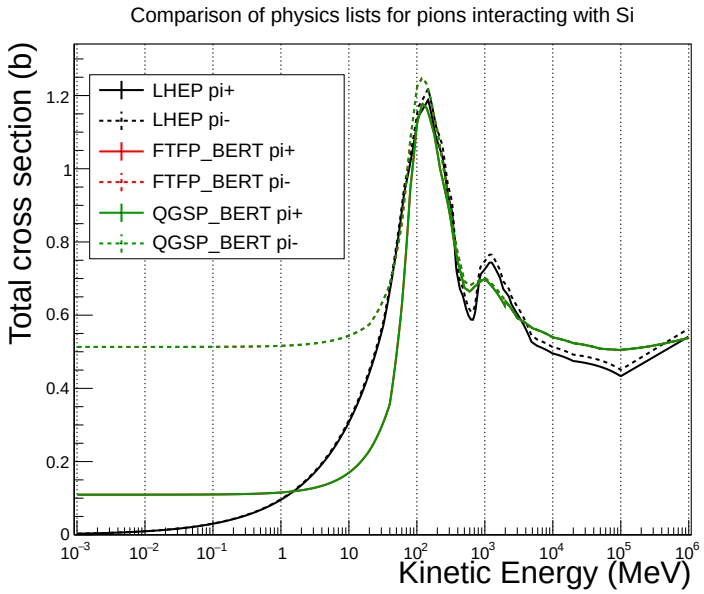

Figure 4.3: Comparison of the total interaction cross section for pions interacting in silicon. There is no difference between the FTFP_BERT and QGSP_BERT physics lists.

the interaction of particles with these materials. In addition to the three physics lists, the cross sections measured by the COMPASS collaboration are shown on the plots as triangular points, however, data is only available for aluminium and beryllium. In this plot, the uncertainties on these measurements are smaller than the point size. For protons, there is very little difference between the three physics lists, with no difference between the FTFP_BERT and QGSP_BERT lists, and all lists are consistent with the COMPASS data. For anti-protons, the cross section is much higher at low energies. The LHEP and QGSP_BERT physics lists have similar cross-sections, because they use the same parameterisation for low energies. It can be seen that cross sections predicted by the FTFP_BERT list are more consistent with the COMPASS data for anti-protons than are the other physics lists.

Figure 4.3 shows the cross section as a function of energy for pions interacting with a silicon target. In this case there is no difference between the cross section measured by the FTFP_BERT and QGSP_BERT physics lists. However there are some differences between these lists and the LHEP physics list, particularly when comparing the cross section for π^+ mesons with π^- mesons. The new lists introduce a much larger asymmetry in the cross sections at low energy ($< 10\,\text{MeV}$), but a smaller

asymmetry for hadrons with energy greater than 1 GeV, which is the typical energy of signal tracks in the LHCb detector. Further studies must now be performed by the LHCb collaboration to determine if this behaviour describes the detector asymmetries observed in data.

Due to the improved modelling of low energy interactions, and based on the recommendation of the GEANT4 team, it has been decided that LHCb will use the FTFP_BERT physics list as default in the simulation productions.

4.3 Production and Tracking Cuts in GEANT4

4.3.1 Applying cuts in GEANT4

As a particle interacts with a material in the LHCb detector, new particles can be produced. The nature of these particles could be hadrons produced in nuclear collisions, or electrons from pair production. These new particles are usually called secondary particles, and have a much lower energy than the original particle. In principle, each of these secondary particles must be simulated in the detector and tracked down to zero kinetic energy. This uses a large amount of CPU resources, and has very little impact on the overall behaviour observed in the LHCb detector. In order to simulate these processes in a more efficient way, requirements, or cuts, are applied to particles in GEANT4, so that only the high energy secondary particles are simulated. This section presents a study of the cuts applied in GEANT4, and the effect these cuts have on the distributions of hits in the individual subdetectors.

In the GEANT4 system set up in the LHCb software framework, there are two ways in which cuts can be applied to secondary particles. The first is by applying production cuts. Before a secondary particle is produced, the expected range of that particle is calculated, i.e. how far through the material the particle is expected to travel before it has zero kinetic energy. If this calculated range is below a given

threshold, then the secondary particle will not be simulated, and no tracking will take place. The calculated range depends on the material in which the particle is produced, and on the energy of the particle, so that particles are not produced if they are of too low an energy, or if they are produced in a material with a high cross section of interactions. Alternatively, one can apply tracking cuts, which place requirements on the energy of tracks in the detector, and will not track particles which are below an energy threshold.

By default, LHCb apply production and tracking cuts to photons, electrons and positrons, with a threshold range of 5mm and a threshold kinetic energy of 0.1 MeV. In addition, tracking cuts are applied to muons, pions, kaons and protons, with an energy threshold of 1 MeV. For convenience, the cuts will be referred to as Elecromagnetic (EM) and Hadronic (HAD) cuts for photons, electrons, positrons and muons, pions, kaons, protons respectively. Since these cuts affect the number of secondary particles produced in an event, this will have an impact on the number of hits observed in the detector. The number of hits is an important quantity used in many physics analyses, and it is important to ensure it is modelled correctly in the simulation. This section compares the number of hits observed in data and MC for different tracking and production cuts.

4.3.2 Discrepancies between data and Monte Carlo

In order to investigate the discrepancy between data and MC, comparable samples must be produced. During the data taking period, the LHCb experiment recorded a small amount of data while running a no-bias trigger. This is a special configuration of the LHCb trigger, in which every single pp bunch crossing is recorded. There is no requirement on the presence of tracks with a minimum energy, and so all measurements will be unbiased. However, the LHCb detector cannot run in this configuration for very long, and so only a small amount of data is recorded with these trigger conditions. MC samples are then produced with similar conditions, with the exception that a collision must take place in the simulation, and so this is

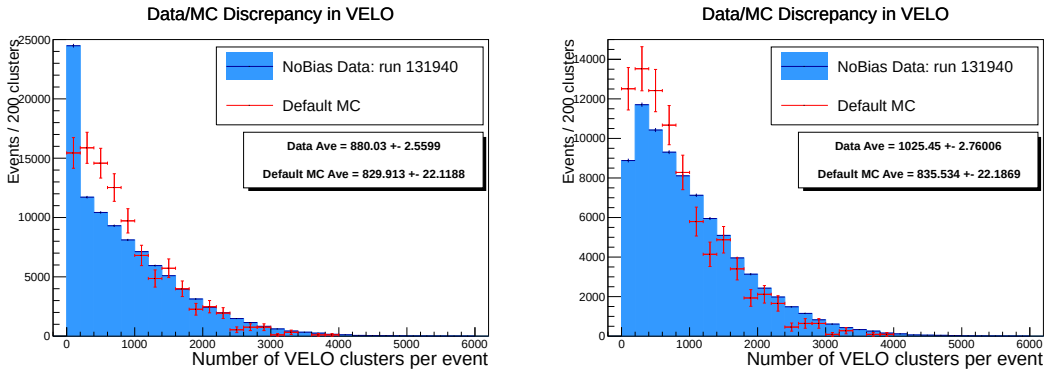

Figure 4.4: Comparison of the number of hits in the VELO for data and MC with no track requirement (left), and with at least one track in the event (right).

referred to as a minimum-bias sample. In the no-bias data, it is possible to record events which are completely empty, i.e. no inelastic collision has occurred, which would not be present in the minimum-bias MC samples. This effect is best shown in Figure 4.4, which shows the number of hits observed in the VELO, for both data and MC. The left plot shows the comparison between the no-bias data and minimum-bias MC with no requirements place on the number of tracks in the event. It can be seen that there is a large discrepancy between the two distributions, particularly in the lowest bin. This bin contains the empty events, which are not modelled by the MC. To overcome this problem, the plots are reproduced with a requirement that at least one reconstructed track is present in the event. With this requirement, the data has a small bias, but can be compared directly with the MC. This is shown in the right plot in Figure 4.4, and it can be seen that the discrepancy is now visibly reduced.

Another effect which needs to be taken into account is that of spillover, where particles from a previous bunch crossing are still in the detector when the next collision takes place. Since pp collisions occur every 50 ns in the LHCb detector, spillover is an effect which will certainly be present in the data, and so may need to be simulated in the MC as well. Spillover has a larger effect on the number of hits observed in the subdetectors which are further downstream of the interaction point; particles will travel very quickly through the first tracking stations, and spend longer in the calorimeters, where they interact to produce secondary particles. However,

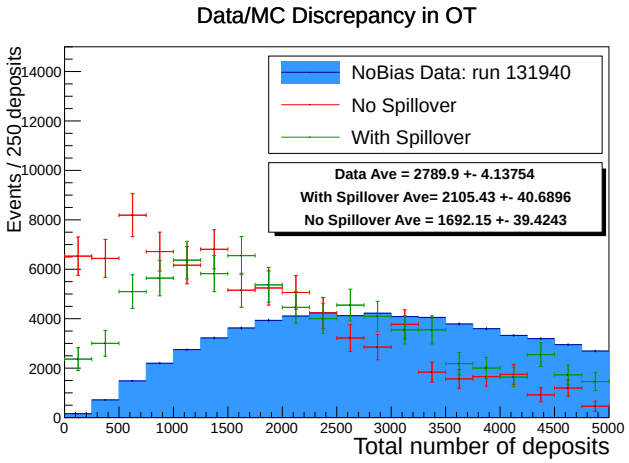

Figure 4.5: Comparison of the number of hits in the OT for data and MC simulated with and without spillover. Both samples have the requirement of at least one track to be present in the event.

since the granularity of the calorimeters is relatively low, a large number of primary and secondary particles will be classified as one hit in the calorimeter. Therefore, the only subdetector which we expect to be affected by spillover is the OT, where the hits are dominated by secondary particles produced by interactions with materials. The comparison of data with MC that has been simulated with and without spillover is shown in Figure 4.5. A clear difference between the two MC samples can be seen, with the average number of hits being higher, and closer to data, when spillover is simulated. In MC productions in 2011 and 2012, spillover was not simulated, which could explain some of the differences between data and MC which were observed. In 2015, the bunch spacing will be reduced to 25 ns, and so the effect of spillover will be enhanced. MC productions for these conditions will therefore have spillover simulated by default. The main problem with this is that the spillover requires a lot of CPU time to simulate, and so this is going to place an extra strain on the computing resources. Investigations into the effect of GEANT4 cuts on both the distribution of hits and the timing of simulations are therefore important. The next subsections look at the possibility of changing the cuts in GEANT4 to improve the agreement between data and MC, or to reduce the timing of simulations without significantly changing measurable distributions.

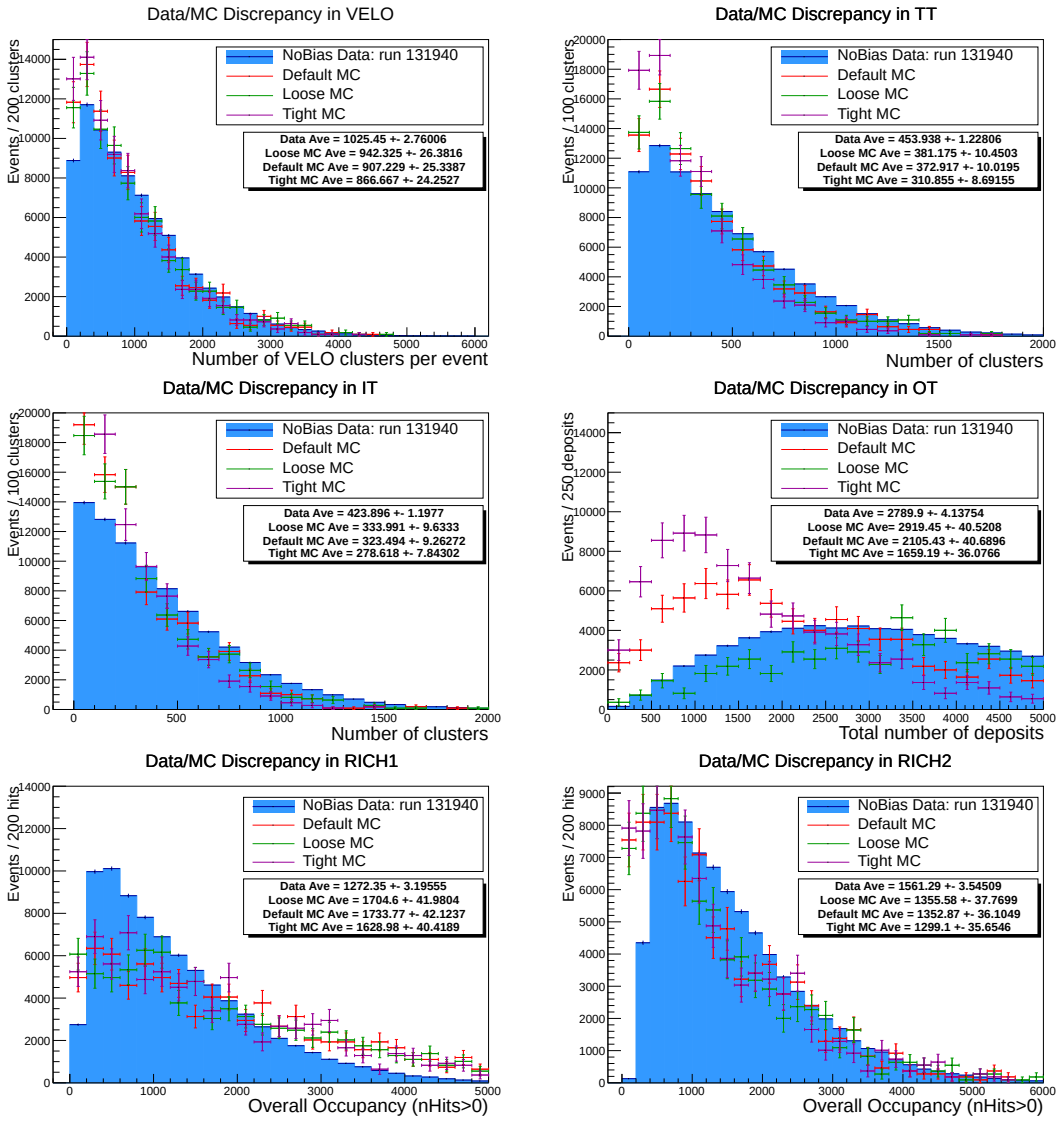

Figure 4.6: Comparison of the number of hits observed in the tracking and RICH subdetectors for data and MC produced with three levels of GEANT4 cuts.

4.3.3 Effect of GEANT4 cuts on hits in subdetectors

Figure 4.6 and Figure 4.7 shows the effect of different GEANT4 cuts applied on the predicted number of hits observed in each of the LHCb subdetectors. Each plot shows MC produced with three levels of GEANT4 cuts applied: the default cuts are 5mm production cuts as already mentioned, 1 MeV HAD tracking cuts and 0.1 MeV EM tracking cuts; the loose cuts are 0.01mm production cuts with no tracking cuts; and the tight cuts are 50mm production cuts, 100 MeV HAD tracking cuts and 10 MeV EM tracking cuts. The aim of these cuts is to look at the effect that

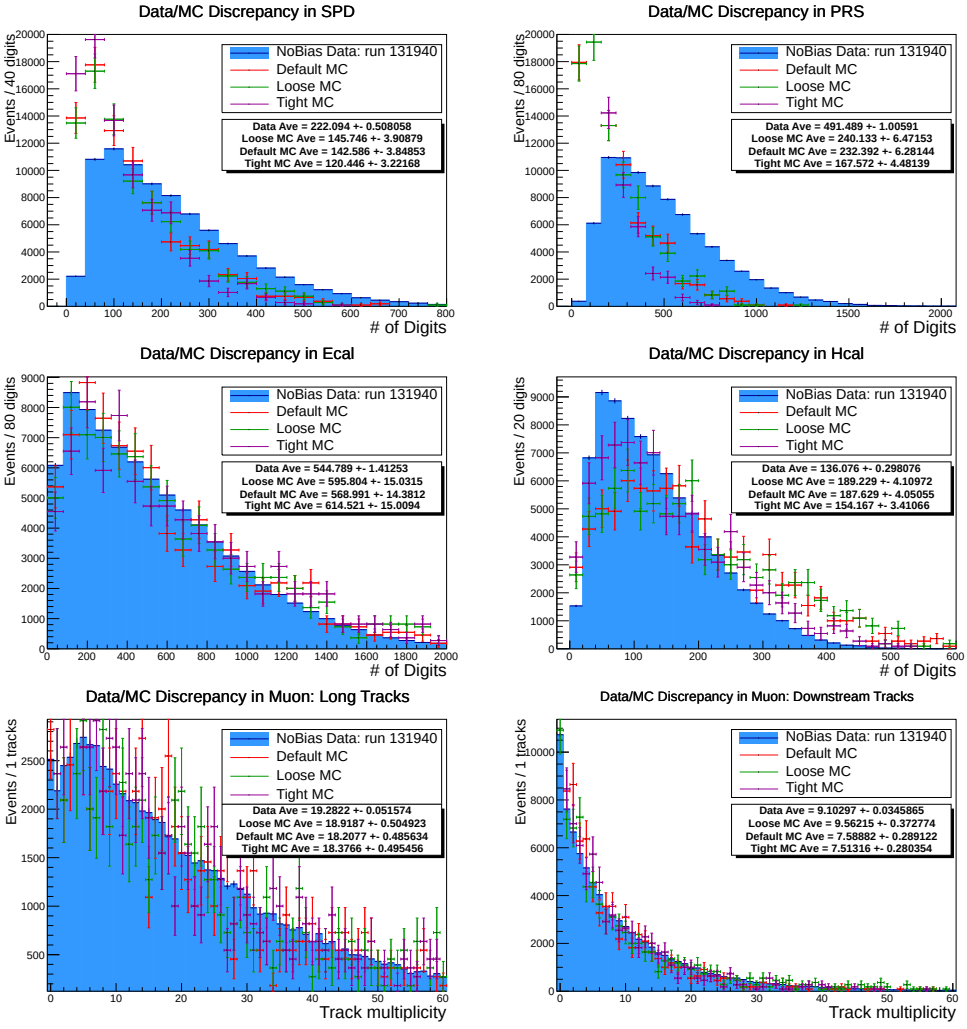

Figure 4.7: Comparison of the number of hits observed in the calorimeter and muon subdetectors for data and MC produced with three levels of GEANT4 cuts.

extreme cuts have on the distributions. The tight cuts are extremely tight, and would not be used in productions. The loose cuts are very loose, and the time taken to simulate events with these cuts increases dramatically. Using these extreme cuts highlights which subdetectors can be improved by changing the GEANT4 cuts, and which are not affected by the cuts.

From the plots in Figure 4.6 and Figure 4.7 we can see which subdetectors suffer from the largest discrepancy between data and MC, and of these, which can be improved by changing the GEANT4 cuts.

- In the tracking stations close to the beam pipe, i.e. the VELO, TT and IT, the discrepancies between data and MC are small, the number of hits is not affected by the GEANT4 cuts. This is to be expected since hits in these subdetectors are dominated by primary particles, and the GEANT4 cuts only affect secondary particles.

- This is not the case in the OT, since the hits in the OT are predominantly secondary hits. The number of hits observed in this subdetector is therefore strongly dependent on the GEANT4 cuts, and by loosening these cuts, a better agreement between data and MC can be achieved.

- There are large discrepancies in the number of hits observed in the SPD and PRS, but the number of hits observed in the MC is not affected by the GEANT4 cuts, which indicates the problem is with the number of tracks produced by PYTHIA, and not the number of secondary hits produced.

- In the ECAL, there is no discrepancy between data and MC, and in the HCAL there is a small discrepancy. The number of hits observed in the MC is higher than in data, but this is improved with tighter GEANT4 cuts. This implies that extra secondary hits are being simulated, which are not present in the data due to an energy threshold not being present in the simulation. Therefore with tighter GEANT4 cuts, there are fewer hits per event produced, which is consistent with observations in data.

- In the RICH detectors, there is again some discrepancy, but the number of hits is not affected by the GEANT4 cuts, and the number of hits in the muon chambers matches between data and MC.

In order to find the optimum cuts to apply to tracks in the OT and HCAL, dedicated MC samples were produced with a range of production and tracking cuts. Figures 4.8 and 4.9 show the average number of hits observed in each subdetector as a function of different GEANT4 cuts. Figure 4.8 shows the average number of MC hits observed, that is the number of particles simulated by GEANT4. Figure 4.9 shows the number of reconstructed hits observed after the full reconstruction chain is applied, as in Figure 4.6. There is clearly a big difference between the number of MC hits and the number of reconstructed hits. This is because the simulation software registers a new MC hit for every interaction simulated by GEANT4. However, these will not all be registered as separate hits by the reconstruction code, which will group all secondary particles into one hit. This is particularly emphasised in the calorimeters, which generally have a low granularity, and so many secondary particles will contribute to one energy deposit. This is interesting to note, and it is important to see that in many subdetectors one can reduce the number of secondary particles produced in the simulation without affecting the number of reconstructed hits in the detector simply by tightening the cuts in GEANT4.

As expected, the only distribution which is affected by the GEANT4 cuts is the number of reconstructed hits in the OT, which is mostly sensitive to the electron production cuts, with a slight dependence on the electron tracking cuts. The dashed line on Figure 4.9 shows the average number of hits observed in the no-bias data sample. It can be seen that agreement between the MC and the data can be achieved with a cut of 0.03 mm on the electron production cut in the OT. For the cuts applied in the other subdetectors, the situation is less clear. For example, the average number of hits in the HCAL decreases as the tracking cut on photons is tightened, but there is still no agreement between the MC and data, in which we find an average number of hits of 136. It is therefore not clear if tightening the GEANT4 cuts in the

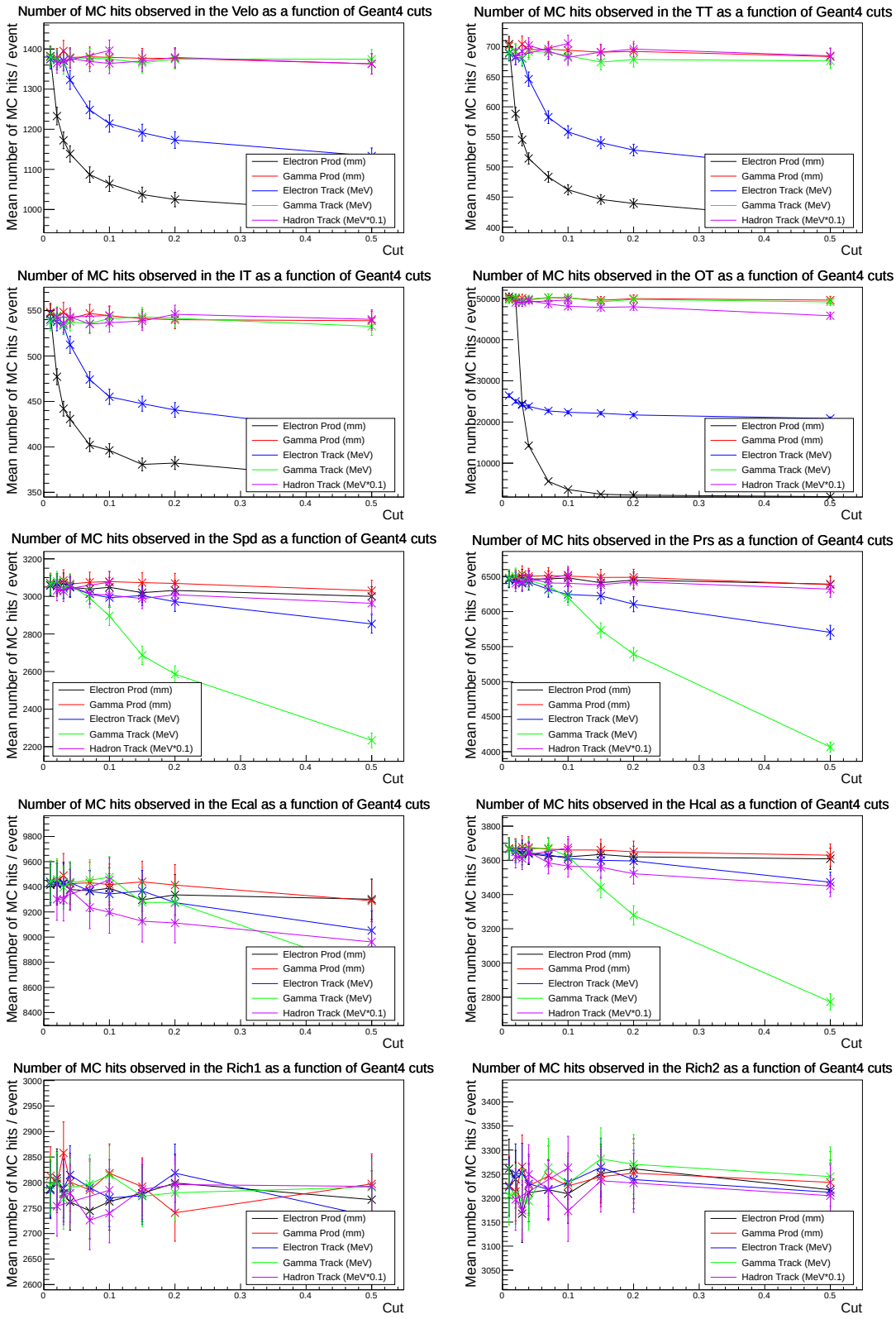

Figure 4.8: The number of MC hits observed in each subdetector as a function of GEANT4 cuts. Note that the hadronic tracking cuts have been scaled by 0.1 to fit on the same axis.

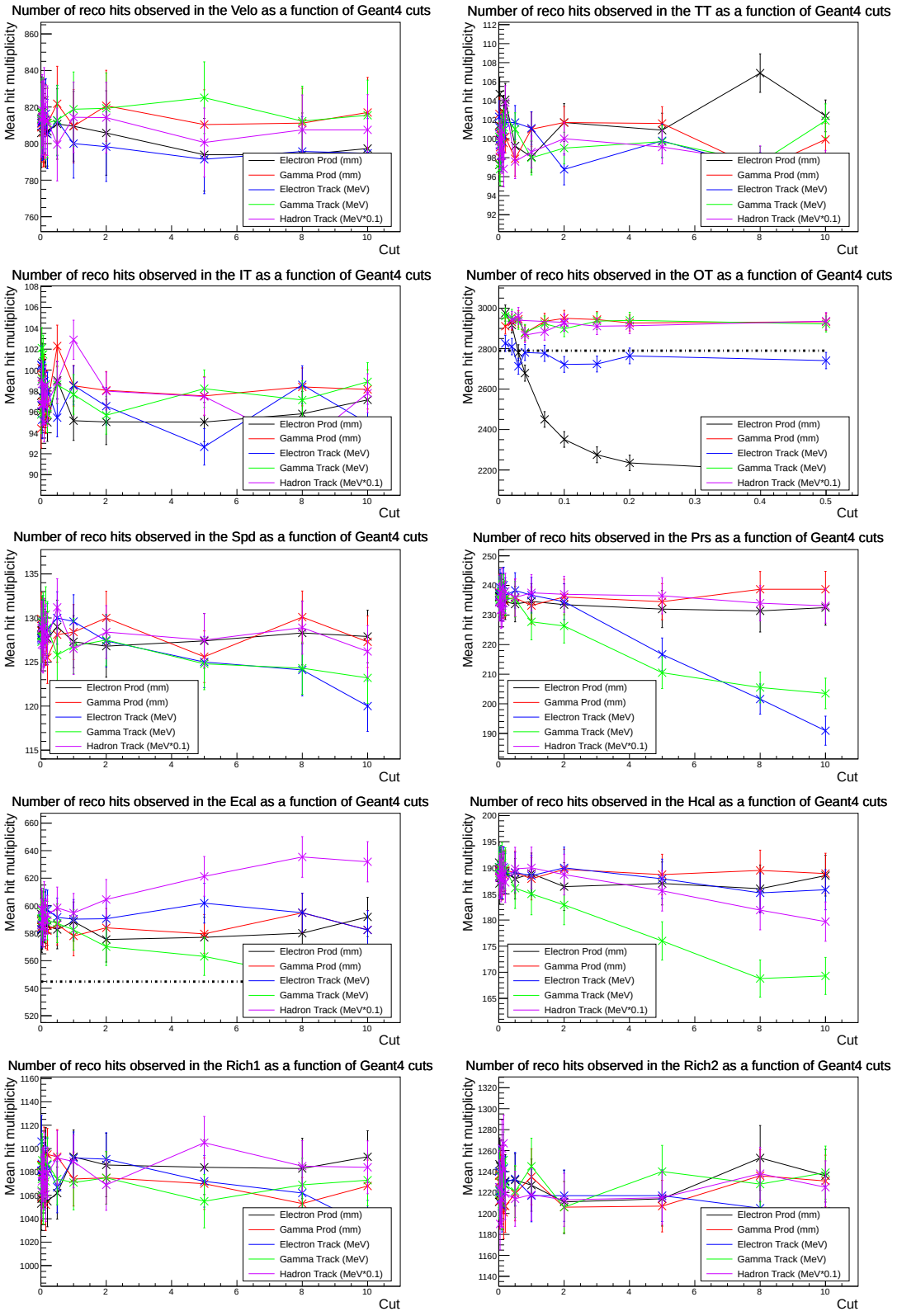

Figure 4.9: The effect of tightening cuts in GEANT4 on the average number of reconstructed hits detected in subdetectors. Note that the hadronic tracking cuts have been scaled by 0.1 to fit on the same axis. Where there is a dashed line, this indicates the number of hits observed in data.

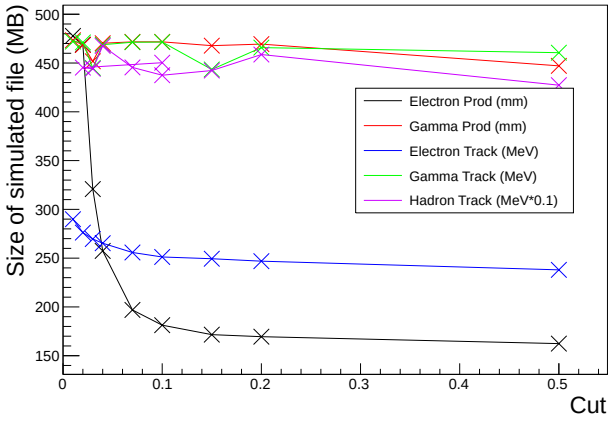

Figure 4.10: Size of simulated files produced as a function of GEANT4 cuts.

HCAL will improve the agreement between data and MC, and more investigation is required.

4.3.4 Effect of GEANT4 cuts on computing resources

In addition to achieving agreement between MC and data, it is important to consider the resources used to simulate MC samples. Simulating MC samples requires two major resources: disk space and computing time. These are both limited for the LHCb experiment, and so the simulations need to be optimised to reduce these resources while maintaining sufficient detail to reliably model data. This section describes how the GEANT4 cuts can affect the computing resources used by the simulation.

The largest effect on the disk space used by simulated samples is the information about the simulated events which is stored in the MC file, but looser GEANT4 cuts will produce more secondary particles, and this will increase the size of the files produced. This is evident in Figure 4.10, which shows the size of the file produced when 100 events are simulated. The size rapidly increases as the electron production cuts are loosened, and more secondary particles are being created in the OT and stored.

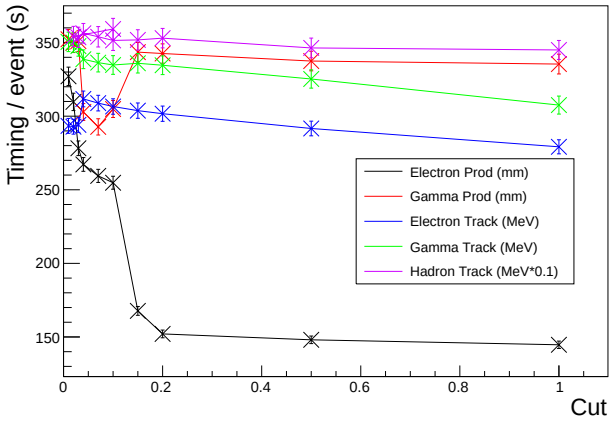

Figure 4.11: Time to simulate an event as a function of GEANT4 cuts.

In addition to the size of the simulation files, the GEANT4 cuts also have a large impact on the time taken to simulate each event. As previously explained, the timing of the simulation can be split into two distinct stages: the time to simulate the pp collision, and the time to simulate the response of the particles in the LHCb detector. The latter stage is by the far the most time consuming, and so optimising the simulation of the detector can have a big impact on the simulation time. The effect of the GEANT4 cuts on the simulation can be seen in Figure 4.11. The time is measured using an internal timing monitor available in the simulation software and varies with the CPU used to run the simulation. All the jobs are run on a local batch system, where the speed of each CPU is known, and all batch nodes have the same processing power. In addition, the simulation job is split into 10 subjobs, with each subjob generating 100 events, and the results of these jobs are merged after the processing. Each subjob runs on a different batch node, and so by taking an average of the timing of each subjob will mitigate any differences which might exist in the nodes which may be processing more than one job at a time. The plots show the time to simulate one event with a given set of GEANT4 cuts, averaged over all simulations with those cuts. It should be noted that here we are measuring only the time taken to simulate the event. In principle the GEANT4 cuts will have an effect on the time taken for the digitisation and reconstruction stages, but these are much faster than the simulation and so are not as important for computing resources.

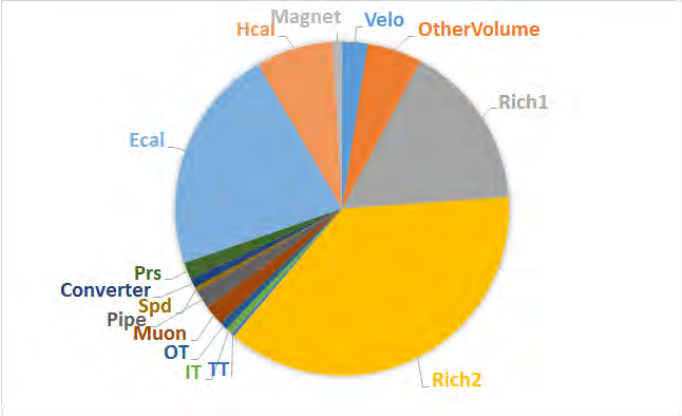

Figure 4.12: Distribution of time spent to simulate an event in each volume of the LHCb detector using the default GEANT4 cuts.

The effect of the GEANT4 cuts on the timing is clear: with looser GEANT4 cuts, more secondary particles are produced in the detector which take time to simulate through the subdetectors. In order to improve the timing of the simulations, it is interesting to see in which subdetectors the processes take longest to simulate. In order to do this, a tool was written to use the SteppingAction method in GEANT4 to find out for each step of the simulation, which subdetector the step is simulated in, and the total time taken to simulate the step. For each event it is then possible to keep track of the total time spent simulating processes in each subdetector. The results of these studies are shown in Figure 4.12, which shows the distribution of time spent in each subdetector when simulating 1000 events using the default GEANT4 cuts used in the LHCb collaboration. It can be seen that the time taken to simulate an event is dominated by the time spent in the RICH, ECAL and HCAL detectors.

4.3.5 Results with optimum cuts

The results of the optimisation indicate that production cuts of 0.03 mm and tracking cuts of 0.01 MeV on electrons and positrons give a good agreement between data and MC in the OT, but also increases the computing resources used by GEANT4. It was therefore decided to apply these loose cuts only in the OT, and keep the default cuts in the other subdetectors. The number of hits observed in the OT using these

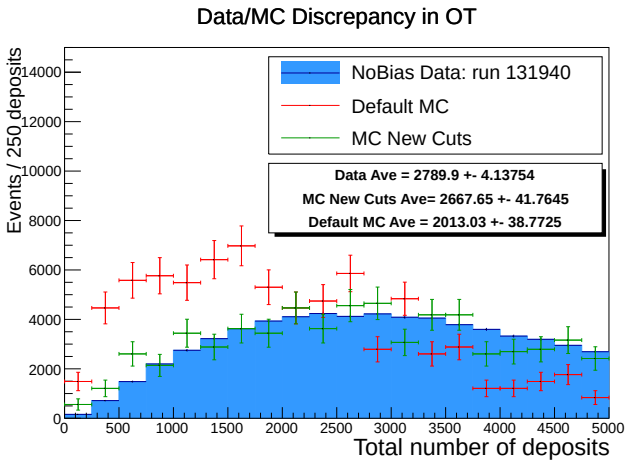

Figure 4.13: Distribution of hits observed in the OT using the optimal cuts found in these investigations, with spillover simulated. For comparison, the distribution of the default cuts and the distribution observed in data is included from Figure 4.5.

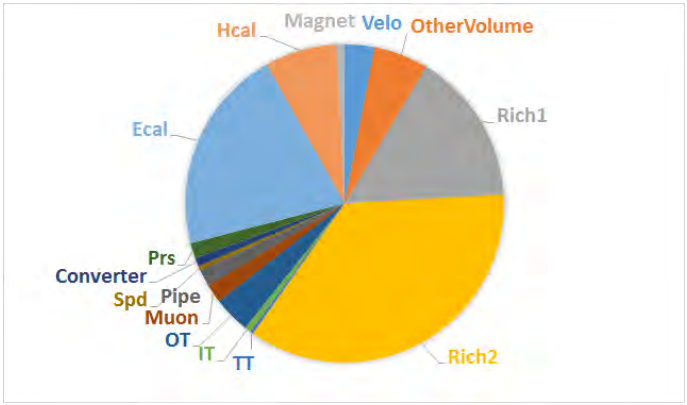

Figure 4.14: Distribution of time spent to simulate an event in each volume of the LHCb detector using the optimal cuts found in these investigations.

new optimised cuts is shown in Figure 4.13. It can be seen that the agreement with the optimised cuts is much better than the default cuts which were previously used. In order to assess the impact of these new cuts on the computing resources, the timing studies described in the previous section were repeated. The time spent in each subdetector is shown in Figure 4.14, and the relative amount of time spent in the OT has increased compared with the other detectors where the cuts applied are tighter, now taking 3.7% of the overall time, compared with 0.7% of the overall time when the default cuts are applied. The average time spent to simulate an event with these new cuts in (267 ± 8) s, which is a small decrease in the time per event with the default cuts applied everywhere, which was (293 ± 9) s. The statistical significance of the difference in the timing per event is only 2.3σ, which indicates that loosening the cuts in the OT only does not have a significant impact on the timing of the simulations.

4.4 Conclusions and Outlook

The effect of the use of GEANT4 in the simulation of the LHCb detector in the production of MC samples has been investigated.

The different physics lists available to describe hadronic processes have been studied and it was found that the FTFP_BERT physics list was most appropriate to describe the hadronic physics involved in interactions in the LHCb detector, and this is now the list currently used by the LHCb collaboration. The process of plotting the cross sections has been automated so that these checks can be carried out to highlight rapidly any changes to the physics lists between different versions of GEANT4, and to compare any new physics lists that might become available in the future.

The effect of the cuts applied to secondary particles by GEANT4 has also been investigated. It is found that applying looser cuts to the particles produced in the OT subdetector gives a better agreement between the MC simulated samples and the data samples for the number of hits observed in the OT. It is found that the

cuts applied in GEANT4 do not affect the number of hits in any other subdetector. These cuts do, however, have a large impact on the computing resources used by the simulation, and so to minimise this effect the looser cuts are only applied in the OT, where the improvement is seen. There are some limitations to the investigation as described, and some further investigation is required before these results can be implemented in the LHCb productions. The studies are carried out using minimum bias simulations and no-bias data. This is the simplest way of investigating the differences between data and MC, but it may not be a full representation of the differences observed in real physics analyses. In particular, after the trigger is applied to the data, this will change the distribution of the number of hits observed in the subdetectors, and may have an impact on the agreement between data and MC. In addition, while the number of hits in a subdetector is interesting, it would also be useful to investigate the impact of this on physics quantities such as the tracking efficiency or impact parameter resolution. This would be a useful study to determine if there is a real advantage in loosening the cuts to improve the agreement between data and MC. These studies are more difficult since they involve the simulation of the LHCb hardware trigger, which is difficult to model. It would therefore be difficult to separate any differences observed into those due to the detector simulation and to the modelling of the trigger. Since the studies presented in this chapter show the effect of the detector simulation on these discrepancies, they can be a useful starting point for further studies of interesting physics channels. Possible decays which would be useful to investigate are the $D^{*\pm} \to D^0\pi^\pm$, $B^\pm \to J/\psi\, K^\pm$ and $B_s^0 \to D_s^\pm\pi^\mp$ decays, which are expected to be clean, high statistics signals in LHCb, and are representative of many similar analyses which use the simulation.

CHAPTER

5

ANALYSIS STRATEGY IN SEARCH FOR $\Lambda_B^0 \to \Lambda^0 \eta^{(\prime)}$

This chapter describes the analysis strategy used in the search for the $\Lambda_b^0 \to \Lambda^0 \eta^{(\prime)}$ decays. The physics interest in searching for this decay has been presented in Chapter 2, and the main aim of this analysis is to observe these decays and measure their branching fractions. In this chapter, the general strategy for reconstructing the decays and measuring the branching fractions is described, before detailing the selection applied to the data to separate the signal and background. The results of the search are presented in Chapter 6.

5.1 Introduction

The general analysis strategy is to use the $B^0 \to K_S^0\eta'$ decay as a control channel in the search for the $\Lambda_b^0 \to \Lambda^0\eta'$ and $\Lambda_b^0 \to \Lambda^0\eta$ decays, as well as the equivalent charge conjugated decays of $\overline{\Lambda}_b^0$ baryons. This channel has a similar decay topology to the signal channels, but a higher branching fraction, which has been measured to be $(6.85 \pm 3.8) \times 10^{-5}$ [7]. Since this channel has been studied by the BaBar and Belle collaborations, the outcome of the selection can be predicted and used to check the analysis being applied. In addition, the ratio of branching fractions for the signal and control channels is measured. By measuring the ratio instead of the absolute branching fractions of the signal modes, many of systematic uncertainties, for example due to the luminosity or the production cross sections, are expected to cancel out and a more precise measurement can be made.

The search is performed using the full dataset recorded by the LHCb experiment, which corresponds to $1.0\,\text{fb}^{-1}$ of pp collision data collected in 2011 at a centre of mass energy $\sqrt{s}$ =7 TeV and $2.0\,\text{fb}^{-1}$ of data recorded in 2012 at $\sqrt{s}$ =8 TeV. The data will contain candidate signal events, and background events which are reconstructed in a similar way to signal events but do not originate from the decays of interest. The main aim of the analysis is to design a selection which will reject as much background as possible whilst keeping a significant number of signal events. The signal can be observed by plotting the invariant mass distribution of the reconstructed B^0 and Λ_b^0 candidates. For signal events this distribution will peak around the B^0 or Λ_b^0 invariant mass, which will not be the case for background events.

The analysis is performed blind, such that events with an invariant mass in the range $5519\,\text{GeV}/c^2 < m(\Lambda^0\eta^{(\prime)}) < 5719\,\text{GeV}/c^2$ are obscured until the selection and fit models have been finalised. If the analysis is not blind, then the selection may be optimised to what is observed in the data, which will bias the measurements made. By not looking at the signal until the selection is fixed, we can be confident that the results are not biased. The selection is optimised using the control channels, MC

signal events and the data sidebands of the Λ_b^0 decays. After unblinding, the yield the Λ_b^0 decays is found from fits to the data. If no significant signal is observed then the yield is used to place a limit on the branching fraction.

5.2 Reconstructing the decays

The first step of the analysis is to use the tracks and energy deposits in the LHCb detector to reconstruct the B^0 and Λ_b^0 candidates, and the intermediate resonances.

The K^0 is reconstructed through the K_S^0 decay $K_S^0 \to \pi^+\pi^-$, which has a branching fraction of 69.2% [7]. The K_S^0 can also decay through $K_S^0 \to \pi^0\pi^0$ with a branching fraction of 30.69% [7]. However, neutral particles such as π^0 mesons are difficult to reconstruct as they leave no tracks in the detector. The K_S^0 will then not be reconstructed accurately, and so decays to neutral particles are not considered in this analysis. K_L^0 mesons have not yet been successfully reconstructed in LHCb as this would require reconstructing a neutral energy deposit in the hadronic calorimeter. Therefore, only K_S^0 mesons are reconstructed in this analysis.

The K_S^0 is classified according to where it decays in the LHCb detector. If a K_S^0 decays within the VELO then both of its tracks will be Long tracks, as described in Chapter 3, so the K_S^0 will be reconstructed as Long-Long (LL). On the other hand, if the K_S^0 decays after the VELO, and the daughter tracks produce hits in just the TT and the IT/OT, then the tracks will be reconstructed as downstream tracks and the K_S^0 is classified as Downstream-Downstream (DD). The resolution of the measured variables will be poorer for the DD category compared with the LL category, and therefore the selection optimisation is split into a LL and DD selection to make the best use of the data.

The Λ^0 baryon is reconstructed through the $\Lambda^0 \to p\pi^-$ decay (plus the charge conjugated decay), which happens with a branching fraction of 63.9% [7]. The Λ^0 also decays through $\Lambda^0 \to n\pi^0$, with a branching fraction of 35.8% [7], however, only

Table 5.1: Decays of η' and η mesons [7].

Decay	Branching Fraction (%)
$\eta' \to \pi^+\pi^-\gamma$	29.1 ± 0.5
$\eta' \to \pi^+\pi^-\eta$	42.9 ± 0.7
$\eta' \to \pi^0\pi^0\eta$	22.2 ± 0.8
$\eta \to \gamma\gamma$	39.41 ± 0.20
$\eta \to \pi^+\pi^-\pi^0$	22.92 ± 0.28
$\eta \to \pi^0\pi^0\pi^0$	32.68 ± 0.23
$\eta \to \pi^+\pi^-\gamma$	4.22 ± 0.08

the decay to charged particles is used. As with the K_S^0 meson, the Λ^0 baryon is a long lived particle, and so can be reconstructed with both long and downstream tracks. The selection is therefore split into LL and DD selections.

There are three main decays which can be used to reconstruct the η', which are summarised in Table 5.1. Since the decay to $\pi^0\pi^0\eta$ contains only neutral particles, the reconstruction of this decay very difficult. It will be difficult to reconstruct a secondary vertex for this decay, making it difficult to trigger on and introducing a large amount of background. It will therefore not be used in this analysis. Instead, the η' meson is reconstructed using the two final states, $\pi^+\pi^-\gamma$, which is dominated by the $\eta' \to \rho^0\gamma$ decay, and $\pi^+\pi^-\eta$, where the η subsequently decays to two photons.

Of the main decays of the η meson, two of the final states, $\gamma\gamma$ and $\pi^0\pi^0\pi^0$ contain only neutral particles, and hence are not useful for this analysis. One decay channel with charged particles is the $\eta \to \pi^+\pi^-\gamma$ decay, which has a relatively low branching fraction, and so is not expected to add much to the analysis. Therefore, only the $\eta \to \pi^+\pi^-\pi^0$ decays are used in the search.

Finally, the intermediate resonances are combined to reconstruct the B^0 or Λ_b^0 candidate. In order to improve the resolution of the reconstructed invariant mass, the decay is refitted using the DecayTreeFitter tool [96]. A least squares fit is performed to refit the momentum vectors of the tracks, with the mass of the intermediate resonances fixed to their known masses and the position of the PV fixed to that found from tracks in the underlying event which are not signal tracks.

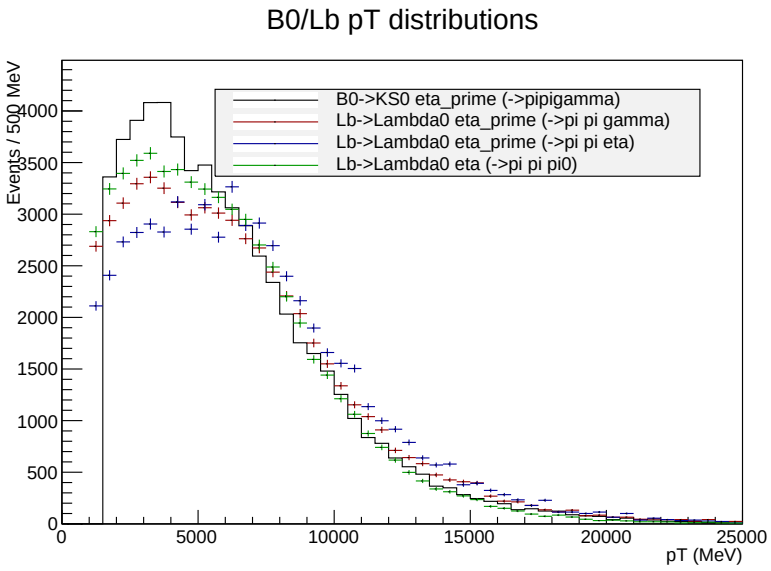

Figure 5.1: Comparison of B^0 p_T distribution with the Λ_b^0 p_T distribution in each of the signal channels.

Since both the branching fraction of the $B^0 \to K_S^0\eta$ decay, and the trigger efficiency for $B^0 \to K_S^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) are too low, the $B^0 \to K_S^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay will be used as a control channel in the search for all three signal channels, $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$), $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) with $\eta \to \gamma\gamma$, and $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$).

5.3 Comparison of control and signal channels

The control channel is compared in detail with each of the signal channels in order to check that the decay kinematics are similar. The reconstruction and selection efficiencies in general depend on the kinematics of the decay particles. It follows that any correction for, or uncertainty on, this efficiency is also a function of the kinematics. However, if the kinematics of the signal and control channels are the same then the uncertainties will cancel out when the ratio of branching fractions is measured.

Figure 5.1 shows the comparison of the p_T distribution of the B^0 from the control

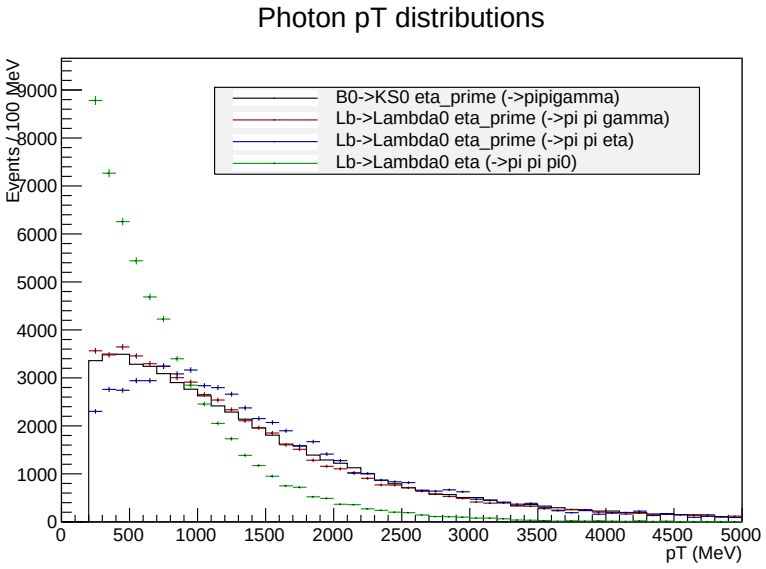

Figure 5.2: Comparison of photon p_{T} distributions for each decay channel.

channel, with the corresponding p_{T} distribution of the Λ_{b}^0 baryons from the signal channels. Since the distributions are very similar, we can assume that any systematic uncertainty on the measurement of the efficiencies using MC samples will cancel when the ratio of efficiencies is calculated.

The efficiency of reconstructing photons is a quantity which is not modelled well in the MC simulations. There will be an uncertainty on the measured efficiency due to this difference. However, if the photon p_{T} distribution is the same for both signal and control channels, then this uncertainty will largely cancel out in the measurement of the ratio of the branching fractions. The p_{T} distribution of photons in each of each signal channels is shown in Figure 5.2. The p_{T} distribution of photons from $\eta' \to \pi^+\pi^-\gamma$ decays is the same for signal and control channels, and so no correction needs to be applied for the $\Lambda_{\mathrm{b}}^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel. For the $\Lambda_{\mathrm{b}}^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) and $\Lambda_{\mathrm{b}}^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) decays there are more photons in the signal channels than the control channel, and the p_{T} distribution is slightly different. This means the uncertainty will not completely cancel, and to account for this a data driven correction to the photon reconstruction efficiency is applied. Details of the calculation of this correction for each signal channel is given in Section 5.4.

5.4 Branching fraction measurements

The branching fractions will be measured relative to the control channel. By measuring the ratio:

$$R \equiv \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta^{(\prime)})}{\mathcal{B}(B^0 \to K^0 \eta')}, \tag{5.1}$$

many of the systematic uncertainties and acceptance factors will cancel. The branching fractions are calculated using the following formula:

$$\mathcal{B} = \frac{N_S}{2\mathcal{L}_{int}\sigma_{b\bar{b}}\epsilon_{acc} f_x \epsilon_{tot} \times C_\gamma \times \mathcal{B}(\eta^{(\prime)}) \times g_l \times \mathcal{B}(K_S^0 \to \pi^+\pi^-)\{\mathcal{B}(\Lambda^0 \to p\pi^-)\}}, \tag{5.2}$$

where N_S is the number of signal decays which have been observed in the data, and $\mathcal{L}_{int}$ is the total integrated luminosity, which is $(3.0 \pm 0.1)\,fb^{-1}$ for this data sample. $\sigma_{b\bar{b}}$ is the total cross section for producing $b\bar{b}$ quark pairs. This has been measured within the LHCb acceptance, and then extrapolated to the full 4π acceptance using the extrapolation factor given in PYTHIA 6.4. The full cross-section at $\sqrt{s}$ =7 TeV is then $\sigma_{b\bar{b}} = (284 \pm 53)\,\mu b$ [97]. At $\sqrt{s}$ =8 TeV, the cross section is found to be $(298 \pm 36)\,\mu b$ [98]. This cross-section must then be multiplied by the acceptance factor, ϵ_{acc}, i.e. the fraction of signal decays which are in the LHCb acceptance. This has been estimated with MC simulations, and the acceptance factor for each decay mode is presented in Table 5.2. C_γ is the correction factor to apply due to the photon efficiency, described in Section 5.3. $\mathcal{B}(\eta^{(\prime)})$ is the branching fraction for the η' or η decays, shown in Table 5.1. $\mathcal{B}(K_S^0 \to \pi^+\pi^-) = (69.2 \pm 0.05)\%$ is the branching fraction for the K_S^0 decay, and $\mathcal{B}(\Lambda^0 \to p\pi^-) = (63.9 \pm 0.5)\%$ is the branching fraction of the Λ^0 decay [7]. The factor g_l, where $g_l = 0.5$ for K_S^0 decays and $g_l = 1$ for Λ^0 decays, accounts for the fact that only half of the K^0 mesons decay as K_S^0. ϵ_{tot} is the total efficiency for selecting signal events, which is determined by applying the selection to a sample of MC simulated signal events. f_x are the fragmentation fractions, the fraction of time the b– or $\bar{b}$– quark hadronises with quark of time x, f_{B^0} for the B^0 decays and $f_{\Lambda_b^0}$ for the Λ_b^0 decays. The ratio $f_{\Lambda_b^0}/f_{B^0}$ has been calculated by LHCb, and has been found to have a dependence on the

Table 5.2: Acceptance efficiencies of each decay.

Decay	Year	ϵ_{acc} %
$B^0 \to K_S^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	2011	0.237 ± 0.004
	2012	0.238 ± 0.004
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	2011	0.236 ± 0.002
	2012	0.237 ± 0.001
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\eta)$	2011	0.216 ± 0.001
	2012	0.218 ± 0.001
$\Lambda_b^0 \to \Lambda^0 \eta$ $(\eta \to \pi^+\pi^-\pi^0)$	2011	0.231 ± 0.001
	2012	0.232 ± 0.001

pseudorapidity, η of the Λ_b^0 baryon [99], with the dependence having the functional form

$$\frac{f_{\Lambda_b^0}}{f_{B^0}} = (0.387 \pm 0.033) + (0.067 \pm 0.013)(\eta - \bar{\eta}), \tag{5.3}$$

where $\bar{\eta}$ is the average pseudorapidity, and is fixed at $\bar{\eta} = 3.198$.

The weighted average of the η distribution for the Λ_b^0 baryon is calculated, and averaging over all Λ_b^0 channels gives $\langle\eta\rangle = 3.404 \pm 0.001$. Using the value in Equation 5.3 gives:

$$\frac{f_{\Lambda_b^0}}{f_{B^0}} = 0.40 \pm 0.03. \tag{5.4}$$

Under the assumption that $f_{\text{baryon}} = f_{\Lambda_b^0}$ and $f_{B^\pm} = f_{B^0}$, one can use the unitarity relation below to calculate $f_{\Lambda_b^0}$ and f_{B^0}.

$$f_{B^0} + f_{B^\pm} + f_{B_s^0} + f_{\Lambda_b^0} = 1. \tag{5.5}$$

Substituting the recent LHCb measurement for $f_{B_s^0}/f_{B^0}$ [100] gives $f_{B^0} = 0.377 \pm 0.005$ and $f_{\Lambda_b^0} = 0.151 \pm 0.011$.

The ratio of signal channel to control channel branching fractions can be calculated using:

$$R = \frac{N_S(\Lambda_b^0)}{N_S(B^0)} \times \frac{\epsilon_{acc}(B^0)}{\epsilon_{acc}(\Lambda_b^0)} \times \frac{\epsilon_{tot}(B^0)}{\epsilon_{tot}(\Lambda_b^0)} \times \frac{f_{B^0}}{f_{\Lambda_b^0}} \times \frac{(C_\gamma)^{N_\gamma(B^0)}}{(C_\gamma)^{N_\gamma(\Lambda_b^0)}} \times \frac{0.5 \times \mathcal{B}(K_S^0)}{\mathcal{B}(\Lambda^0)} \times \frac{\mathcal{B}(\eta')}{\mathcal{B}(\eta^{(\prime)})}. \tag{5.6}$$

The acceptance factors have been calculated using MC productions, and shown in Table 5.2.

The ratio of branching fractions can be calculated from the PDG values (and their uncertainties), giving:

$$\frac{\mathcal{B}(\mathrm{K}^0_\mathrm{S} \to \pi^+\pi^-)}{\mathcal{B}(\Lambda^0 \to \mathrm{p}\pi^-)} = 1.083 \pm 0.009 \tag{5.7}$$

for the K^0_S and Λ^0 decays, and

$$\frac{\mathcal{B}(\eta' \to \pi^+\pi^-\gamma)}{\mathcal{B}(\eta' \to \pi^+\pi^-\eta)\ [\mathcal{B}(\eta \to \pi^+\pi^-\pi^0)]} = 1.71 \pm 0.05\ [1.31 \pm 0.03] \tag{5.8}$$

for the $\eta^{(\prime)}$decays. The ratio of production fractions can be calculated from the LHCb measurement, assuming $f_{\mathrm{B}^\pm} = f_{\mathrm{B}^0}$:

$$\frac{f_{\mathrm{B}^0}}{f_{\Lambda^0_\mathrm{b}}} = 2.5 \pm 0.2 \tag{5.9}$$

which is just the inverse of Equation 5.4.

Finally, the ratio of photon efficiency corrections, $(C_\gamma)^{N_\gamma(\mathrm{B}^0)}/(C_\gamma)^{N_\gamma(\Lambda^0_\mathrm{b})}$, is determined. Here, N_γ is the number of photons in the decay, so is equal to one for $\eta' \to \pi^+\pi^-\gamma$ decays and two for photons from η and π^0 decays. For the $\Lambda^0_\mathrm{b} \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, there will be no correction factor applied, since there are the same number of photons in the control channel and the signal channel, and the photon p_T distribution is similar. For the $\Lambda^0_\mathrm{b} \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) and $\Lambda^0_\mathrm{b} \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) decays, a correction factor of $1/C_\gamma$ is applied, where C_γ is the efficiency correction given in [101], which is calculated in bins of photon p_T. The correction is shown as a function of p_T in Figure 5.3, and it can be seen it is fairly flat and close to one. The correction is then given by

$$\frac{1}{C_\gamma} = \sum_i \frac{N_i(\mathrm{B}^0)}{N_i(\Lambda^0_\mathrm{b})} \times \frac{N_{tot}(\Lambda^0_\mathrm{b})}{N_{tot}(\mathrm{B}^0)} \times \frac{1}{C_{i\gamma}} \times \frac{N_i(\mathrm{B}^0) + N_i(\Lambda^0_\mathrm{b})}{N_{tot}(\mathrm{B}^0) + N_{tot}(\Lambda^0_\mathrm{b})}, \tag{5.10}$$

where the sum is over the bins of p_T, $C_{i\gamma}$ is the correction factor in p_T bin i, N_i is the

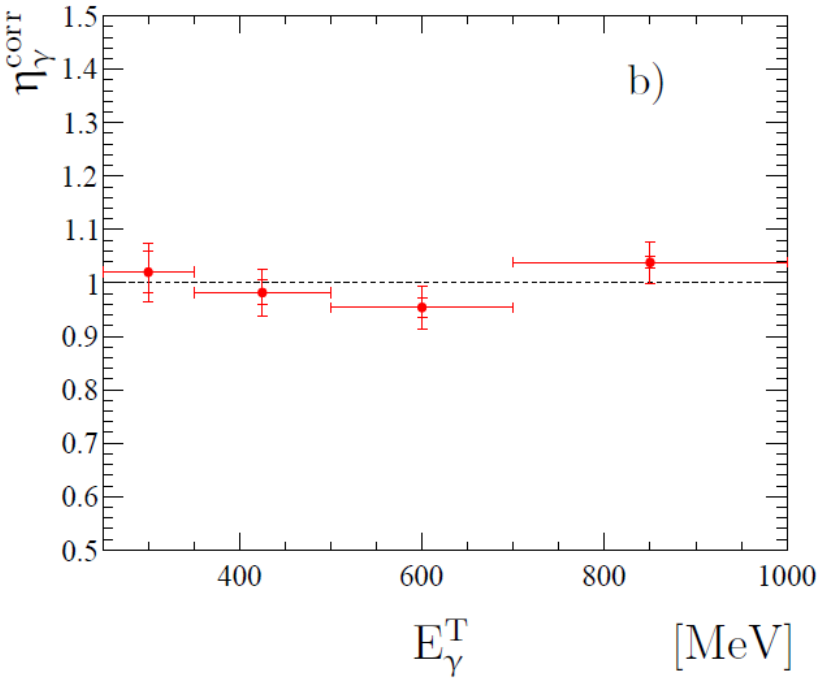

Figure 5.3: Photon reconstruction correction factor as a function of photon p_T [101].

number of events in the bin and N_{tot} is the total number of events. The correction has been calculated to be 0.96 ± 0.04 for the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) decay and 1.07 ± 0.04 for the $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) decay [101].

5.5 Selection

The selection applied to the data to reduce the number of background events is applied in several stages. The first stage is the trigger selection, which is followed by a loose kinematic selection to reduce the data sample to a reasonable size. This is followed by a PID selection, which identifies the correct signal tracks, and a multivariate selection which produces a powerful discriminant to separate signal and background. Finally, an additional selection is applied for events which contain multiple candidates. This section describes each of these stages.

5.5.1 Trigger selection

The first selection that the data must pass is the trigger selection, which is described in detail in Section 3.10. The selection is complicated for this analysis by a change in the trigger conditions in the 2012 data samples. During the data taking in 2011 and the first part of 2012 there was a bug in the trigger which meant that long lived DD particles, like the K^0_S and Λ^0 particles, were not used in the trigger selection. Since the signal modes rely heavily on the presence of the K^0_S to trigger, this has a large impact on the efficiency of the trigger selection on signal events. The bug was fixed during a technical stop in June 2012, so the efficiency for triggering on DD events will be higher for the latter 2/3 of the data sample. Calculations of the trigger efficiencies, which are shown in Section 6.1, must therefore be split into three distinct periods: 2011, pre-June 2012 (2012a), and post-June 2012 (2012b).

5.5.2 Pre-selection

In order to reduce the size of the data samples a loose pre-selection, known as the stripping selection, is applied to the data that passes the trigger selection. Requirements are placed on the tracks used to reconstruct signal candidates to ensure they are associated with genuine particles in the detector. It is possible that a random selection of hits in the detector can be reconstructed as tracks. These are referred to as "Ghost" tracks, and are removed from the data sample by placing cuts on the χ^2/ndf of the track fit, and on a multivariate quantity, called the "Ghost Probability", which defines the probability that a track is a ghost track. Tracks which pass these requirements are used to reconstruct the intermediate resonances as described above. Requirements are then placed on the minimum p_T of the tracks, the photons and reconstructed resonances. Signal particles are generally produced with higher p_T than background events, and so a minimum requirement on this quantity reduces the amount of background in the sample. To further reduce this, requirements are placed on the reconstructed resonances.

Table 5.3: Summary of the pre-selection requirements for $B^0 \to K_S^0 \eta'$ and $\Lambda_b^0 \to \Lambda^0 \eta'$ lines.

Particle	Variable	Cut
B^0	χ^2_{vtx} <	15
	DOCA χ^2 <	20
	p_T >	1500 MeV
	DIRA >	0.9995
	χ^2_{IP} <	20
	m(B^0)±	500 MeV
K_S^0	χ^2_{vtx} <	15
	DOCA χ^2 <	25
	p_T >	1200 MeV
	FD χ^2 >	50 (LL) 20 (DD)
	m(K_S^0)±	14 MeV (LL) 23 MeV (DD)
Λ_b^0	χ^2_{vtx} <	20
	DOCA χ^2 <	20
	p_T >	1000 MeV
	DIRA >	0.9995
	χ^2_{IP} <	20
	m(Λ_b^0)±	750 MeV
Λ^0	χ^2_{vtx} <	15
	DOCA χ^2 <	30 (LL) 25 (DD)
	p_T >	1000 MeV
	m(Λ^0)±	15 MeV (LL) 20 MeV (DD)
η'	χ^2_{vtx} <	15
	DOCA χ^2 <	20
	p_T >	2000 MeV
	m(η')±	150 MeV
η	p_T >	200 MeV
	m(η)±	50 MeV
γ	p_T >	200 MeV
Track	p_T >	300 MeV
	Ghost Prob <	0.5
	χ^2/ndf <	3

Table 5.4: Summary of the pre-selection requirements in $\Lambda_b^0 \to \Lambda^0 \eta$ lines.

Particle	Variable	Cut
Λ_b^0	χ^2_{vtx} <	20
	DOCA χ^2 <	20
	p_T >	1000 MeV
	χ^2_{IP} <	6
	m(Λ_b^0)±	750 MeV
Λ^0	χ^2_{vtx} <	15
	DOCA χ^2 <	30 (LL) 25 (DD)
	p_T >	1000 MeV
	m(Λ^0)±	15 MeV (LL) 20 MeV (DD)
η	χ^2_{vtx} <	15
	DOCA χ^2 <	20
	p_T >	2000 MeV
	m(η)±	150 MeV
π^0	p_T >	200 MeV
	m(π^0)±	50 MeV
γ	p_T >	200 MeV
Track	p_T >	300 MeV
	Ghost Prob <	0.5
	χ^2/ndf <	3

In background events, the resonances can be reconstructed using "combinatorial tracks", i.e. tracks that originate from the underlying event, and not the decay of the intermediate particles. To remove these events, requirements are placed on the quality of the reconstructed vertex, χ^2_{vtx}, the χ^2 of the Distance of Closest Approach (DOCA) of two tracks to each other, and the invariant mass of the reconstructed candidate. Reconstructed Λ_b^0 and B^0 candidates must satisfy two further criteria to ensure they originate from the PV. A requirement is placed on the Direction Angle (DIRA) of the candidates, which is defined as the cosine of the angle between the momentum vector and the vector that joins the PV to the decay vertex of the B^0 or Λ_b^0 candidate. If the candidate originates from the PV as expected, this angle will be close to zero, and the cosine will be close to one. If the candidate is built from combinatorial tracks then the momentum vector will not necessarily point to the PV. Finally, a requirement is placed on the χ^2 of the impact parameter, χ^2_{IP}, which is defined as the difference in the χ^2 of the PV when the particle of interest is added to the fit. For particles originating from the PV this difference will be small.

All the cuts applied in the stripping selection are shown in Table 5.3 for $B^0 \to K_S^0 \eta'$ and $\Lambda_b^0 \to \Lambda^0 \eta'$ candidates, and in Table 5.4 for $\Lambda_b^0 \to \Lambda^0 \eta$ candidates.

In addition to the stripping selection, there is an extra pre-selection applied to the data. This includes kinematic cuts which are not easy to apply at the stripping stage. The χ^2 of the decay vertex of the reconstructed B^0 is required to be less than 20 and the Confidence Level (CL) of photons is required to be greater than 0.05. The CL is a multivariate quantity formed from a combination of information from the calorimeters, and is expressed as a probability that the particle is a single photon, and not, for example, a misidentified π^0. A cut on the flight distance, $c\tau$, calculated relative to the refitted PV, is placed on the B^0 in the form $\ln(c\tau/\sigma(c\tau)) > 2$. Finally the χ^2/ndf of the fit from DecayTreeFitter is required to be less than 20. Since these cuts are similar to the stripping selection, they are treated together, and only the efficiency of the combined selection will be calculated.

5.5.3 PID selection

After the stripping selection, PID cuts are applied to charged tracks to reduce the background from misidentified particles. These cuts make use of the excellent particle identification provided by the RICH detectors, as discussed in Section 3.6. The RICH detectors make use of various Cherenkov radiators, and by comparing the Cherenkov angle with the track momentum, it is possible to see some separation between pions, kaons, protons and to some extent electrons and muons. Each radiator shows good separation in different momentum regimes. One can use a Neural Network (NN) to construct a probability that the track corresponds to a given particle species. The NN is trained using information from the RICH detectors, along with other subdetectors such as the SPD, and tracking system to provide information on the quality of the track. For each track, the NN then produces a $ProbNN_{\pi,\mathrm{K,p}}$ variable corresponding to the probability that the track is from a pion, kaon or proton respectively. So for a real pion, we expect $ProbNN_{\pi}$ to peak at a value of one, and $ProbNN_{\mathrm{K,p}}$ to both peak at a value of zero. These PID variables depend on the kinematics of the track, particularly on variables such as momentum, p, and pseudorapidity, η.

It is well-known within the collaboration that the MC does not model well the PID distributions of particles. One of the variables which is used to train the NN is the number of hits observed in the SPD subdetector. It has been shown in Chapter 4 that there is a disagreement between the number of hits in the SPD observed in data and in the MC. This leads to a difference in the $ProbNN$ variables calculated using the MC samples compared with data. Because of this, a data driven method is used to optimise the selection and calculate selection efficiencies, using a calibration data sample. Calibration pions come from a sample of $\mathrm{D}^{*0} \to \mathrm{K}^{+}\pi^{-}$ decays, which have very little contamination from $\mathrm{D}^{*0} \to \mathrm{K}^{-}\pi^{+}$ decays and so the charge of the track defines the track type, and protons from Λ_{c}^{+} decays, with each sample coming from 2011 and 2012 data separately. The combinatoric background from these decays is removed using the *sPlot* technique [102], described in Section 6.4.1. The important

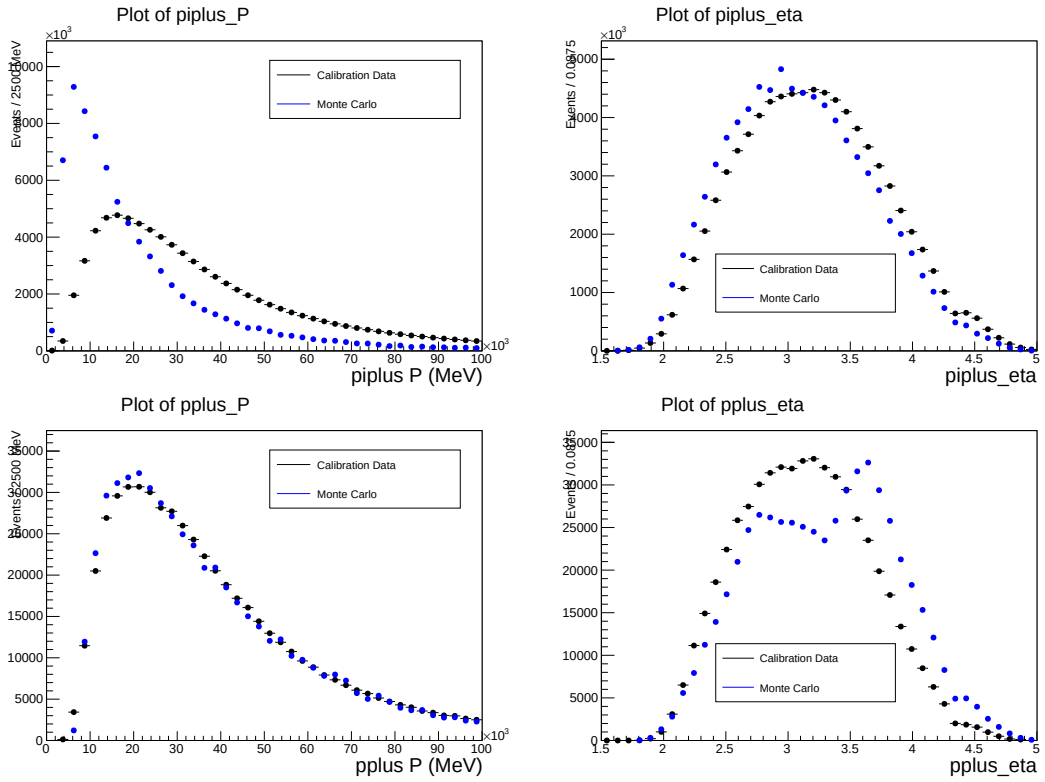

Figure 5.4: Distribution of p and η comparing the calibration data samples (black) and signal MC samples (blue) for pions (top) and protons (bottom).

issue when using these calibration samples is that the calibration tracks will have different kinematics from the signal tracks, and this will have an effect on the PID distributions. Figure 5.4 shows the p and η distributions of the π tracks from signal $B^0 \to K_S^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays and p tracks from the $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays, along with π and p from the calibration data. Some differences in the distributions are observed, which will affect the PID distributions of these tracks.

The first step of the analysis is to reweight the calibration data such that the kinematic distributions match those of the signal samples. This is done by binning both the calibration data and MC signal in two-dimensional bins of p and η. The binning scheme must be chosen such that there are sufficient statistics in each bin to calculate a weight, and so a non-uniform scheme is chosen, with finer binning at low p and in areas where the PID efficiency varies rapidly, and coarser binning at high p where statistics are more limited. 24 bins in p, and 14 bins in η are chosen, and the

bin boundaries are:

$$p\ (\mathrm{GeV}/c) : [\ 2\ ,\ 3\ ,\ 3.5\ ,\ 4\ ,\ 4.5\ ,\ 5\ ,\ 5.5\ ,\ 6\ ,\ 6.5\ ,\ 7\ ,\ 7.5\ ,\ 8.2\ ,\ 9.3\ , \\ 12\ ,\ 13.5\ ,\ 15.6\ ,\ 17.5\ ,\ 20\ ,\ 25\ ,\ 30\ ,\ 35\ ,\ 40\ ,\ 55\ ,\ 100\ ,\ 400\ ] \tag{5.11}$$

for pions, and:

$$p\ (\mathrm{GeV}/c) : [\ 8\ ,\ 9\ ,\ 10\ ,\ 12\ ,\ 13\ ,\ 14\ ,\ 15\ ,\ 16\ ,\ 17\ ,\ 18\ ,\ 19\ ,\ 20\ , \\ 22.5\ ,\ 25\ ,\ 27.5\ ,\ 30\ ,\ 35\ ,\ 40\ ,\ 45\ ,\ 50\ ,\ 60\ ,\ 70\ ,\ 100\ ,\ 400\ ] \tag{5.12}$$

for protons. For both pions and protons, 14 bins in η are chosen:

$$\eta : [\ 1.5\ ,\ 2\ ,\ 2.25\ ,\ 2.35\ ,\ 2.5\ ,\ 2.75\ ,\ 3.0\ ,\ 3.1\ ,\ 3.25\ ,\ 3.4\ , \\ 3.5\ ,\ 3.75\ ,\ 4\ ,\ 4.25\ ,\ 5\ ] \tag{5.13}$$

Figure 5.5 shows the distribution of events for pions and protons from both the MC signal samples and the calibration data samples, binned in 2D according to the scheme described above. It can be seen that generally the coverage is quite good. There are only a few events in the corners of the plots (i.e. extremely high/low p or η), but since events in these bins should be weighted to zero this should not be a problem.

These distributions are used to reweight the calibration tracks to match the kinematics of each track in the signal sample separately. Each signal track is binned as above, and for each event in the calibration samples, a weight is calculated as a function of p and η of that event:

$$W(p,\eta) = \frac{N_{sig}(p,\eta)}{N_{cal}(p,\eta)} \times \frac{N_{cal}(tot)}{N_{sig}(tot)}, \tag{5.14}$$

where $N_{sig(cal)}(p,\eta)$ is the number of signal (calibration) events in a particular p,η bin and $N_{sig(cal)}(tot)$ is the total number of events in the signal (calibration) sample.

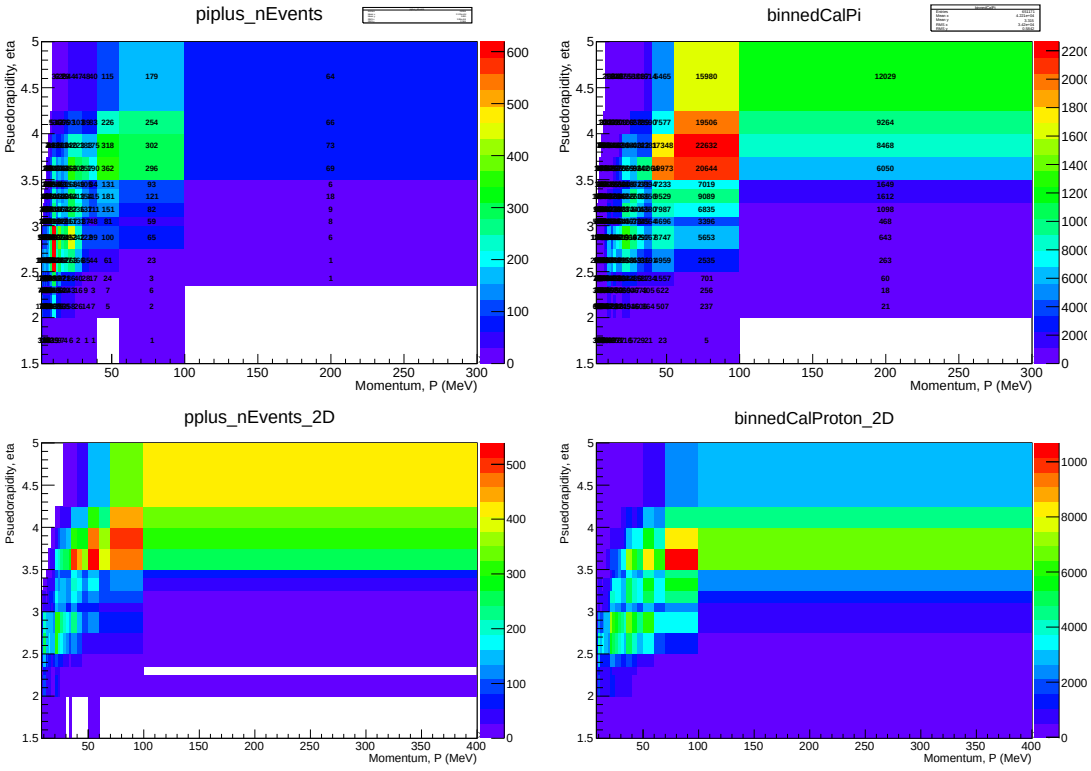

Figure 5.5: Distribution of pions (top) and protons (bottom) from MC signal (left) and calibration data (right), binned in two dimensions of p and η according to the schema described in the text.

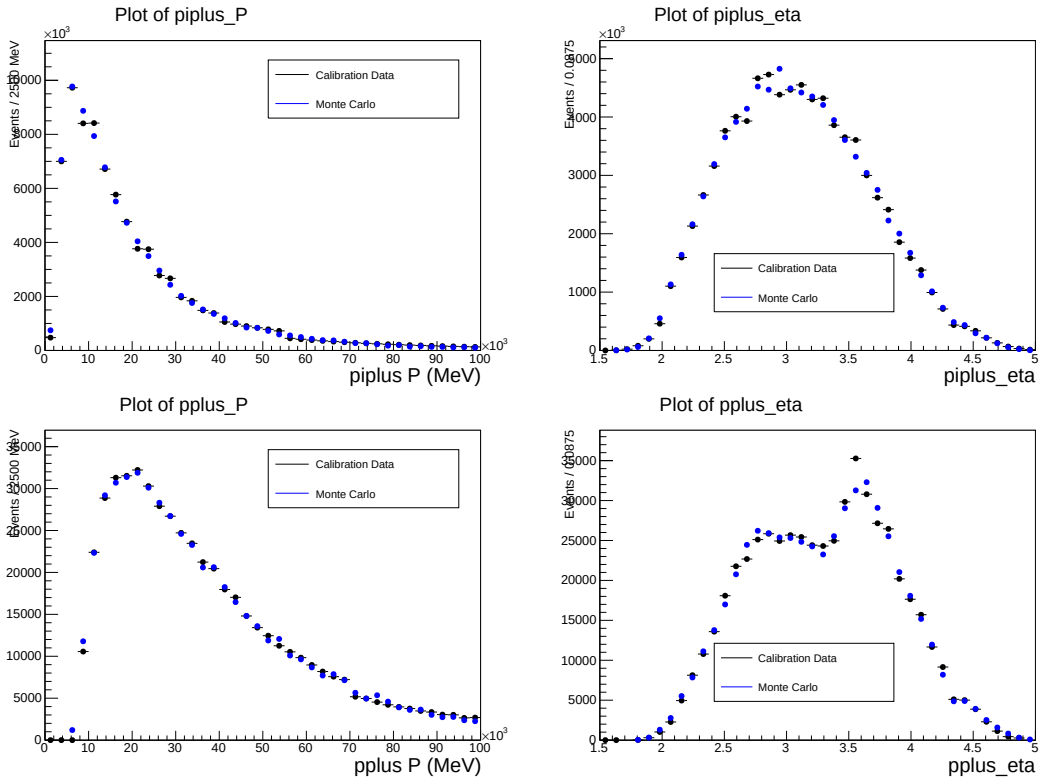

Figure 5.6: Distribution of p and η comparing the calibration data samples (black) and signal MC samples (blue) for pions (top) and protons (bottom), after the calibration data has been reweighted to match the signal kinematics.

Table 5.5: Summary of PID cuts.

Track	**Cut**
All $\pi^{\pm}$	$ProbNN_{\pi}$ >0.2
Long $\pi^{\pm}$	$ProbNN_{\mathrm{K}}$ <0.4
All p	$ProbNN_{\mathrm{p}}$ >0.1
Long p	$ProbNN_{\mathrm{K}}$ <0.8

Figure 5.6 shows the pion and proton kinematics after the reweighting has been applied to the calibration samples, demonstrating that the kinematics of these samples match more closely the kinematics of the signal samples.

Figures 5.7 and 5.8 show the distributions of various PID distributions of tracks in the LL and DD categories respectively. It can be seen that the agreement between MC (blue) and the calibration data (black) is not perfect, and although the differences are not too significant in the LL category, there are some large discrepancies in the proton $ProbNN$ variables, and also for pions from DD $\mathrm{K}_{\mathrm{S}}^0$ meson decays. We therefore need to reweight the MC to match the PID distributions of the calibration data.

Also shown in red on the plots is the distribution of PID variables for background events. These come from the sidebands of the data, and are events with high or low invariant $\mathrm{K}_{\mathrm{S}}^0\eta'$ invariant mass (or $\Lambda^0\eta^{(\prime)}$ invariant mass), and so these samples contain only background events, mainly combinatoric background, but also possible backgrounds with misidentified particles. By comparing the red and black curves, it is possible to find cuts on the PID distributions which remove background events, whilst keeping a large fraction of signal events. It should be noted that the $ProbNN_{\mathrm{K}}$ distributions of downstream pions and protons show no separation between signal and background, and so a cut is placed on this quantity for long tracks only. For all distributions, the optimal cut was chosen by eye to remove as much of the background as possible, while retaining most of the signal events. The cuts chosen are summarised in Table 5.5 and are shown as green dashed lines in the plots.

To check that the binning in p and η we have chosen is suitable for these PID cuts, the

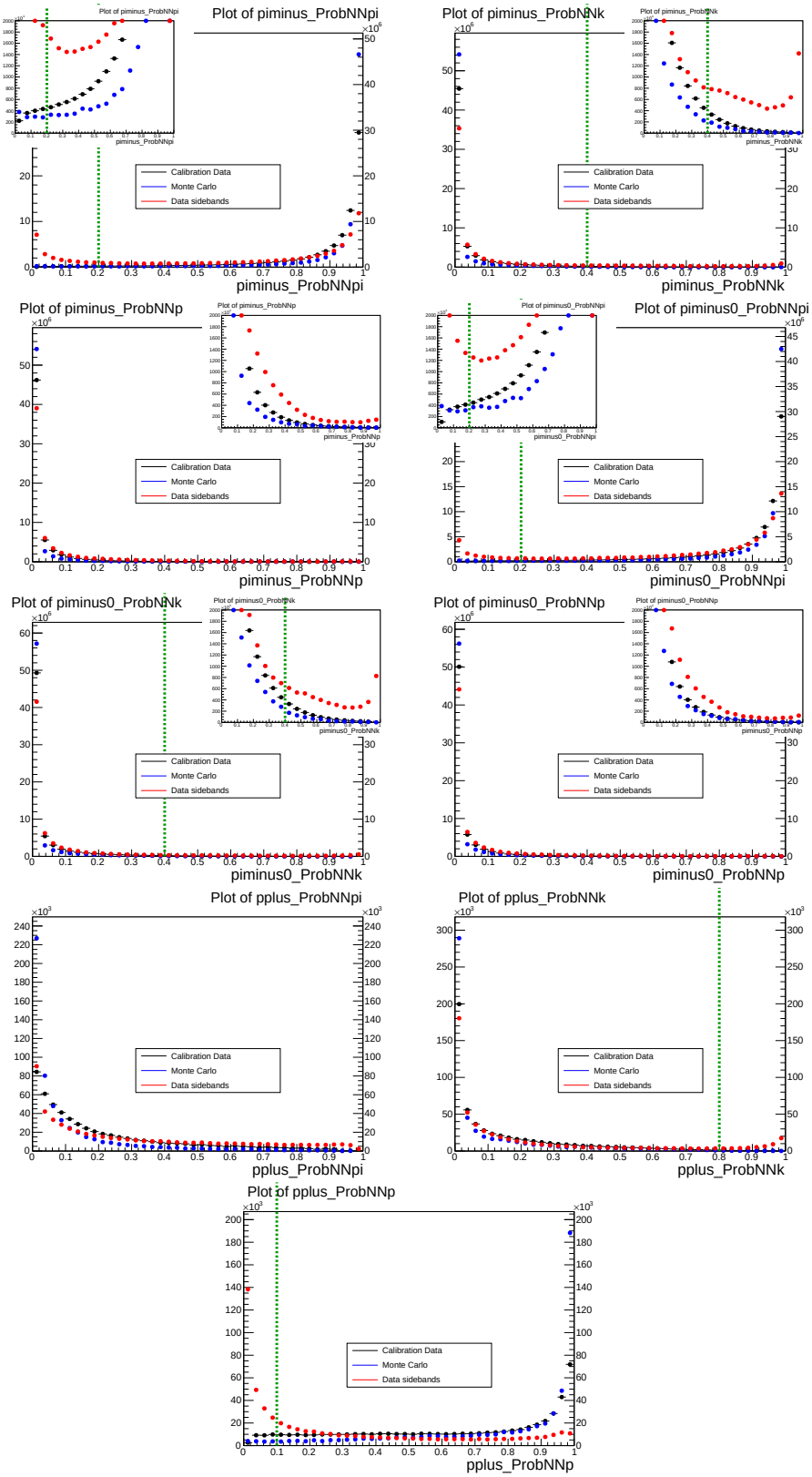

Figure 5.7: PID distributions for pions from η' decays (top), pions from K^0_S decays (middle) and protons (bottom), in the LL category. The plots show the kinematically reweighted calibration data (black), the MC signal (blue), and the background from the sidebands of data (red). The insert of each plot shows the same plot zoomed on the y axis, which shows the number of events per probability of 0.025.

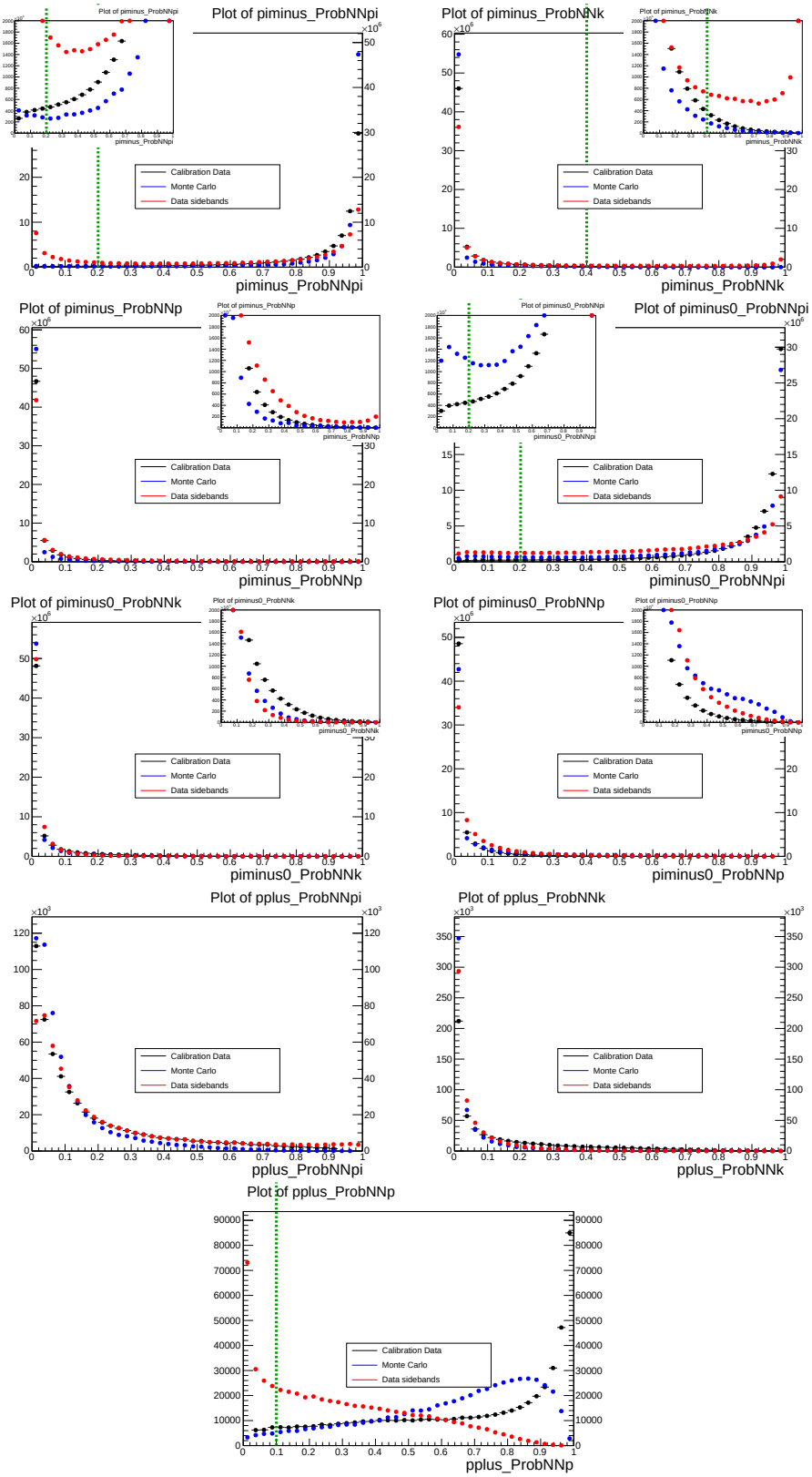

Figure 5.8: PID distributions for pions from η' decays (top), pions from K_S^0 decays (middle) and protons (bottom), in the DD category. The plots show the kinematically reweighted calibration data (black), the MC signal (blue), and the background from the sidebands of data (red). The insert of each plot shows the same plot zoomed on the y axis, which shows the number of events per probability of 0.025.

overall efficiency of these cuts is measured to make sure it does not vary significantly across bin boundaries. The efficiencies are calculated using the calibration samples, and are plotted as a function of p in each bin of η, and also as a function of η in each bin of p. The efficiencies are shown in Figure 5.9 for pions and Figure 5.10 for protons. The dashed lines show the bin boundaries chosen in the variable being plotted. To quantify the difference across the bins, we measure the PID efficiency at each of the bin boundaries, and calculate the difference between the efficiencies measured at consecutive boundaries. In this way we have a measure of the variation of PID efficiency in each bin, and we take an average of the absolute efficiency differences in each bin, weighted by the number of events in that bin, to find an average variation across all bins. For pions, this gives an average variation of 0.32%, and for protons the average variation is 2.7%. The efficiency shows a stronger dependence on PID cuts for protons, but if more bin boundaries are added then the statistics in each bin is too low and so a compromise must be reached and the 2.7% variation must be accepted.

In this analysis, the PID cuts are applied before variables are used in a multivariate selection (described in Section 5.5.4). It is therefore important that the distributions of the MC are correct, so that the events that pass the PID selection are a true representation of the signal events that would pass the selection. In order to do this, we reweight the MC to match the calibration data. This is done in the same way as the kinematic reweighting: the MC and calibration samples are binned in PID variables, and for each variable an event weight is calculated corresponding to the ratio of the number of events in the particular PID bin. The overall event weight is then the product of the weights calculated using the $ProbNN_{\pi}$ variable for signal pions and the $ProbNN_{\mathrm{p}}$ variable for signal protons. Since the PID variables are highly correlated with each other, reweighting one variable does improve the agreement in the other variables. However, as the correlation is not 100% there will still be some residual difference in the $ProbNN_{\mathrm{K}}$ variables. There is very little correlation in the PID variables between the different signal particles, and so the overall weight is the product of the weight for each track. Figures 5.11 and 5.12 show the PID

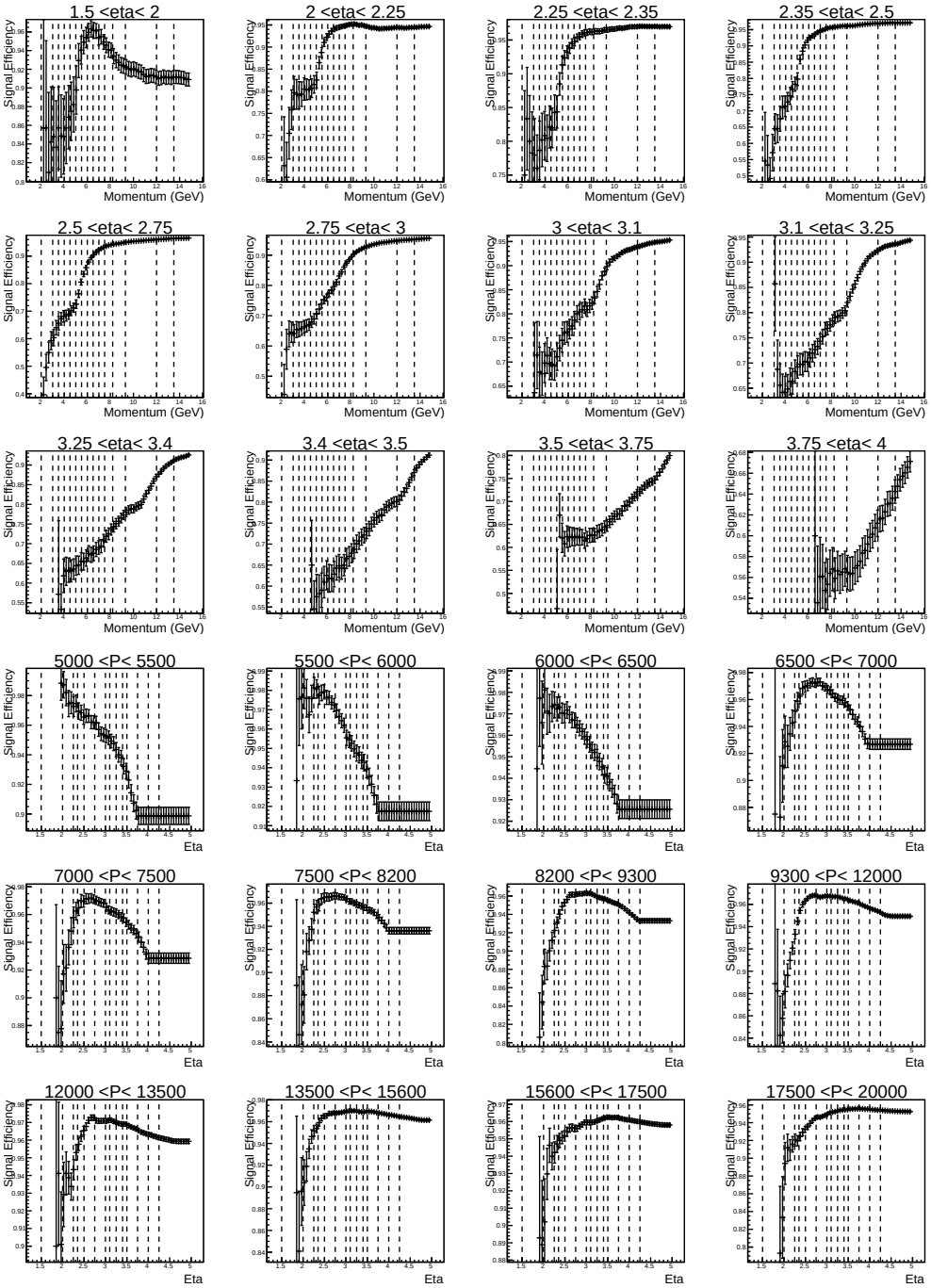

Figure 5.9: Efficiency of PID cuts in bins of p and η for calibration pions. Top) shows the efficiency as a function of p in bins of η. Bottom) shows the efficiency as a function of η in bins of p. The dashed lines show bin boundaries in the other variable.

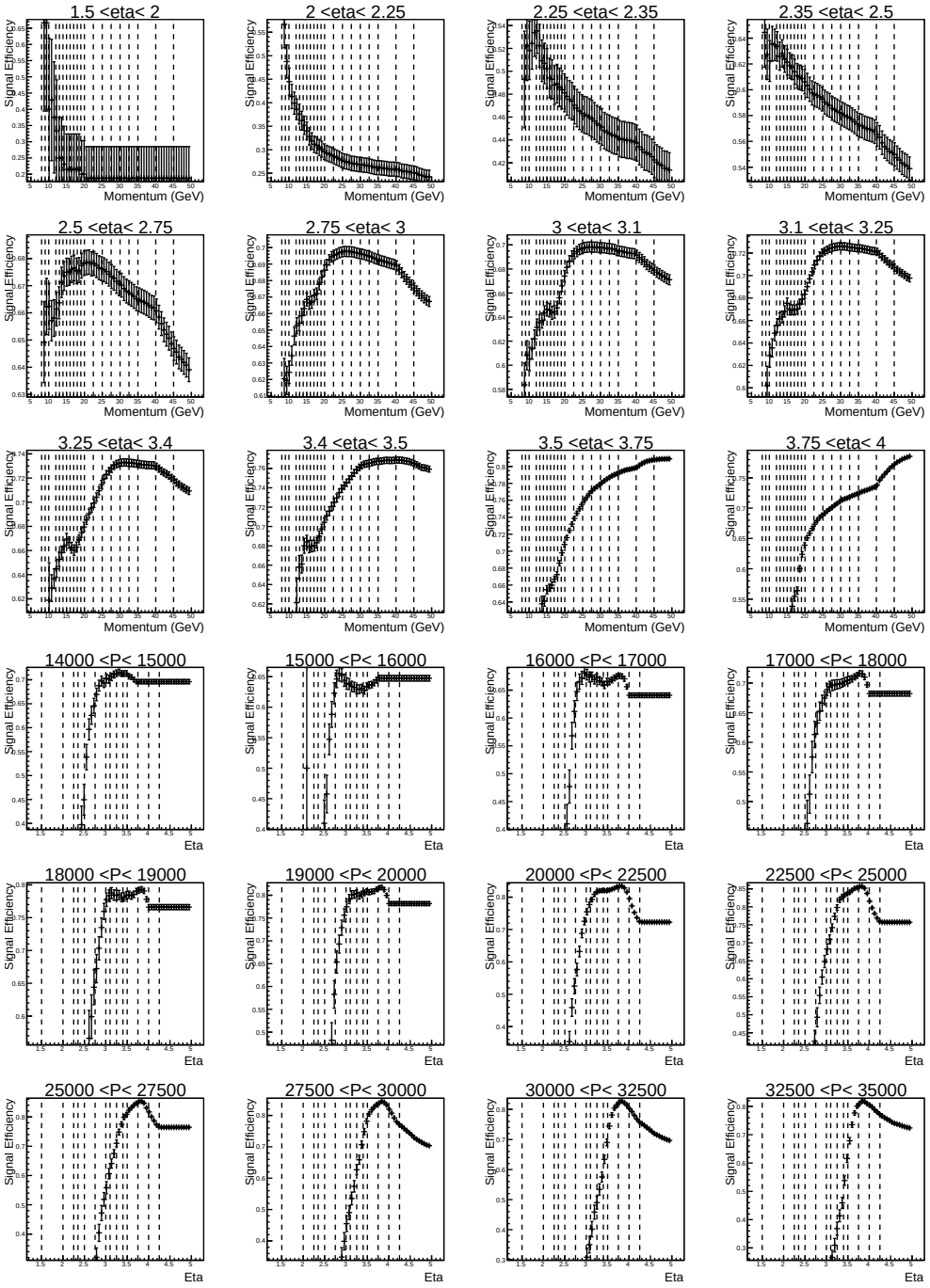

Figure 5.10: Efficiency of PID cuts in bins of p and η for calibration protons. Top) shows the efficiency as a function of p in bins of η. Bottom) shows the efficiency as a function of η in bins of p. The dashed lines show bin boundaries in the other variable.

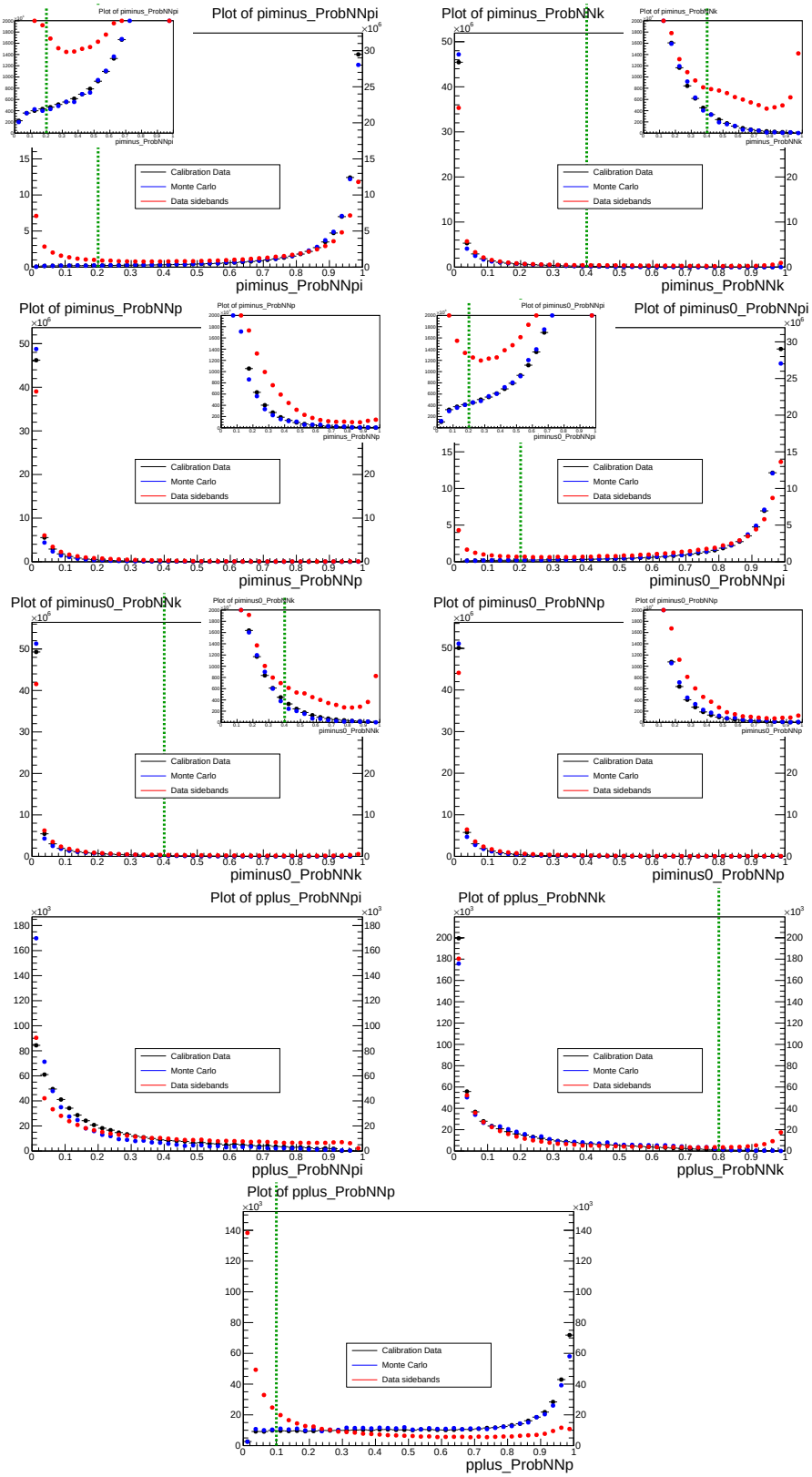

Figure 5.11: PID distributions for pions from η' decays (top), pions from K_S^0 decays (middle) and protons (bottom), in the LL category. The plots show the reweighted calibration data (black), the PID reweighted MC signal (blue), and the background from the sidebands of data (red). The insert of each plot shows the same plot zoomed on the y axis, which shows the number of events per probability of 0.025.

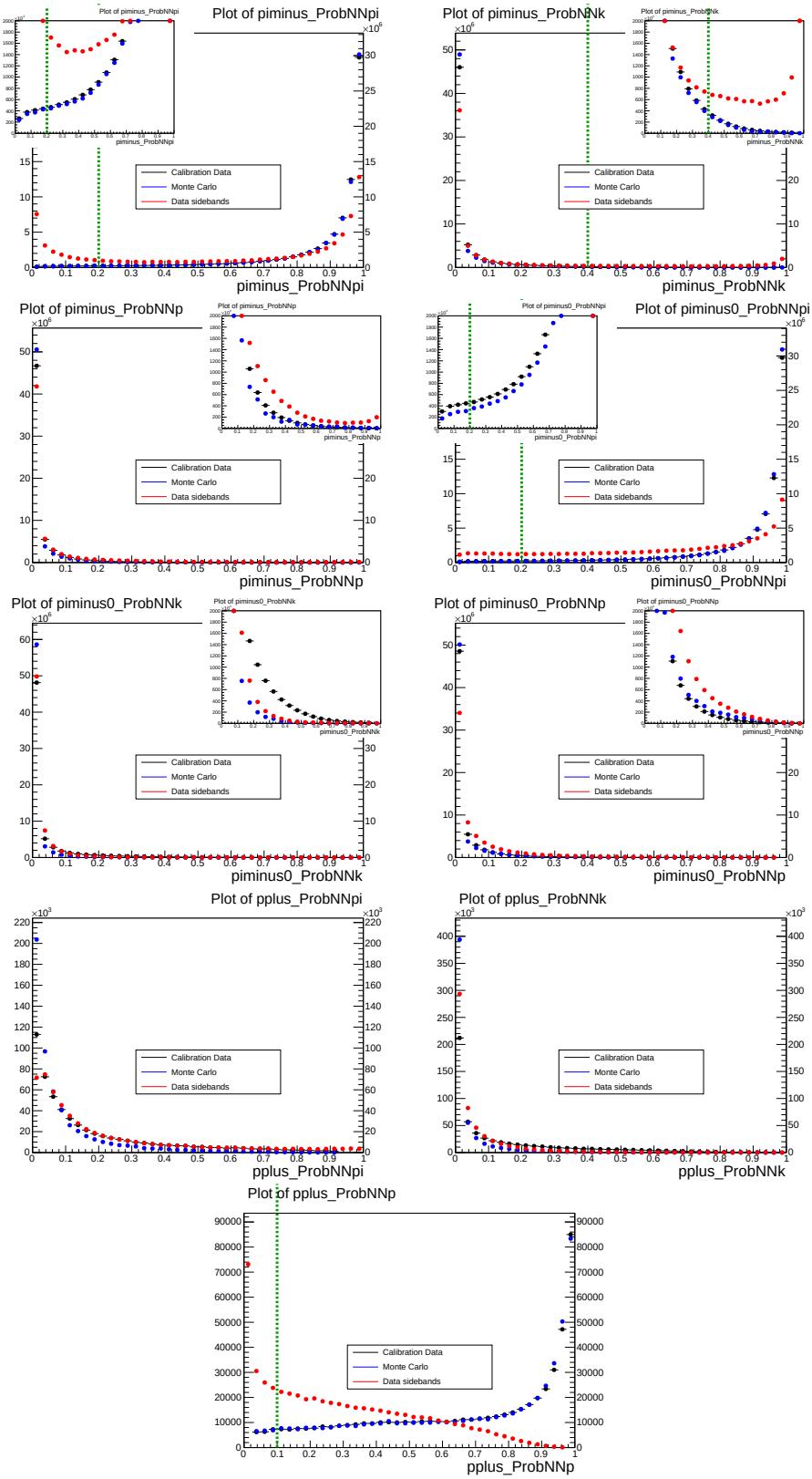

Figure 5.12: PID distributions for pions from η' decays (top), pions from K^0_S decays (middle) and protons (bottom), in the DD category. The plots show the kinematically reweighted calibration data (black), the PID reweighted MC signal (blue), and the background from the sidebands of data (red). The insert of each plot shows the same plot zoomed on the y axis, which shows the number of events per probability of 0.025.

distributions after the reweighting is applied. It is clear that the agreement between MC and the calibration data is now much better, for the variables for which cuts are applied. For the variables for which cuts are not applied, i.e. the $ProbNN_p$ variables and the $ProbNN_K$ variables for downstream tracks, the agreement is not good and is actually worse after the reweighting. This is to be expected, as the procedure used does not account for correlations between the different $ProbNN$ variables. So reweighting the $ProbNN_\pi$ variable will also have an effect on the $ProbNN_K$ and $ProbNN_p$ distributions. However, this is not a problem, as no cut is placed on this variable, and so agreement between data and MC is not important. The differences are quantified in Table 5.6, which shows the weighted average difference between the number of events in the calibration data and the MC across each bin in the PID distributions. It can be seen that, generally, the difference is a few percent, except for tracks from signal Λ^0 decays.

The same procedure was repeated for all decay channels, and the PID distributions were examined for each track. No difference from what has been shown here was observed, which demonstrates the agreement for each PID variable across the whole distribution. However, it is also interesting to check the agreement in the region where we are applying the cut. One way to do this is to measure the total efficiency of the PID selection when applied to all signal particles. This is calculated by applying the selection to the reweighted MC samples and using the calibration samples, and the differences are presented in Table 5.7. The average difference between the measured efficiencies is 1.3% for B^0 decays in the LL category, 0.6% for B^0 decays in the DD category, 1.4% for Λ_b^0 decays in the LL category and 0.34% for Λ_b^0 decays in the DD category. These remaining differences will be taken into account in the systematic uncertainties.

5.5.4 Multivariate selection

This analysis makes use of a multivariate selection to separate signal candidates from the combinatorial background. The selection considers many variables which have

Table 5.6: Weighted average difference over all bins in PID and kinematic distributions of reweighted Monte Carlo compared with Calibration data for $B^0 \to K^0_S \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) and $\Lambda^0_b \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays. $\pi^\pm$ refers to pions from η' decays, and $\pi^\pm_{V0}$ refers to pions from K^0_S or Λ^0 decays.

Channel	**Cat**	**Year**	**Track**	$ProbNN_\pi$ (%)	$ProbNN_K$ (%)	$ProbNN_p$ (%)
B^0	LL	2011	π^+	3.242 ± 0.003	2.078 ± 0.002	–
			π^-	2.898 ± 0.003	1.381 ± 0.002	–
			π^+_{V0}	1.432 ± 0.002	6.973 ± 0.004	–
			π^-_{V0}	1.557 ± 0.002	6.605 ± 0.004	–
B^0	LL	2012	π^+	1.757 ± 0.001	4.860 ± 0.003	–
			π^-	0.351 ± 0.001	5.431 ± 0.003	–
			π^+_{V0}	1.162 ± 0.001	5.050 ± 0.003	–
			π^-_{V0}	1.428 ± 0.001	4.296 ± 0.002	–
B^0	DD	2011	π^+	0.719 ± 0.001	5.625 ± 0.004	–
			π^-	0.093 ± 0.000	5.364 ± 0.004	–
			π^+_{V0}	2.719 ± 0.003	–	–
			π^-_{V0}	1.262 ± 0.002	–	–
B^0	DD	2012	π^+	2.097 ± 0.002	6.048 ± 0.003	–
			π^-	1.507 ± 0.001	6.774 ± 0.003	–
			π^+_{V0}	3.633 ± 0.002	–	–
			π^-_{V0}	1.506 ± 0.001	–	–
Λ^0_b	LL	2011	π^+	6.265 ± 0.004	2.470 ± 0.003	–
			π^-	1.919 ± 0.002	1.238 ± 0.002	–
			p	–	19.690 ± 0.006	3.359 ± 0.003
			π^-_{V0}	7.497 ± 0.004	3.335 ± 0.003	–
Λ^0_b	LL	2012	π^+	2.201 ± 0.002	3.572 ± 0.002	–
			π^-	2.149 ± 0.002	5.192 ± 0.003	–
			p	–	12.169 ± 0.003	4.173 ± 0.002
			π^-_{V0}	6.899 ± 0.003	10.532 ± 0.003	–
Λ^0_b	DD	2011	π^+	1.348 ± 0.002	4.594 ± 0.003	–
			π^-	2.389 ± 0.002	5.053 ± 0.004	–
			p	–	–	1.863 ± 0.002
			π^-_{V0}	2.976 ± 0.003	–	–
Λ^0_b	DD	2012	π^+	0.520 ± 0.001	9.690 ± 0.004	–
			π^-	3.945 ± 0.002	9.516 ± 0.004	–
			p	–	–	2.154 ± 0.001
			π^-_{V0}	2.282 ± 0.002	–	–

Table 5.7: Difference in PID efficiencies measured in reweighted MC and calibration data.

Channel	**Cat**	**Year**	$\epsilon_{\rm PID}$**(Calib) (%)**	$\epsilon_{\rm PID}$**(MC) (%)**	**Difference (%)**
$B^0 \to K_S^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	2011	89.683 ± 0.006	89.8 ± 0.2	0.2 ± 0.2
	LL	2012	90.109 ± 0.003	88.6 ± 0.2	–1.5 ± 0.2
	DD	2011	90.489 ± 0.003	91.8 ± 0.1	1.3 ± 0.1
	DD	2012	91.148 ± 0.002	91.6 ± 0.1	0.4 ± 0.1
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	2011	87.4 ± 0.5	85.1 ± 0.2	–2.4 ± 0.5
	LL	2012	88.16 ± 0.04	86.5 ± 0.2	–1.7 ± 0.2
	DD	2011	88.7 ± 0.4	89.1 ± 0.1	0.4 ± 0.4
	DD	2012	90.1 ± 0.9	89.9 ± 0.1	–0.2 ± 0.9
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\eta)$	LL	2011	88.9 ± 0.7	86.3 ± 0.3	–2.6 ± 0.7
	LL	2012	89.76 ± 0.08	88.8 ± 0.3	–1.0 ± 0.3
	DD	2011	90.5 ± 0.6	90.3 ± 0.2	–0.2 ± 0.7
	DD	2012	91.8 ± 1.9	91.0 ± 0.2	–0.7 ± 1.9
$\Lambda_b^0 \to \Lambda^0 \eta$ $(\eta \to \pi^+\pi^-\pi^0)$	LL	2011	88.3 ± 0.4	86.6 ± 0.2	–1.7 ± 0.4
	LL	2012	89.06 ± 0.03	87.2 ± 0.2	–1.8 ± 0.2
	DD	2011	89.6 ± 0.4	89.6 ± 0.1	0.0 ± 0.4
	DD	2012	91.4 ± 0.9	90.4 ± 0.1	–1.0 ± 0.9

some ability to discriminate between signal and background. By optimising and applying cuts to each of these variables in turn, the background will be reduced but the signal may also be reduced by an unacceptable amount. A multivariate selection will take all of these variables and combine them into one discriminating variable which has good separating power. By accounting for the correlations between these variables, the selection will generally reduce the background to an acceptable level whilst having a high efficiency for selecting signal events.

The specific multivariate technique adopted for this analysis is the BDT [103, 104] which is implemented using the TMVA package in the ROOT framework [105]. The BDT is trained using two mutually exclusive samples which contain only signal or only background events. The decision tree is built up by defining a set of variables, x_i, and applying cuts on the variables in turn to each sample. A typical decision tree is shown in Figure 5.13. The root node starts with a mix of signal events and background events. A cut variable is chosen and events enter one of two leaf nodes, depending on whether they pass or fail the cut. The variable and value of the cut

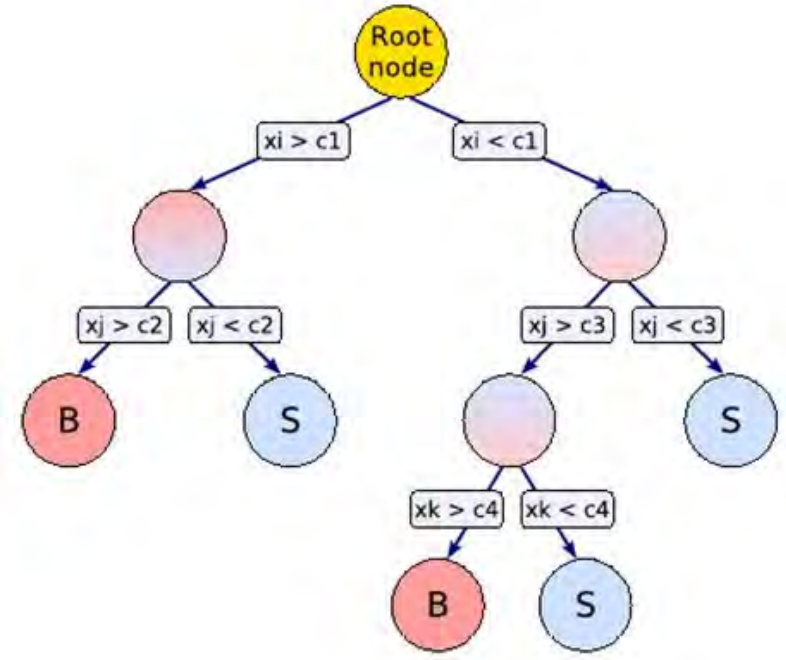

Figure 5.13: Typical make up of a decision tree.

is chosen to optimise the purity of the leaf node, defined as

$$P = \frac{\sum_s W_s}{\sum_s W_s + \sum_b W_b}, \tag{5.15}$$

where $\sum_{s(b)}$ is a sum over signal (background) events, and $W_{s(b)}$ is the event weight (e.g the PID weight for this analysis). If the leaf contains purely signal or background events then the quantity $P(1-P) = 0$, and so a leaf is defined as signal or background if this quantity is below a certain value. If it is above this value then a different variable is chosen and the same procedure is repeated for events in this leaf. The building of the tree can proceed in a number of ways: until each node satisfies the criteria of signal or background; up to a maximum tree depth, until there are too few events in a leaf to produce statistically meaningful results; or until there is a predefined number of leaves in the tree. Unless the first condition is chosen, the leaves which have not yet been defined will be classified as signal (background) if $P > (<)0.5$. The decision tree is applied to a sample which contains an unknown contribution of signal and background events. Each event in this sample follows the

cuts in the decision tree until it reaches a final leaf, and is classified as signal or background according to the leaf it lands on.

Decision trees are useful for combining variables in a way that exploits the correlations between variables. The main problem with this method is that the trees produced are generally unstable, with slight changes in the training samples producing very different decision trees. The method to overcome this is called "boosting", which builds up a large number of trees and combines the results of all these trees into one discriminant. In addition, the trees are built such that they learn from the performance of the previous tree so that each new tree performs better than the previous one. This technique is known as "Adaptive Boosting" [106] and is used in the training of the BDT used in this analysis.

The first tree is built in the same way as above, and a number of new quantities are defined for that tree which influence the training of the second tree. For each of the N events in the combined signal and background sample, the quantity I is defined, which is given the value of zero if the event lands in the correct leaf at the end of the tree (i.e. a signal events lands in a signal leaf or a background event in a background leaf), and a value of one otherwise. The performance of the tree is then defined by two variables:

$$err = \frac{\sum_{i=1}^{N} W_i \times I_i}{\sum_{i=1}^{N} W_i} \tag{5.16}$$

$$\alpha = ln\left(\frac{1 - err}{err}\right), \tag{5.17}$$

so that trees with a small number of incorrectly classified events have a larger value of α. The event weights are then changed to account for this, with the weight of each event becoming

$$W_i \to W_i \times e^{\alpha I_i} \tag{5.18}$$

By doing this, events which are classified incorrectly are given a larger weight, so that the next decision tree gives a higher priority to these events. This is then repeated for a large number of trees to get the best performing decision trees possible. These trees are combined into a BDT which is applied to an unknown sample of signal and

background events. Each event in this sample passes through each tree and is given a score of $T = 1$ if it lands in a signal leaf and $T = -1$ if it lands in a background leaf. The weighted sum of the score of each tree is calculated, and this is called the BDT_response,

$$\text{BDT\_response} = \sum_{m=1}^{N_{tree}} \alpha_m T_m. \tag{5.19}$$

Events which are more background like have a low BDT_response value, and signal like events have a high BDT_response value. Cutting on this value can therefore separate signal and background events.

Creating a BDT in this way produces a selection which is much more stable than using an individual decision tree. The only issue which still needs to be considered is that of overtraining. This can occur if only a small number of events is used to train the BDT. Due to the nature of the boosting, if there are too few events then the BDT will be trained on statistical fluctuations, and will give different results when applied to a different sample that does not share the same properties as those of the training sample.

In this analysis, a BDT is trained for each of the control channel and signal channels, separated into the LL and DD categories and 2011 and 2012 data taking periods. The specific configuration and performance of each BDT is described in this section.

MC samples are used to model the signal events, and the upper mass sidebands of data are used to model background events, where the sideband is defined as events with a reconstructed invariant mass more than 100 MeV above than the nominal B^0 or Λ_b^0 mass, which is more than three times the mass resolution. The pre-selection and PID cuts are applied to MC and data prior to the training of the BDT. This means the BDT is trained using events which look most similar to signal events in data. However, when the trigger cuts are applied, there are not enough events in the background samples to train a BDT, and the statistics in the MC sample is reduced by a factor of $\approx$4. For this reason, no trigger cuts are applied prior to training.

A consequence of this is that some of the distributions of variables used to train the

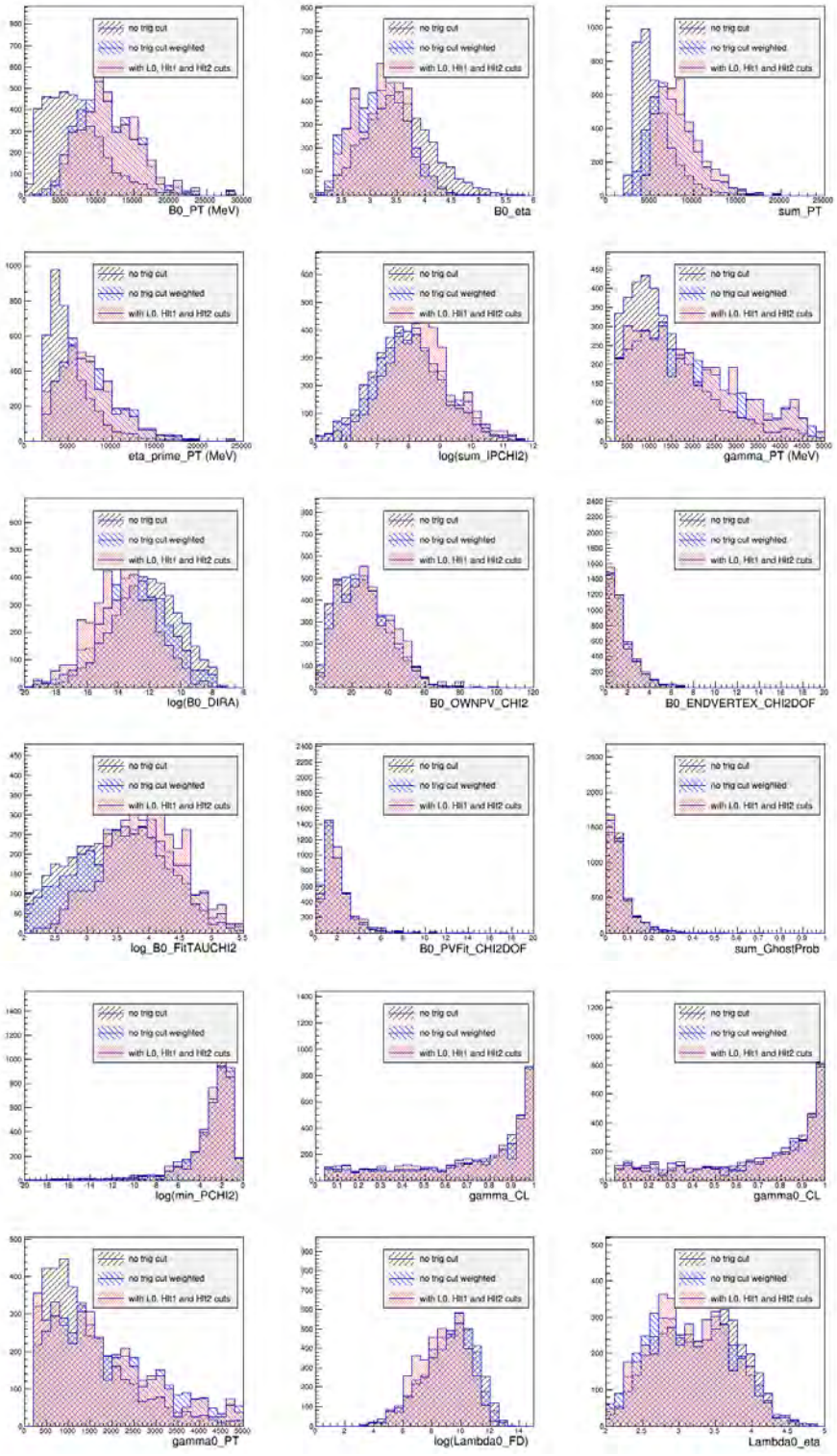

Figure 5.14: Comparison of distributions for $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\eta$) variables with different trigger cuts applied, using an MC signal sample.

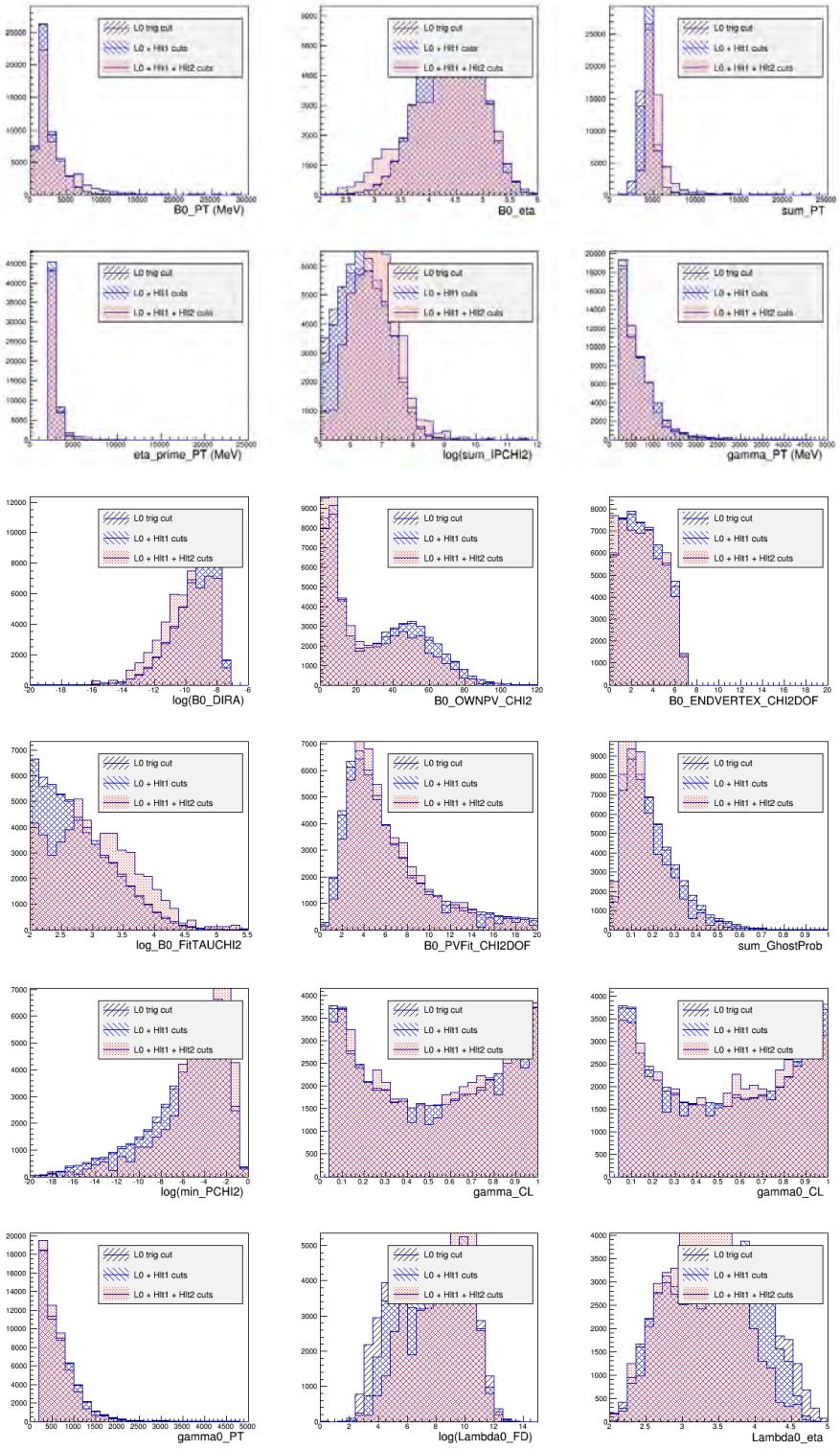

Figure 5.15: Comparison of distributions for $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\eta$) variables with different trigger cuts applied, using a background sample from the upper sidebands of data.

Table 5.8: Number of signal and background events used for training the BDTs for each channel, and trigger cuts applied prior to training.

Decay	Category	Year	max-depth	No. Signal	No. Background
$B^0 \to K_s^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	2011	3	11662	168343
		2012	3	23753	476429
	DD	2011	3	26141	150821
		2012	3	40573	421017
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	2011	3	13408	52246
		2012	3	13084	204289
	DD	2011	3	29001	53111
		2012	3	30883	163219
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\eta)$	LL	2011	3	4835	6176
		2012	3	5106	27530
	DD	2011	2	9213	4376
		2012	3	10166	14094
$\Lambda_b^0 \to \Lambda^0 \eta$ $(\eta \to \pi^+\pi^-\pi^0)$	LL	2011	2	17366	2298
		2012	3	15151	8994
	DD	2011	2	40389	2051
		2012	2	27538	6305

BDT will be different in the absence of trigger cuts , particularly p_T variables. To combat this, the distribution of B^0 p_T is compared with and without trigger cuts applied. An event weight is then calculated, which reproduces the p_T distributions when the trigger cuts are applied. This event weight is applied to the MC samples when the BDT is trained, and the reweighted distributions, as well as the distributions with and without trigger cuts, are shown in Figure 5.14. This reweighting is not applied to the combinatoric background sample, since the trigger cuts do not affect the p_T distributions of the background, as can be seen in Figure 5.15.

Finally, to deal with the fact the statistics available for training each channel is different, the parameters of the BDT training, including the number of trees used to train and the maximum depth of each tree, can be changed. Table 5.8 summarises which parameters are used for each channel, and how many signal and background events are used for training the BDT.

In order to make the best use of the statistics available, two BDTs are trained for each channel. Each event in the signal and data samples is assigned a random number between 0 and 1, calculated using the global event number. In this way,

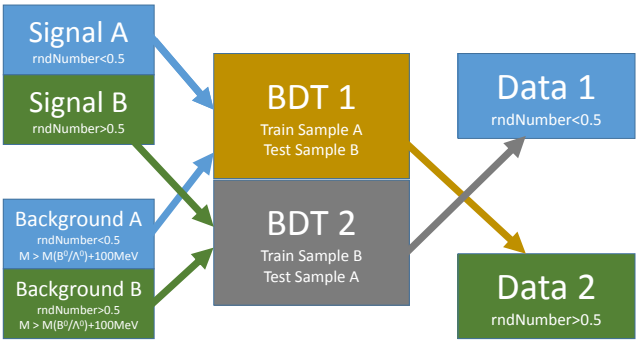

Figure 5.16: Schematic view of split BDT process.

the random numbers are reproducible, but it has been checked that they are not correlated with the event number. The samples are split into 2 sub-samples, based on whether the random number is less than 0.5 (sub-sample 1) or greater than 0.5 (sub-sample 2). A BDT is trained on sub-sample 1, with sub-sample 2 used to test there is no overtraining of the BDT. This is done by plotting the BDT response of both the test sample and training sample together and checking the distributions are the same. A second BDT is then trained using sub-sample 2, and tested with sub-sample 1, so that the full statistics of data and MC are used. The first BDT is applied to sub-sample 2 of the data and the second BDT is applied to sub-sample 1, so that the BDTs are not applied to the same events they were trained on, reducing any possibility of biasing the BDT selection. This process is summarised schematically in Figure 5.16.

The variables used to train the BDT are described bellow.

- The pseudorapidity, η, of the B^0/Λ_b^0, $\eta'/\eta/\pi^0$, charged particles and LLK_S^0/Λ^0 hadrons.
- Transverse momentum, p_T of the B^0/Λ_b^0, $\eta/\eta'/\pi^0$, photon and the sum of the charged particle p_T.
- The log of the lifetime significance of the B^0/Λ_b^0. This is the lifetime calculated for the DecayTreeFitter tool divided by its error.
- The DIRA of the B^0/Λ_b^0.

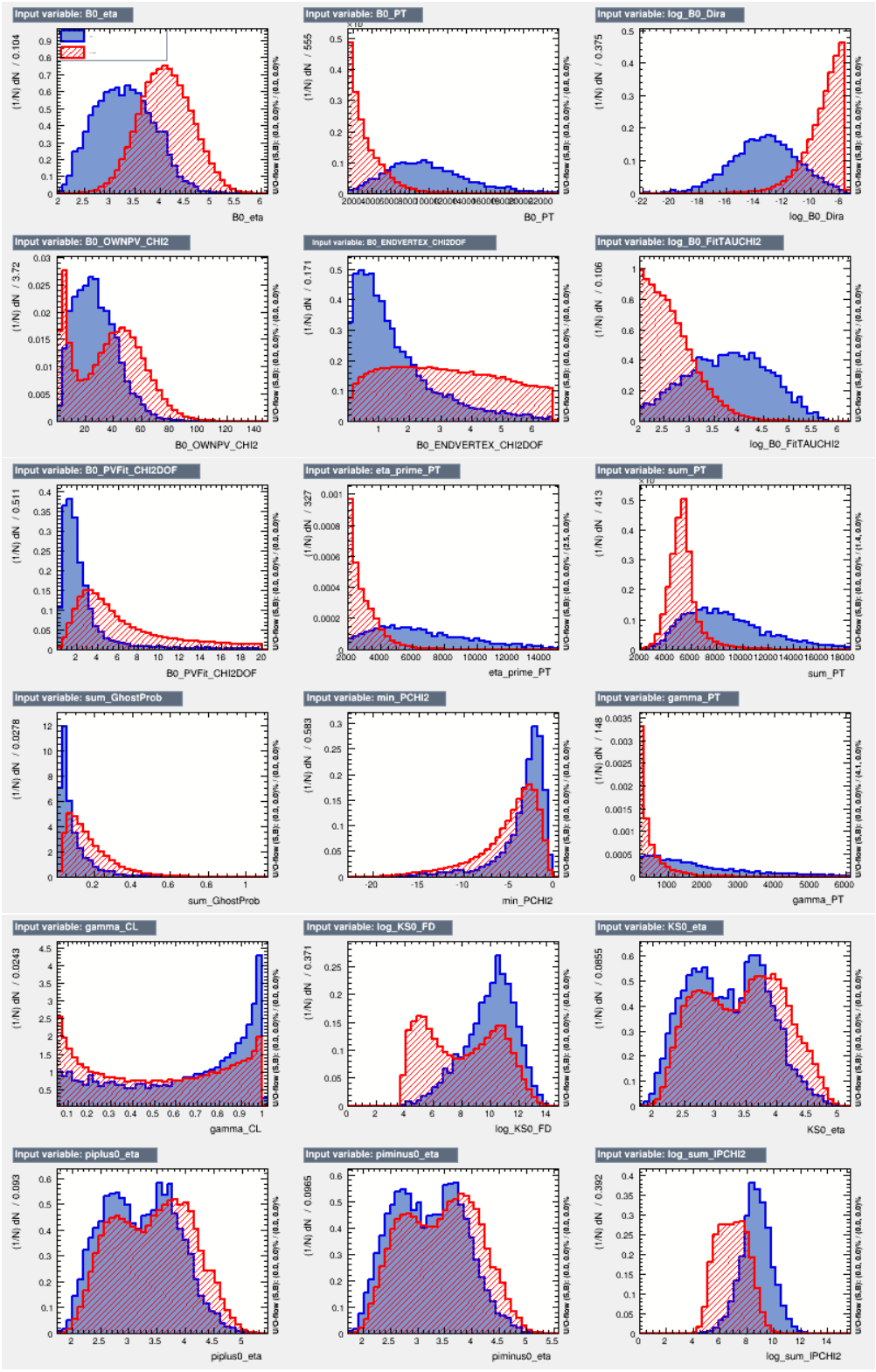

Figure 5.17: Output from TMVA comparing signal (blue) and background (red) for each variable used in the BDT for $B^0 \to K_S^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays in the LL category using 2012 data.

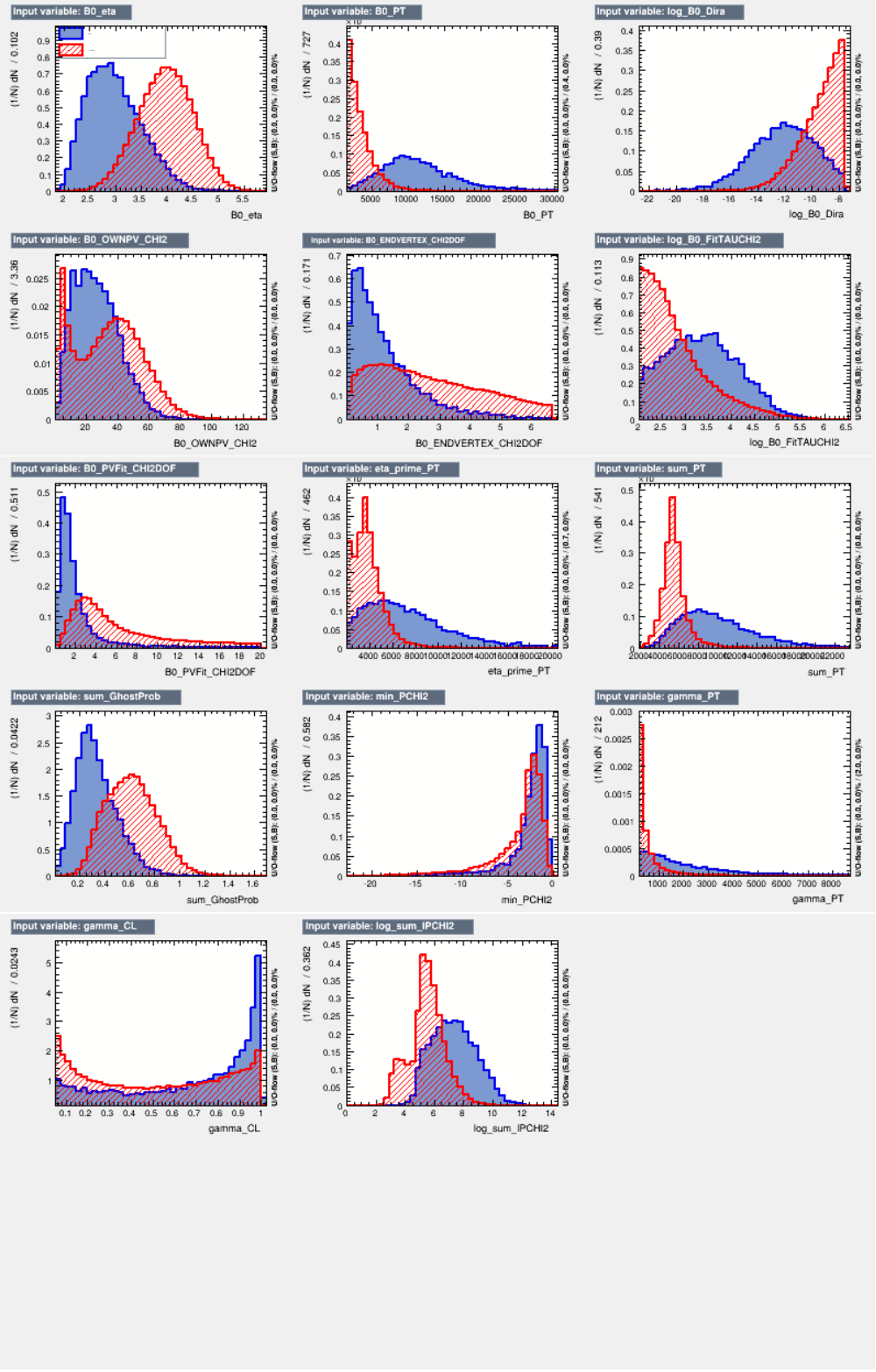

Figure 5.18: Output from TMVA comparing signal (blue) and background (red) for each variable used in the BDT for $B^0 \to K^0_S \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays in the DD category using 2012 data.

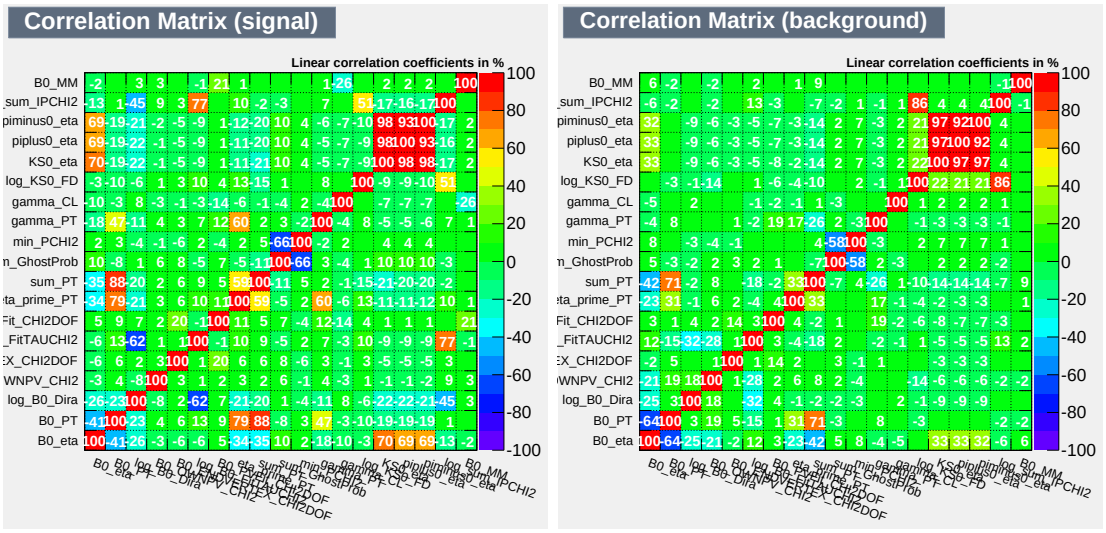

Figure 5.19: Correlation between varaibles for signal (left) and background (right), for $B^0 \to K_S^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays using 2012 data in the LL category.

- The χ^2/ndf of the decay vertex of the B^0/Λ_b^0.
- The χ^2 of the primary vertex.
- The χ^2/ndf of the fit of the full decay chain.
- The minimum probability χ^2 of the charged particles.
- The sum of the Ghost Probability of the charged particles.
- The sum of the impact parameter significance of all daughter particles for LL events. For DD events the sum does not include the K_S^0 or pions from K_S^0/Λ^0 decays.
- The flight distance significance of LL K_S^0/Λ^0 particles.

There are 18 variables used to train the LL BDT, and 14 for the DD BDT. Figure 5.17 and Figure 5.18 show the distribution of these variables for the LL and DD categories respectively, comparing the signal and background distributions. The correlations between the variables is shown in Figure 5.19. For interest, the mass of the B^0/Λ_b^0 candidate is included and no variable is highly correlated with this mass.

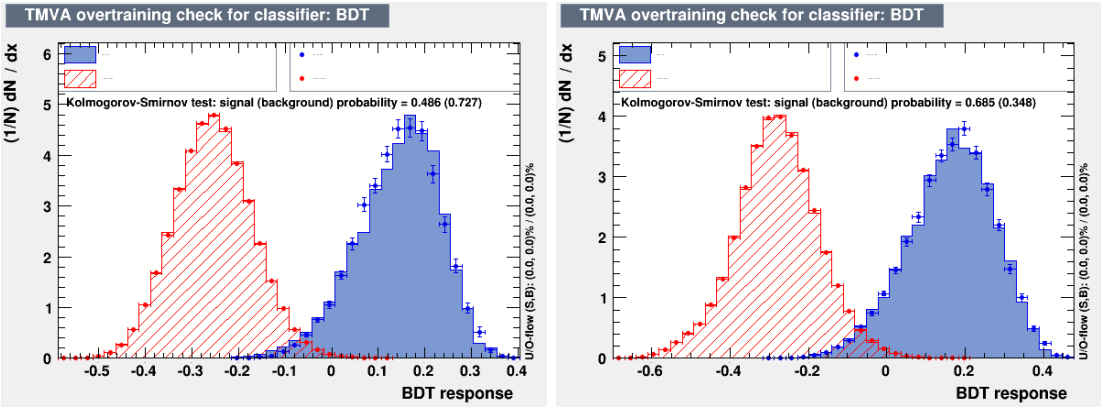

Figure 5.20: BDT_response for signal (blue) and background (red), for $B^0 \to K^0_S\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays using 2012 data in the LL (left) and DD (right) categories.

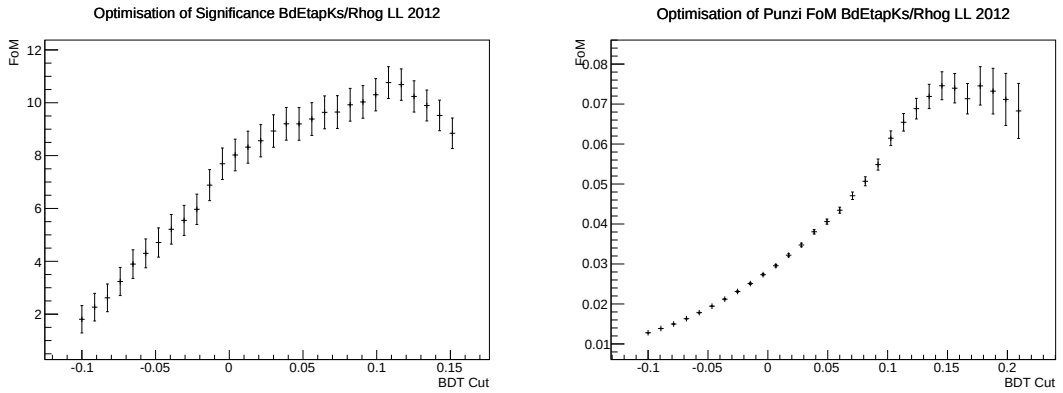

Figure 5.21: Significance (left) and Punzi figure of merit (right) as a function of the cut on the BDT_response for $B^0 \to K^0_S\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays using 2012 data in the LL category.

The BDT_response is shown in Figure 5.20 for the BDT trained to the $B^0 \to K^0_S\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays using 2012 data. The distributions show good separation between signal and background, and the test and training samples are superimposed to show there is no overtraining of the BDT.

In order to find the optimum cut on the BDT_response, two methods of optimisation are used. The first is by maximising the significance of the signal observed in data, defined as $S/\sqrt{S+B}$, where S is the number of signal events observed, and B is the number of background events in the signal mass window. Both quantities are obtained from fits to the data, described in Section 6.2.2. The second method is

Table 5.9: Optimum cut on the BDT_response.

Decay	Category	Year	BDT Cut
$B^0 \to K_S^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	2011	0.07
		2012	0.11
	DD	2011	0.09
		2012	0.10
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	2011	0.15
		2012	0.15
	DD	2011	0.15
		2012	0.15
$\Lambda_b^0 \to \Lambda^0 \eta'$ $(\eta' \to \pi^+\pi^-\eta)$	LL	2011	0.04
		2012	0.09
		2012	0.02
$\Lambda_b^0 \to \Lambda^0 \eta$ $(\eta \to \pi^+\pi^-\pi^0)$	LL	2011	0.09
		2012	0.16
		2012	0.09

to maximise the Punzi figure of merit [107], defined as $\epsilon_{\text{BDT|PID}}/(a/2 + \sqrt{B})$, where $\epsilon_{\text{BDT|PID}}$ is the efficiency of a given BDT cut with respect to the PID cut, calculated from MC samples, B is the expected number of background events in the signal window which pass the BDT cut, calculated by fitting the sidebands of the data sample and extrapolating into the signal region, and a is the significance we are optimising for, in this case $a = 3$. These two quantities are shown as a function of the cut on the BDT_response in Figure 5.21 for $B^0 \to K_S^0 \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ decays using 2012 data in the LL category. Using the significance of the signal gives a slightly looser optimum cut than using the Punzi figure of merit. In order to retain high statistics, the BDT cut will be optimised using the significance of the signal for the control channel, and the Punzi figure of merit will be used for the Λ_b^0 signal channels, where the significance is not known. The optimum BDT cuts are shown in Table 5.9

5.5.5 Multiple candidates

The final stage of the selection is to remove events that have more than one candidate in them, known as multiple candidates. Multiple candidates are a problem in this analysis due to the soft p_{T} spectrum of the photons and the wide η' mass window. This means that random photons produced in the underlying event, which generally have low p_{T}, could have a similar p_{T} to the photons from genuine η' decays, and so two or more η' candidates can be reconstructed from each event. In the 2012 data sample of $\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays, 1096 candidates pass the full selection chain with no restriction on the reconstructed B^0 mass. Of these, there were 81 events with multiple candidates. These events were spread across the whole $\mathrm{K}^0_{\mathrm{S}}\eta'$ invariant mass region, as shown in Figure 5.22. Different methods of choosing the best candidate were investigated. By choosing the candidate with the highest p_{T} photon, the photon which gave an invariant mass closest to the η' mass, or the candidate with the best fit χ^2 from the DecayTreeFitter tool, it was found that there were 211–213 signal events in the data sample, extracted from the signal component of the fit to the data, described in Section 6.2. However, using these variables may introduce biases to the background shape or the measurement of the branching fractions. To avoid this possibility entirely, a candidate is chosen at random from those which pass the selection. Following this procedure for the $\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) sample in 2012 data, 204 signal events pass the selection.

The reduction in signal efficiency by randomly choosing the candidate to keep is around 4%, which is larger than the expected size of the systematic uncertainty introduced by using a more efficient method. For this reason, the candidate which has the photon with the highest p_{T} was kept. The efficiency of selecting the correct candidate can be calculated using a sample of MC in which the underlying event is simulated, so that multiple candidates are modelled. The efficiency is given by:

$$\epsilon_{\mathrm{m}} = 1 - f \times (1 - \epsilon_{\mathrm{c}}), \tag{5.20}$$

where ϵ_{m} is the efficiency of applying the multiple candidates selection to signal

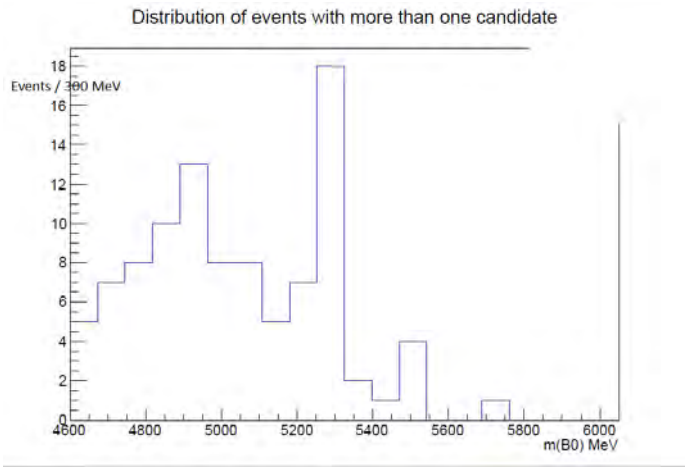

Figure 5.22: Invariant mass distribution of events with more than one candidate in the 2012 data sample of $B^0 \to K_S^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays.

events, f is the fraction of signal events which have more than one candidates, and ϵ_c is the probability of selecting the correct signal candidate from an event which contains more than one candidate. These numbers can be obtained from the MC samples, and are shown in Table 5.10. There will be a systematic uncertainty associated with this, due to the modelling of the underlying event in the MC, and this is discussed in Section 6.5.

Finally, a study is performed to ensure this method of rejecting candidates does not bias the distribution of background events. In this, the invariant mass distribution of all events with exactly one candidate is fit, with the background being modelled with an exponential shape. From this fit, we find that the background is described by an exponential shape of the form $e^{-\tau t}$, with a characteristic decay constant of $\tau = 0.0033 \pm 0.0002\,\mathrm{MeV}/c^{2^{-1}}$. The fit is then performed to the invariant mass distribution of all events which are selected by the multiple candidate selection. In this case the background shape is found to have a characteristic decay constant of $\tau = 0.0032 \pm 0.0002\,\mathrm{MeV}/c^{2^{-1}}$. If the multiple candidates are being removed from a particular mass region, then the shape of the background will change after they are removed. However, as expected the parameters are consistent, and this shows that the method of choosing the best candidate does not bias the background shape.

Table 5.10: Efficiencies for applying multiple candidates cuts on signal events.

Decay	**Year**	**Cat**	f	ϵ_{c}	ϵ_{m}
$\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{S}}\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	2011	LL	0.18	0.81	0.97
	2012	LL	0.075	0.87	0.99
	2011	DD	0.28	0.83	0.95
	2012	DD	0.27	0.86	0.96
$\Lambda^0_{\mathrm{b}} \to \Lambda^0\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	2011	LL	0.16	0.87	0.98
	2012	LL	0.15	0.87	0.98
	2011	DD	0.18	0.86	0.98
	2012	DD	0.21	0.86	0.97
$\Lambda^0_{\mathrm{b}} \to \Lambda^0\eta'$ $(\eta' \to \pi^+\pi^-\eta)$	2011	LL	0.24	0.71	0.93
	2012	LL	0.26	0.74	0.93
	2012	DD	0.17	0.75	0.95
$\Lambda^0_{\mathrm{b}} \to \Lambda^0\eta$ $(\eta \to \pi^+\pi^-\pi^0)$	2011	LL	0.22	0.75	0.95
	2012	LL	0.24	0.70	0.93
	2011	DD	0.21	0.75	0.95
	2012	DD	0.21	0.73	0.94

5.6 Summary

This chapter described the analysis strategy used in the search for the $\Lambda^0_{\mathrm{b}} \to \Lambda^0\eta^{(\prime)}$ decays. The analysis is performed blind to avoid any bias in the measured results, and the $\mathrm{B}^0 \to \mathrm{K}^0_{\mathrm{s}}\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ decay is used as a control channel to optimise the selection and measure the relative branching fractions of the signal channels. The search is performed using the full dataset recorded by the LHCb experiment in 2011 and 2012 corresponding to $3.0\,\mathrm{fb}^{-1}$ of pp collisions. A selection is applied to the data to reduce the background to a level which makes it possible to observe the signal events. The selection consists of a loose kinematic pre-selection, a PID selection to remove misidentified tracks, and a multivariate selection to further separate signal from combinatorial background. Each stage of the selection is described in detail in this chapter. The selection is applied to both MC signal samples and data to test the performance of the selection and to search for the signal decays, and the results of this are presented in Chapter 6.

CHAPTER

6

RESULTS OF SEARCH FOR $\Lambda_B^0 \to \Lambda^0 \eta^{(\prime)}$

This chapter presents the results in the search for the $\Lambda_b^0 \to \Lambda^0 \eta$ and $\Lambda_b^0 \to \Lambda^0 \eta'$ decays. The B^0 selection described in Chapter 5 is applied to the simulated MC samples to measure the efficiency of the selection and to describe the fits to the invariant mass distribution. The selection is then applied to the data sample and the yield of the control channel is found. Cross checks are performed on the control channel and physics backgrounds are investigated. The Λ_b^0 selections are applied and the signal yields are extracted and used to determine the branching fractions of the signal channels. The statistical treatment of these results are described and the values are discussed in relation to the theoretical predictions shown in Chapter 2.

6.1 Efficiencies

The efficiencies of each stage of the selection are calculated using the MC samples by measuring the fraction of events in the sample which pass the requirements. The measured efficiencies are presented in Table 6.1, which shows the combined stripping and reconstruction efficiency, ϵ_{STRIP}. The stripping efficiency is measured using truth-matched MC samples, with different samples used for the 2011 and 2012 conditions, and reweighting the sample by the PID variables as described in Section 5.5.3. In this analysis, the MC samples are truth-matched, that is they have cuts applied to remove any events in the sample which are not true $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\eta'$ or $\Lambda^0_\mathrm{b} \to \Lambda^0\eta^{(\prime)}$ decays. This is done using the information provided by PYTHIA which matches the particle that was generated to the reconstructed particle in the sample. The efficiency is found by applying the stripping selection to the MC, counting the number of events which pass the selection, and dividing by the total number of events generated.

Also shown is the efficiency of the PID cuts with respect to the pre-selection, $\epsilon_{\text{PID|STRIP}}$, and the efficiency of the BDT selection with respect to the PID selection, $\epsilon_{\text{BDT|PID}}$, measured in the same way as the stripping efficiency. Table 6.2 shows the efficiency of the trigger selection with respect to the BDT selection, $\epsilon_{\text{TRIG|BDT}}$, broken down into the L0, HLT1 and HLT2 efficiencies. For the 2012 data, 2012a refers to the trigger efficiencies before the June technical stop, and 2012b refers to trigger efficiencies after the June technical stop. It is known that the MC samples used to calculate the 2012a trigger efficiencies contain a bug, which causes the L0 efficiency to be calculated incorrectly. For example, for the $\mathrm{B}^0 \to \mathrm{K}^0_\mathrm{S}\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, the measured L0 efficiency is 41% for the 2012a sample, whereas there is no reason this should be different from that of the 49% in the 2012b sample. To compensate for this bug, the 2012a L0 efficiency is assumed to be the same as the 2012b efficiency, and this is shown in Table 6.2. A 5% systematic uncertainty is assigned to this assumption. The 2012a data sample corresponds to $\approx 17\%$ of the overall data sample, and so the overall systematic uncertainty is $< 1\%$. Since the HLT efficiencies

Table 6.1: Efficiencies of the pre-selection, PID and trigger cuts.

Decay	**Cat**	**Year**	ϵ_{STRIP} %	$\epsilon_{\text{PID\|STRIP}}$ %	$\epsilon_{\text{BDT\|PID}}$ %
	LL	2011	0.314 ± 0.004	89.9 ± 0.2	67.6 ± 0.3
$B^0 \to K_S^0 \eta'$	LL	2012	0.260 ± 0.002	88.5 ± 0.1	47.5 ± 0.2
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	0.769 ± 0.006	91.9 ± 0.1	53.3 ± 0.2
	DD	2012	0.688 ± 0.004	91.62 ± 0.09	52.0 ± 0.2
	LL	2011	0.157 ± 0.003	85.1 ± 0.2	56.7 ± 0.3
$\Lambda_b^0 \to \Lambda^0 \eta'$	LL	2012	0.128 ± 0.002	86.5 ± 0.2	40.6 ± 0.3
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	0.377 ± 0.004	89.1 ± 0.1	52.8 ± 0.2
	DD	2012	0.339 ± 0.003	89.8 ± 0.1	47.8 ± 0.2
	LL	2011	0.080 ± 0.002	86.3 ± 0.3	90.5 ± 0.3
$\Lambda_b^0 \to \Lambda^0 \eta'$	LL	2012	0.065 ± 0.001	88.5 ± 0.3	75.4 ± 0.4
$(\eta' \to \pi^+\pi^-\eta\ )$	DD	2011	0.174 ± 0.003	90.3 ± 0.2	N/A
	DD	2012	0.155 ± 0.002	91.2 ± 0.2	84.3 ± 0.2
	LL	2011	0.083 ± 0.002	86.6 ± 0.2	88.4 ± 0.2
$\Lambda_b^0 \to \Lambda^0 \eta$	LL	2012	0.063 ± 0.001	87.3 ± 0.2	68.1 ± 0.3
$(\eta \to \pi^+\pi^-\pi^0\ )$	DD	2011	0.220 ± 0.003	89.63 ± 0.10	N/A
	DD	2012	0.193 ± 0.002	90.5 ± 0.1	82.1 ± 0.2

are computed correctly, these are measured using the 2012a MC samples. Table 6.3 shows the total efficiency of the reconstruction and selection, and the expected yield in each data set, for each decay channel, calculated using Equation 5.2.

6.2 Fits to invariant mass distributions

In order to find the signal yields in data, a model is defined which describes the expected invariant mass distribution of the B^0 or Λ_b^0 candidates. This model is then fitted to the data to extract the parameters of the model and the signal yields. This section describes how the model is constructed before describing the results of the fit to the MC and data samples.

The model used to describe the invariant mass distribution in data contains two components: a signal component and a background component. For the signal component, the invariant mass is expected to follow a Gaussian probability function

Table 6.2: Efficiencies of the trigger cuts, separated into different trigger conditions for 2012.

Decay	**Cat**	**Year**	$\epsilon_{\text{L0\|BDT}}$ %	$\epsilon_{\text{HLT1\|L0}}$ %	$\epsilon_{\text{HLT2\|HLT1}}$ %	$\epsilon_{\text{TRIG\|BDT}}$ %
	LL	2011	49.3 ± 0.4	65.9 ± 0.5	66.1 ± 0.7	21.5 ± 0.3
	LL	2012a	49 ± 2	68.5 ± 0.6	88.6 ± 0.5	30 ± 2
$B^0 \to K^0_S \eta'$		2012b	49.1 ± 0.3	71.1 ± 0.4	65.7 ± 0.5	22.9 ± 0.3
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	33.1 ± 0.3	58.0 ± 0.5	31.3 ± 0.6	6.0 ± 0.1
	DD	2012a	36 ± 2	53.6 ± 0.4	32.1 ± 0.5	6.2 ± 0.3
		2012b	36.2 ± 0.2	57.9 ± 0.4	55.0 ± 0.5	11.5 ± 0.1
	LL	2011	54.8 ± 0.4	68.2 ± 0.5	55.2 ± 0.7	20.6 ± 0.3
	LL	2012a	56 ± 3	68.7 ± 0.8	84.7 ± 0.7	32 ± 2
$\Lambda^0_b \to \Lambda^0 \eta'$		2012b	55.7 ± 0.5	72.2 ± 0.6	69.5 ± 0.7	28.0 ± 0.4
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	57.8 ± 0.3	57.6 ± 0.4	26.4 ± 0.4	8.8 ± 0.2
	DD	2012a	52 ± 3	51.7 ± 0.6	35.5 ± 0.8	9.4 ± 0.5
		2012b	51.5 ± 0.3	55.6 ± 0.4	60.4 ± 0.5	17.3 ± 0.2
	LL	2011	46.4 ± 0.5	41.4 ± 0.7	33 ± 1	6.3 ± 0.2
	LL	2012a	42 ± 2	43 ± 1	63 ± 2	11.5 ± 0.7
$\Lambda^0_b \to \Lambda^0 \eta'$		2012b	42.5 ± 0.5	51.2 ± 0.8	46 ± 1	10.0 ± 0.3
$(\eta' \to \pi^+\pi^-\eta\ )$	DD	2011	53.1 ± 0.4	30.4 ± 0.5	4.2 ± 0.4	0.67 ± 0.07
	DD	2012a	46 ± 2	19.8 ± 0.5	9.7 ± 0.9	0.89 ± 0.08
		2012b	46.3 ± 0.4	25.2 ± 0.5	39 ± 1	4.6 ± 0.1
	LL	2011	46.8 ± 0.3	56.2 ± 0.4	43.6 ± 0.5	11.5 ± 0.2
	LL	2012a	45 ± 2	57.1 ± 0.6	73.8 ± 0.6	19.2 ± 1.0
$\Lambda^0_b \to \Lambda^0 \eta$		2012b	45.5 ± 0.3	64.4 ± 0.5	58.3 ± 0.6	17.1 ± 0.2
$(\eta \to \pi^+\pi^-\pi^0\ )$	DD	2011	46.2 ± 0.2	39.8 ± 0.2	15.5 ± 0.3	2.85 ± 0.06
	DD	2012a	43 ± 2	39.7 ± 0.3	20.2 ± 0.4	3.5 ± 0.2
		2012b	43.3 ± 0.2	44.0 ± 0.3	46.3 ± 0.5	8.8 ± 0.1

Table 6.3: Total efficiencies and expected yields for each decay channel.

Decay	**Cat**	**Year**	ϵ_{tot} %	**Expected Yield**
	LL	2011	0.041 ± 0.001	140 ± 49
$B^0 \to K_S^0 \eta'$	LL	2012	0.0270 ± 0.0008	193 ± 67
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	0.023 ± 0.001	77 ± 27
	DD	2012	0.0335 ± 0.0009	240 ± 84
	LL	2011	0.0156 ± 0.0009	1.2 ± 0.1
$\Lambda_b^0 \to \Lambda^0 \eta'$	LL	2012	0.0131 ± 0.0006	2.0 ± 0.2
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	0.0156 ± 0.0009	1.2 ± 0.1
	DD	2012	0.0223 ± 0.0007	3.5 ± 0.3
	LL	2011	0.0039 ± 0.0004	0.17 ± 0.02
$\Lambda_b^0 \to \Lambda^0 \eta'$	LL	2012	0.0045 ± 0.0003	0.41 ± 0.04
$(\eta' \to \pi^+\pi^-\eta\ )$	DD	2011	0.0007 ± 0.0002	0.031 ± 0.008
	DD	2012	0.0044 ± 0.0003	0.40 ± 0.04
	LL	2011	0.0073 ± 0.0006	0.42 ± 0.05
$\Lambda_b^0 \to \Lambda^0 \eta$	LL	2012	0.0066 ± 0.0004	0.80 ± 0.08
$(\eta \to \pi^+\pi^-\pi^0\ )$	DD	2011	0.0052 ± 0.0005	0.30 ± 0.04
	DD	2012	0.0107 ± 0.0005	1.3 ± 0.1

of the form

$$Gauss(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{6.1}$$

where μ is the mean and σ is the resolution of the Gaussian. This function is normalised so that the integral over all values of x is one. When fitting the signal invariant mass distribution, the mean, μ of the Gaussian will be the mass of the B^0 or Λ_b^0 candidate and σ will be defined as the resolution of the invariant mass. In this analysis the resolution of the invariant mass will include contributions from two detector effects: the resolution of the momentum measurement of the tracks and the resolution of the energy measurement of the photons. The signal can therefore not be described by a single Gaussian function. Instead a double Gaussian function is used, which is defined as the sum of two Gaussian functions with the same mean but independent resolutions. For the background component, the combinatoric background is described by an exponential function of the form

$$Bkg(x) = N_0 \cdot exp\left(-\tau x\right), \tag{6.2}$$

where τ is the characteristic strength of the exponential and N_0 is the normalisation.

These signal and background models are added together to form a combined probability density function (PDF) to describe the invariant mass spectrum. The PDF is defined as

$$PDF(x) = [f_s \cdot Sig(x) + (1 - f_s) \cdot Bkg(x)], \tag{6.3}$$

where f_s is the fraction of events in the signal model.

The PDF is fitted to the spectrum by performing an unbinned extended maximum likelihood fit. The fit is unbinned, and so each event with invariant mass x_i in the dataset of N events is compared with the PDF with a given set of parameters, $\mathbf{p}$. A likelihood is built up from the product of the probabilities of each point, such that the likelihood is

$$\mathcal{L} = \prod_{i=1}^{N} [PDF(x_i; \mathbf{p})]. \tag{6.4}$$

The parameters of the fit are then varied along with f_s to maximise this likelihood. In practice, it is computationally more efficient to do this by minimising the negative log-likelihood, defined as

$$NLL = -\sum_{i=1}^{N} \ln[PDF(x_i; \mathbf{p})]. \tag{6.5}$$

In order to extract yields from the fit, we extend the PDF to include the number of events as another free parameter following a Poisson probability. The likelihood then becomes

$$\mathcal{L} = \frac{\nu^n}{n!} e^{-\nu} \prod_{i=1}^{N} [PDF(x_i; \mathbf{p})] \tag{6.6}$$

where ν is the number of signal events and n is the total number of events.

The unbinned extended maximum likelihood fit is performed to both MC samples and the data samples, and the results of these fits are presented in the following sections.

When studying the invariant mass distribution of intermediate states, rather than

the overall B^0 or Λ_b^0 mass distribution, it is found that a special treatment is required for the signal η' decays, where the shape is slightly different due to the presence of radiative photons which are not included in the invariant mass calculation. This has the effect of extending the Gaussian function to include a larger tail at low values. There is also a tail at the high invariant mass, since these tend to have high photon p_T and so are reconstructed with a high efficiency. To account for this a Crystal Ball function [108] is used to model these decays, which consists of a core Gaussian function with an exponential tail, defined as

$$CB(x) = N \cdot \left\{ \begin{array}{ll} exp(-\frac{(x-\mu)^2}{2\sigma^2}), & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\mu}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} \leq -\alpha, \end{array} \right| \tag{6.7}$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot exp\left(-\frac{|\alpha|^2}{2}\right) \tag{6.8}$$

$$B = \frac{n}{|\alpha|} - |\alpha| \tag{6.9}$$

and both functions are normalised.

6.2.1 Mass fits to MC

The selection is applied to the signal and control channel MC samples, and the invariant mass distribution of the B^0 and Λ_b^0 candidates is plotted. An unbinned extended maximum likelihood fit is performed with the signal model alone to determine the resolution obtained with this selection. The reconstructed B^0, Λ_b^0, K_S^0 and Λ^0 invariant mass distributions are fitted with a double Gaussian function and the η' invariant mass is fitted with a sum of two Crystal Ball functions with the same mean and same width, but with different tail parameters. The results of these fits are shown in Figure 6.1 and 6.2 for the $B^0 \to K_S^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, and Figure 6.3 and 6.4 for the $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, in the LL and DD categories, respectively. The fits to the intermediate resonances, the η', K_S^0 and Λ^0

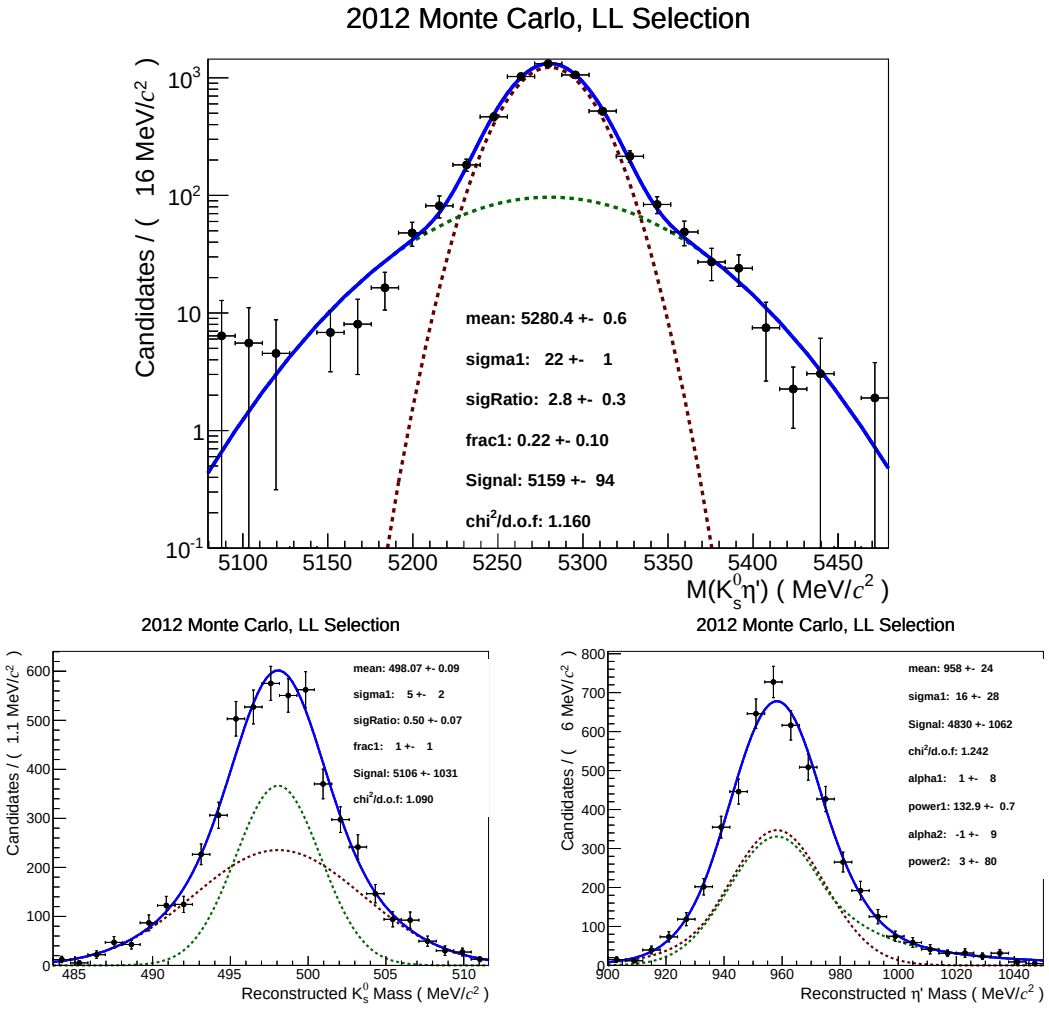

Figure 6.1: Fits to the invariant mass distribution of the reconstructed B^0, K^0_S and η' using MC samples of $B^0 \to K^0_S\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays in 2012 in the LL category. The solid blue line shows the overall fit and the dashed lines show each component of the fit.

particles, are used to define the mass window cuts applied to data, and as a cross check to compare with the resonance observed in data. The parameters of the B^0 mass fit are used in the fits to data, where the ratio of the width of the two Gaussian functions, σ_1/σ_2, and the ratio of the number of events in each Gaussian function, N_1/N_2, are fixed to the values found in MC when performing fits to the B^0 and Λ^0_b mass distributions in data. The fit is performed to MC samples of each of the decay channels, and the parameters obtained from the fit to the B^0 and Λ^0_b invariant mass are given in Table 6.4.

Upon further investigation of the control channel, it was found that there were some

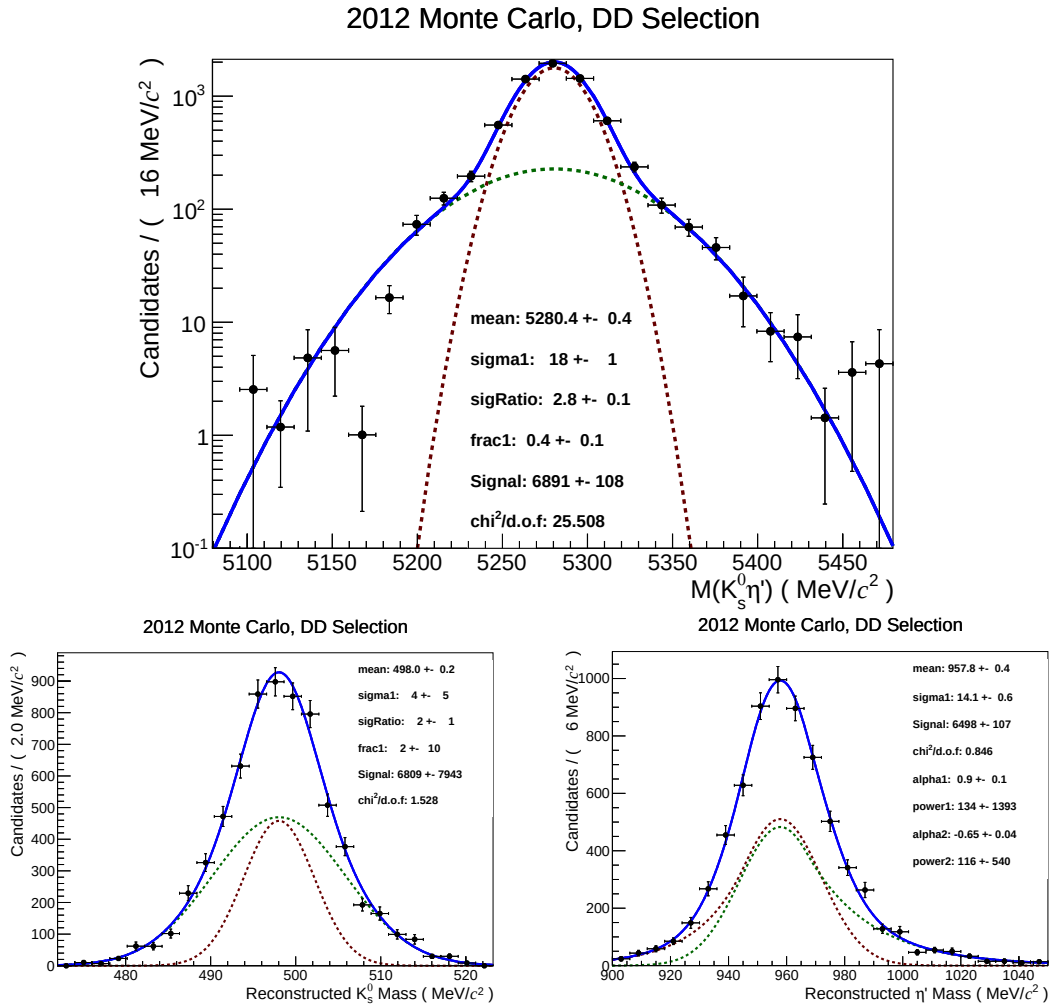

Figure 6.2: Fits to the invariant mass distribution of the reconstructed B^0, K^0_S and η' using MC samples of $B^0 \to K^0_S\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays in 2012 in the DD category. The solid blue line shows the overall fit and the dashed lines show each component of the fit.

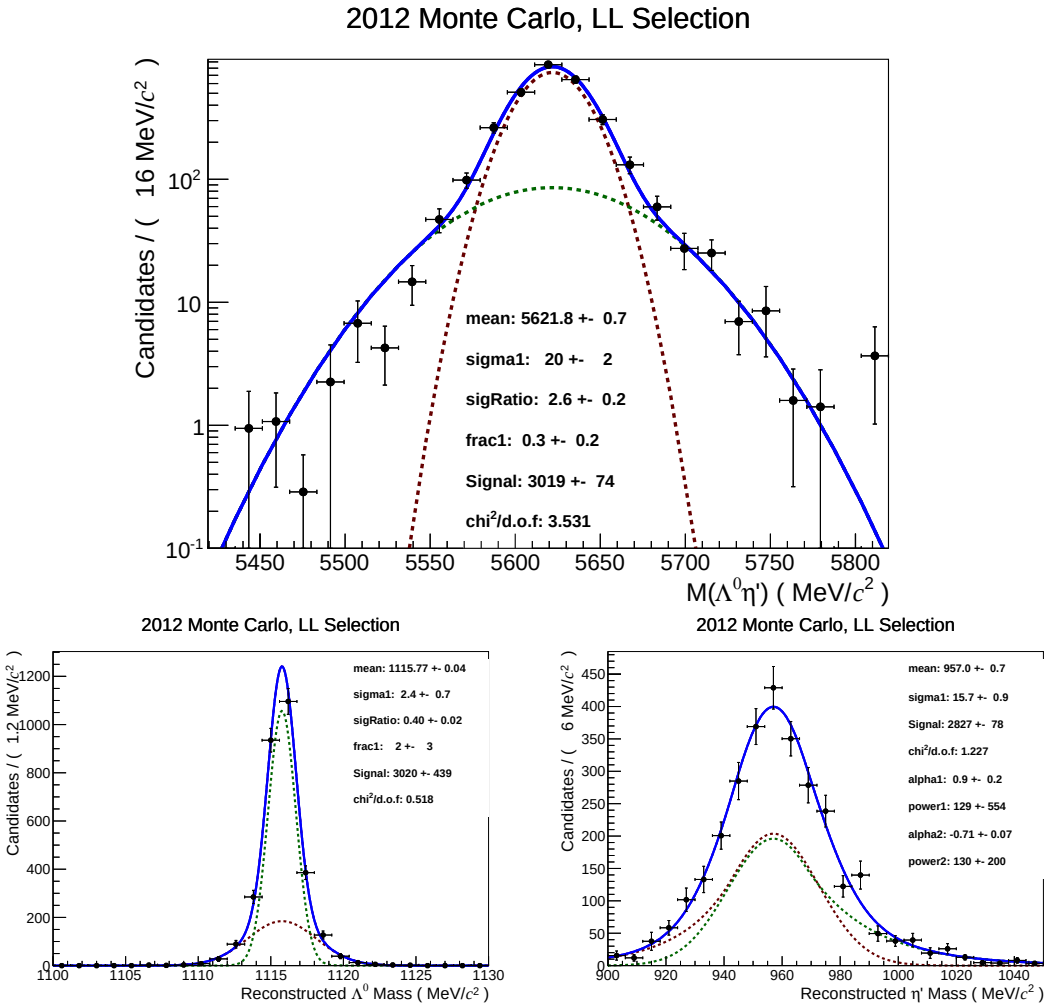

Figure 6.3: Fits to the invariant mass distribution of the reconstructed Λ_b^0, Λ^0 and η' using MC samples of $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays in 2012 in the LL category. The solid blue line shows the overall fit and the dashed lines show each component of the fit.

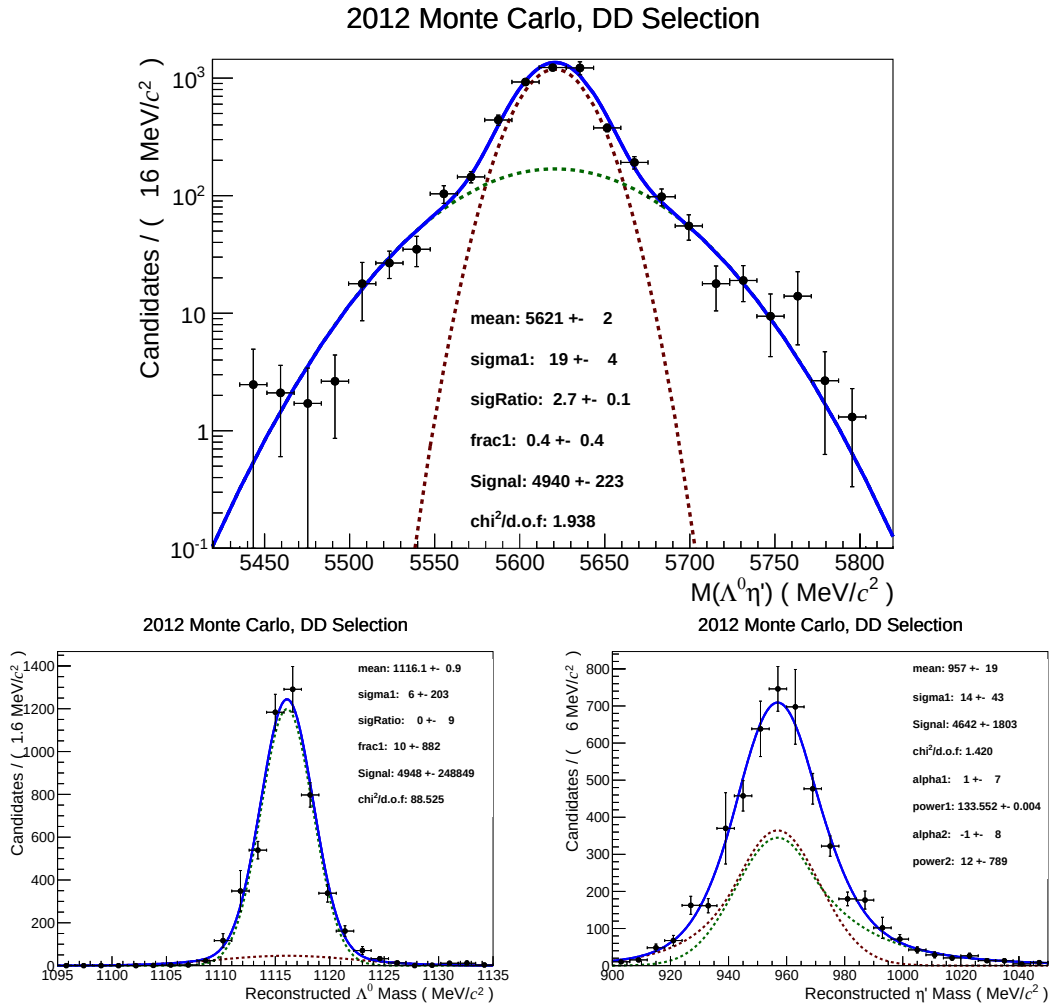

Figure 6.4: Fits to the invariant mass distribution of the reconstructed Λ_b^0, Λ^0 and η' using MC samples of $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays in 2012 in the DD category. The solid blue line shows the overall fit and the dashed lines show each component of the fit.

Table 6.4: Parameters obtained from an unbinned maximum likelihood fit to the B^0 and Λ_b^0 invariant mass distribution in MC.

Decay	Category	Year	$\sigma_1(\,\mathrm{MeV}/c^2)$	σ_2/σ_1	N_2/N_1
	LL	2011	19.348 ± 1.199	2.882 ± 0.207	0.246 ± 0.097
	LL	2012	22.0 ± 1.4	2.8 ± 0.3	0.2 ± 0.1
$B^0 \to K_S^0\eta'$	LL	11+12	21.1 ± 1.0	2.8 ± 0.2	0.2 ± 0.1
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	18.8 ± 1.2	3.1 ± 0.2	0.3 ± 0.1
	DD	2012	18.3 ± 1.0	2.8 ± 0.1	0.4 ± 0.1
	DD	11+12	18.4 ± 0.8	2.9 ± 0.1	0.3 ± 0.1
	LL	2011	22.0 ± 1.6	3.6 ± 1.11	0.1 ± 0.08
	LL	2012	20.0 ± 2.3	2.6 ± 0.2	0.3 ± 0.2
$\Lambda_b^0 \to \Lambda^0\eta'$	LL	11+12	21.4 ± 1.3	3.0 ± 0.3	0.18 ± 0.08
$(\eta' \to \pi^+\pi^-\gamma\ )$	DD	2011	20.9 ± 4.5	2.5 ± 0.2	0.3 ± 0.3
	DD	2012	19.1 ± 4.3	2.7 ± 0.1	0.4 ± 0.4
	DD	11+12	19.6 ± 2.8	2.6 ± 0.1	0.3 ± 0.2
	LL	2011	30.8 ± 63.1	2.1 ± 1.1	0.7 ± 6.9
$\Lambda_b^0 \to \Lambda^0\eta'$	LL	2012	40.6 ± 7.8	2.4 ± 3.4	0.1 ± 0.4
$(\eta' \to \pi^+\pi^-\eta\ )$	LL	11+12	37.5 ± 19.2	2.0 ± 1.8	0.2 ± 1.5
	DD	2012	31.3 ± 23.7	2.2 ± 0.3	0.7 ± 2.6
	LL	2011	29.8 ± 2.6	2.8 ± 0.3	0.3 ± 0.2
$\Lambda_b^0 \to \Lambda^0\eta$	LL	2012	30.4 ± 3.4	2.8 ± 0.3	0.4 ± 0.2
$(\eta \to \pi^+\pi^-\pi^0\ )$	LL	11+12	30.2 ± 2.0	2.8 ± 0.2	0.4 ± 0.1
	DD	2012	28.4 ± 5.5	2.7 ± 0.3	0.4 ± 0.4

inconsistencies in the parameters obtained from the fit, and that they depend on the MC simulation used. Investigations were carried out into the shape parameters obtained using two different MC samples, produced with 2012 conditions, with one sample simulating the η' decay into the non-resonant $\pi^+\pi^-\gamma$ final state using the phase space MC model, and the other sample simulating the η' decaying into the $\rho^0\gamma$ final state, with the $\rho^0 \to \pi^+\pi^-$ decay, and using the SVP MC model, which is a specific model used by EvtGen to simulate a scalar particle decaying into a pseudoscalar and a vector particle. The fits to the η' mass shape using these two samples yield consistent results, however the fits to the B^0 mass show some differences, as shown in Figure 6.5. It should be noted that the statistics in these samples are much lower than the phase space samples used, and so no selection is applied to these samples. The fits are therefore not expected to be directly comparable with the results in Table 6.4, but they can be meaningfully compared with each other. It can also be seen that the fits to the SVP samples reconstructed in the LL category are consistent in both samples. In the DD category, all the parameters are consistent except for the relative number of events in each of the Gaussian functions. It can be seen that there is a larger fraction of events in the narrower Gaussian when considering the phase space decay ($N_2/N_1 = 0.248$), than when considering the resonant decay ($N_2/N_1 = 0.070$). Since the η' decay is expected to be dominated by the resonant decay, this difference could cause problems in the fit to data. The parameters are different by a factor of 3.5, and so the number obtained in the final column of Table 6.4 will be divided by 3.5 when fitting the data in the DD category. No correction will be applied to the LL category and a systematic uncertainty will be assigned to this correction.

6.2.2 Mass fits to data for control channel

Figure 6.6 shows the reconstructed $K_S^0\eta'$ invariant mass distributions of events reconstructed as $B^0 \to K_S^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays. An unbinned maximum likelihood fit is performed to the data, using an exponential function to model the combinatoric

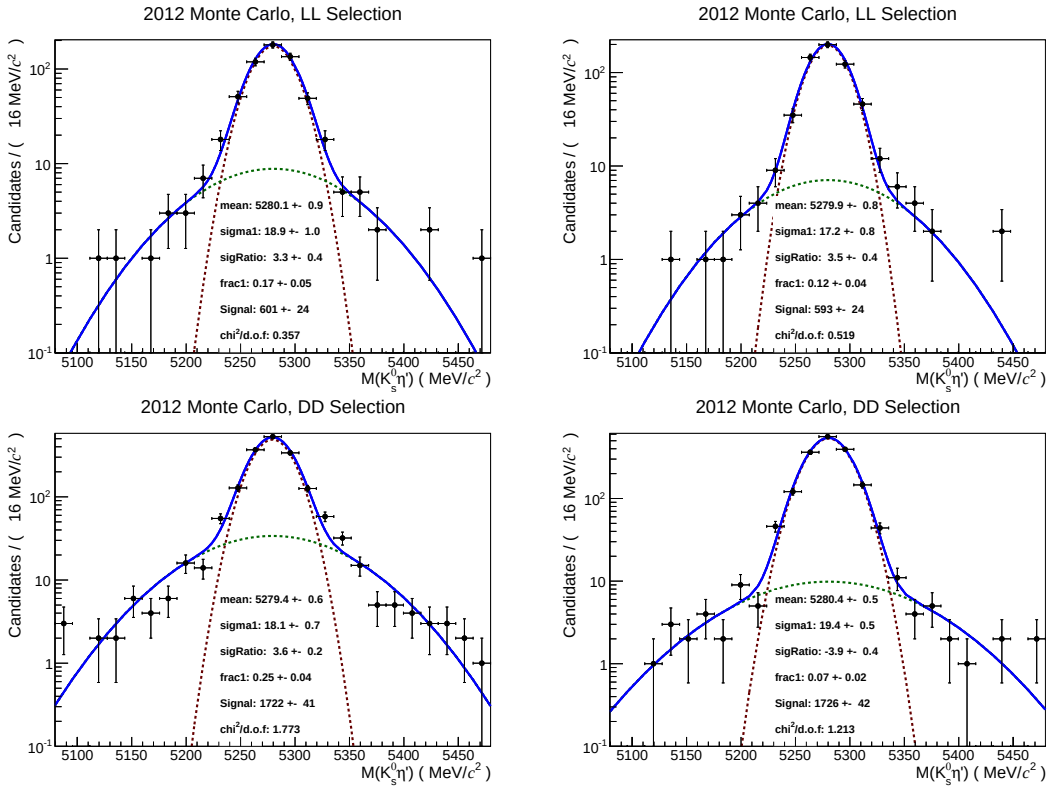

Figure 6.5: Fits to the B^0 mass using MC samples, with the η' decaying into the non-resonant $\pi^+\pi^-\gamma$ decay (left) and the $\rho^0\gamma$ decay (right), split into LL category (Top) and DD category (Bottom).

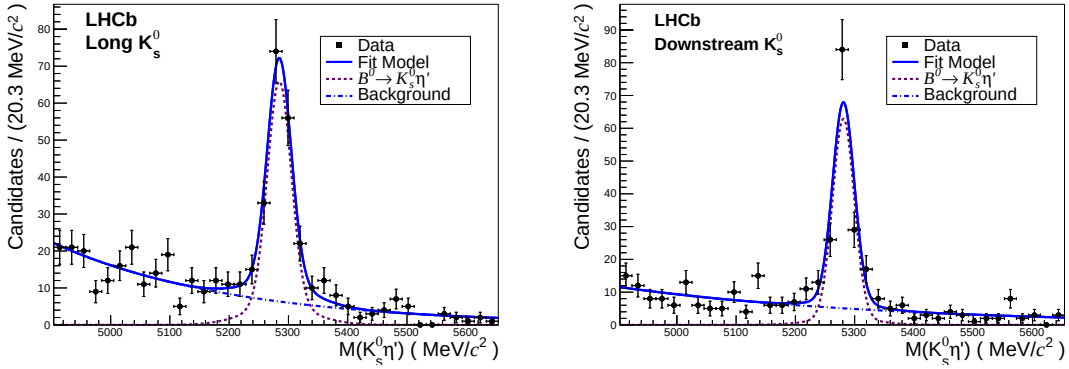

Figure 6.6: Reconstructed $K^0_S\eta'$ invariant mass from the $B^0 \to K^0_S\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, combining 2011 and 2012 data in the LL (left) and DD (right) categories.

Table 6.5: Yields obtained from the unbinned maximum likelihood fit to data.

Decay	Cat	$\sigma_1(\,\mathrm{MeV}/c^2)$	Yield, N_S	Significance
$B^0 \to K^0_S \eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	LL	20 ± 2	188 ± 16	19.2
	DD	18 ± 2	149 ± 14	18.6

background, and a sum of two Gaussian functions with a common mean to model the signal. The parameters of the Gaussian functions are fixed to the values found in fits to the Monte Carlo, as described in Section 6.2.1, and only the mean, μ, and the resolution of the narrower Gaussian function, σ_1, are free to float. The parameters of the fits are found to be consistent between the 2011 and 2012 data fits, and so the two datasets are added together and the fit is performed to the combined dataset.

The yields determined for these decays are shown in Table 6.5, along with the statistical significance and the measured σ_1. The yields are obtained from the signal fit, and defined by the number of events in the signal distribution within a mass window of $3 \times \sigma_2$ around the nominal B^0 mass, where σ_2 is the width of the broader Gaussian shape in the fit. The statistical significance is calculated by comparing the negative log likelihood from the minimised fit to the signal plus background ($NLL(S+B)$) to the negative log likelihood of the background only hypothesis ($NLL(B)$), i.e. fitting an exponential alone. The significance is then given by:

$$L = \sqrt{-2 \times \{NLL(S+B) - NLL(B)\}}, \tag{6.10}$$

The quality of the fit is assessed using the χ^2 test statistic. This takes the binned distribution and compares the value of each bin, y_i, with the expected number using the PDF, $PDF(x_i; \hat{\mathbf{p}})$. The χ^2 is defined as

$$\chi^2 = \sum_i \left(\frac{y_i - PDF(x_i; \hat{\mathbf{p}})}{\sigma(y)_i} \right)^2, \tag{6.11}$$

where the sum is over all bins and $\sigma(y)_i$ is the uncertainty on the number of events in that bin. If the fit is good, then any deviations of the observed data from the fit will

be due to statistical fluctuations only. A good fit would then have a χ^2/ndf value of around one. A value much larger than one indicates a poor quality of fit, and a value much less than one indicates that the uncertainties have been over-estimated. The fits in Figure 6.6 have χ^2 values of 0.9 and 1.1 for the LL and DD category respectively.

6.3 Background studies

Different background processes which might pass the selection and enter the invariant mass distributions have been studied. These backgrounds fall into three categories. The first category is the combinatoric background, where random tracks combine and are reconstructed in the decay. These events are modelled by an exponential shape.

The second category is the partially reconstructed background, where a particle from a similar decay is not reconstructed, and so the decay appears the same as the signal decay, but with a lower invariant mass. The main channel investigated in this category is the $B^0 \to K^{*0}\eta'$ decay, where the K^{*0} decays into $K^0_S\pi^0$ and the π^0 is not reconstructed. A MC sample of these decays has been produced, and the full selection is applied to these events. Events which pass the pre-selection are suppressed by the BDT due to the presence of the DIRA variable in the BDT. If an event has a missing particle from the decay, then the four-momentum of the B^0 candidate will be incorrectly calculated, and so will not point back to the primary vertex. The expected yield of this decay in the 2012 data is only ≈ 2 decays. Therefore, this decay is not modelled in the overall fit model.

In addition, the same decay is studied with the K^{*0} decaying to $K^+\pi^-$, and the K^+ being misidentified as a π^+. In this case, the background would only be present in the LL category, and not the DD category, as the K^{*0} is not a long lived particle. However, this background is hugely suppressed by the PID requirements on the tracks from the K^0_S, and so it is expected there will be fewer than one candidate passing

the selection in the full data sample. This background is also not modelled in the fit.

Another mode which was tested is the $B^0 \to \omega\eta'$ decay, where the ω decays into $\pi^+\pi^-\pi^0$ and the π^0 meson is not reconstructed. However, this mode has a low branching fraction of just 1×10^{-6}, and it was found to be suppressed by the BDT, so that only two events are expected in the mass range considered. This can be included in the fit if necessary, but is expected to contribute negligibly to the overall fit.

There is no similar decay to consider for the Λ_b^0 decays, and so it is assumed that there are no partially reconstructed background for these signal decays.

The final background category is the peaking background, where tracks are reconstructed, but do not originate from the intermediate resonances. These events will therefore have the same reconstructed invariant mass as the signal decay. The decays considered in this category are $B^0 \to K_s^0\pi^+\pi^-\gamma$ and $\Lambda_b^0 \to \Lambda^0\pi^+\pi^-\gamma$, and for each of these, in both the LL and DD category, there is expected to be fewer than one event passing the pre-selection. It can therefore be assumed that there are no such events present in the data, and therefore this contribution does not need to be added to the fit.

In summary, the background model consists of only an exponential function to model the combinatoric background. It is assumed that no other background processes pass the selection. The validity of this assumption is shown by the quality of the fits in Figure 6.6.

6.4 Cross checks

A number of cross checks are carried out using the control channel to ensure the selection is working as expected and a clean sample of the decay we are interested in is obtained.

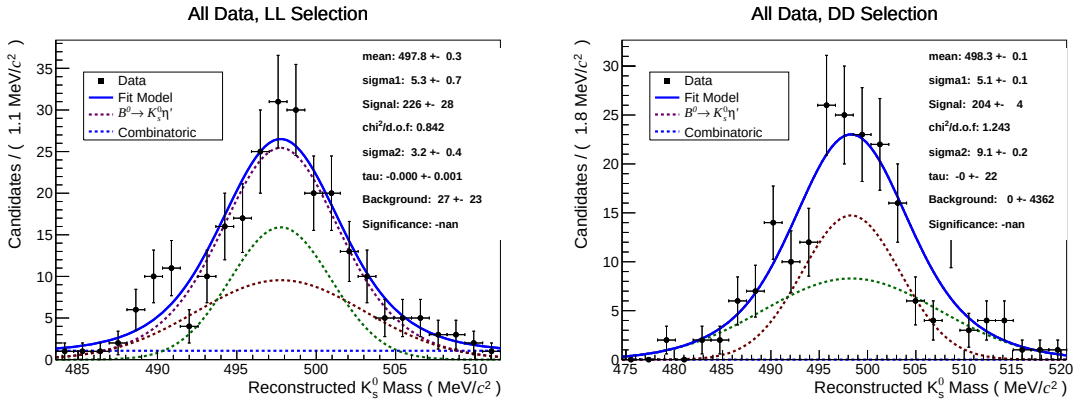

Figure 6.7: Fit to the K^0_S mass distribution for events in the B^0 mass region for the LL (left) and DD (right) categories in data. The solid blue line shows the overall fit and the dashed lines show each component of the fit.

The first check is to compare the actual yields obtained, shown in Table 6.5, with those expected from the efficiency measurements, presented in Table 6.3. In the LL category, we expect a total of 333 ± 83 events and observe 188 ± 16 events. In the DD category, we expect a total of 317 ± 88 events and observe 149 ± 14 events. With the large uncertainties on the expected yield, this is a deviation of less than two standard deviations in each category. This difference has been observed in the analysis of similar decays in LHCb and a detailed investigation has revealed that the efficiency of the L0 trigger is not well modelled in the MC. The L0 hadron trigger used in this analysis has a cut on the number of hits observed in the SPD. Since it is known that the MC does not model the number of hits in this subdetector, the efficiency of this cut is unreliable. However, assuming that the modelling will behave similarly for the control channel and signal channels, any uncertainty due to this mis-modelling will cancel in the ratio of efficiencies. The difference in expected and observed yields is therefore not considered to be a problem.

It is also interesting to compare the resolution of the fit to data with that of the fit to MC. Since the resolution of the first Gaussian function is allowed to float in the data, it should return a similar value to the fit to the MC, and indeed the values in Table 6.4 with Table 6.5 show excellent agreement between the resolution in data and MC.

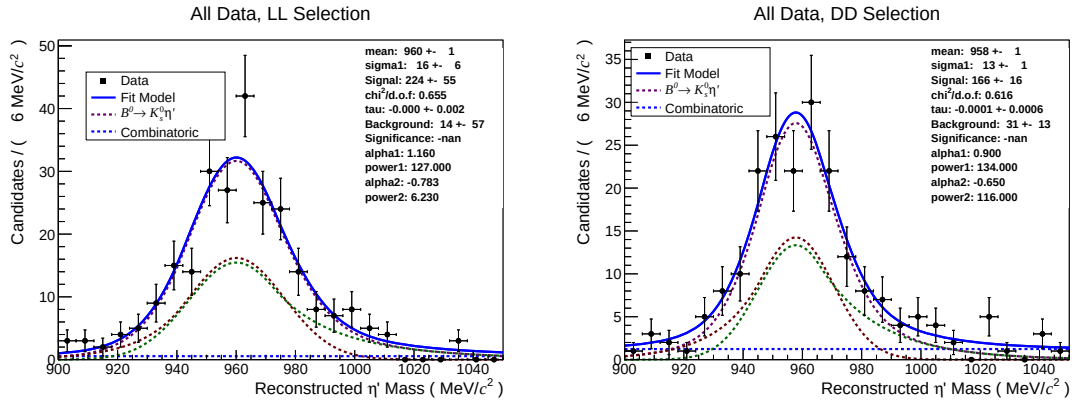

Figure 6.8: Fit to the η' mass distribution for events in the B^0 mass region for the LL (left) and DD (right) categories in data. The solid blue line shows the overall fit and the dashed lines show each component of the fit.

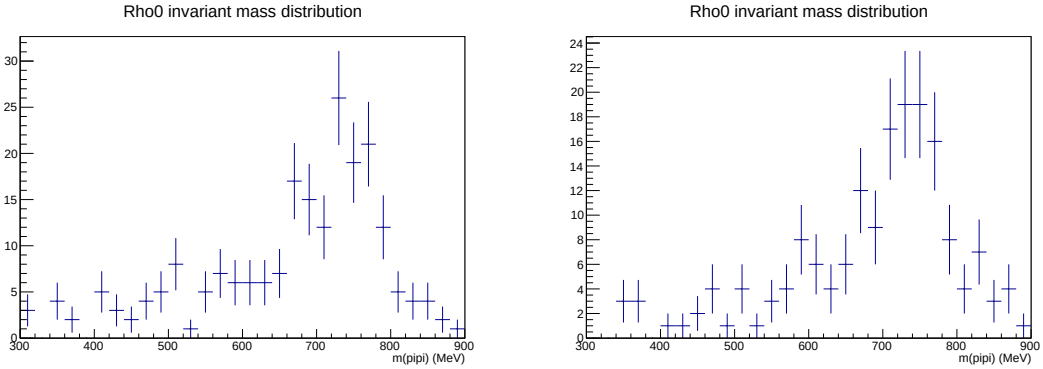

Figure 6.9: Reconstructed $\pi^+\pi^-$ invariant mass from the $B^0 \to K^0_S \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay in the signal window in 2012 data in the LL (left) and DD (right) categories.

To ensure we have a clean sample of $B^0 \to K^0_S \eta'$ decays with no peaking or non-resonant backgrounds, the reconstructed resonances are investigated. The η' and K^0_S masses in the signal mass region are plotted and fit. The fit to the K^0_S masses is shown in Figure 6.7, where the mass is fit with two Gaussian functions with the ratio of widths and the fraction of events in each Gaussian fixed to the values found in the MC fits. The resolutions of the first Gaussian are allowed to float, and in both the LL and DD categories the resolution obtained is consistent with the expectation from MC. The fit to the η' mass is shown in Figure 6.8, and is fit with two Cystal Ball functions, with the tail parameters fixed to the MC values. Once again, the resolution is allowed to float and is found to be consistent with the MC prediction, with the DD events reconstructed with a slightly better resolution than the LL events.

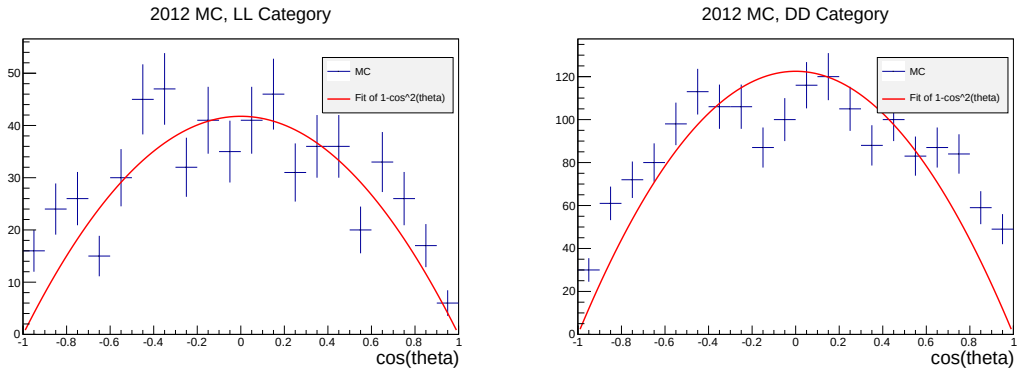

Figure 6.10: Distribution of θ, the angle between the γ and π^+ in the rest frame of the $\pi^+\pi^-$ in the MC in the LL (left) and DD categories (right). Shown in red is the fitted $1-\cos^2(\theta)$ distribution.

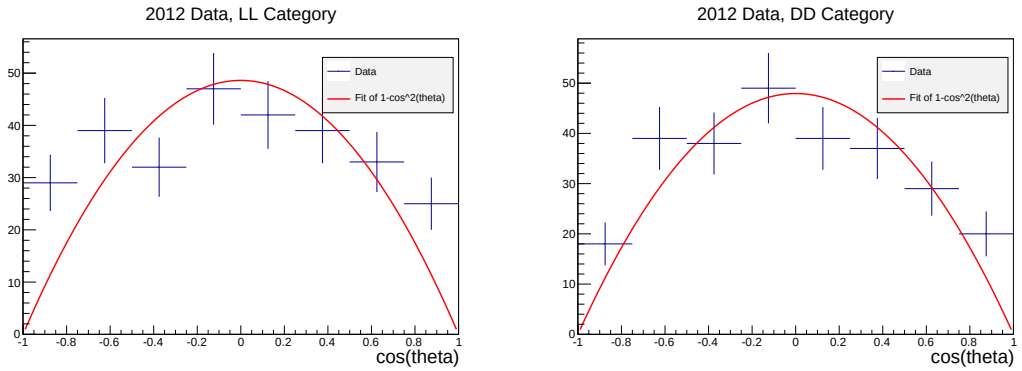

Figure 6.11: Distribution of θ, the angle between the γ and π^+ in the rest frame of the $\pi^+\pi^-$ in the selected data in the LL category (left) and DD category (right). Shown in red is the fitted $1-\cos^2(\theta)$ distribution.

Figure 6.9 shows the reconstructed $\pi^+\pi^-$ invariant mass distribution. This plot shows that the η' decay is dominated by the resonant $\eta' \to \rho^0\gamma$ decay, with a clean ρ^0 signal observed in the data. The signal is asymmetric, with more events at low $\pi^+\pi^-$ masses, as expected from the ρ^0 mass shape which has a radiative tail, and also from the small contribution of non-resonant $\eta' \to \pi^+\pi^-\gamma$ decays.

Another cross-check is to look at the angular distribution of the photon with respect to the pions in the η' decay. Since the photon is a vector particle, its angular distribution in the rest frame of the $\pi^+\pi^-$ pair is well defined, and should follow a $\sin^2$ distribution. To test this, we boost the particles' momenta into the rest frame of the $\pi^+\pi^-$ pair, and measure the angle, θ, between the π^+ and the γ. The expected

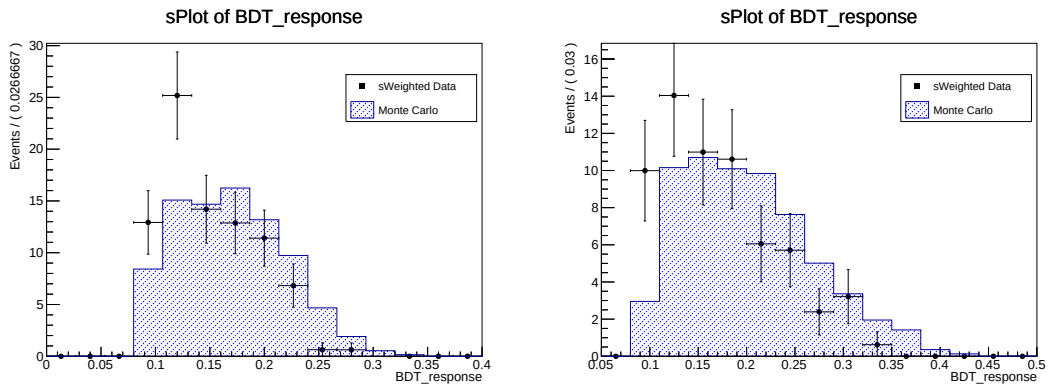

Figure 6.12: sPlots of the BDT response for LL events (left) and DD events (right).

distribution of $\cos(\theta)$ is shown in Figure 6.10, which shows the angular distribution of photons from the MC samples, where the η' decay is modelled using the SVP model. This distribution is fit with a $1 - \cos^2(\theta)$ shape, to show the expected distribution. Figure 6.11 shows the same distributions for the selected events in the 2012 data, limited to events reconstructed within 100 MeV of the signal mass. Shown is each plot is also a fitted $1 - \cos^2(\theta)$ function, which fits the distribution well in both the LL and DD data. It does not fit exactly since the $1 - \cos^2(\theta)$ shape is diluted by the non-resonant $\pi^+\pi^-$ component and the acceptance of the LHCb detector. We can therefore be confident that the events selected in data include a clean sample of $\eta' \to \pi^+\pi^-\gamma$ decays.

6.4.1 sWeighted Data

The final cross-check which is carried out is to compare various distributions in the MC samples with a sample of data from which the background events have been removed. The background is subtracted using the *sPlot* technique [109], which uses the signal and background components of the fit to assign event weights describing how signal-like or background-like the event is. The sample is then reweighted according to the sWeights and a clean sample of signal events is obtained. Of particular interest is the distributions of variables used in the training of the BDT, which are shown in Appendix A.1. The plots show the distributions of each variable

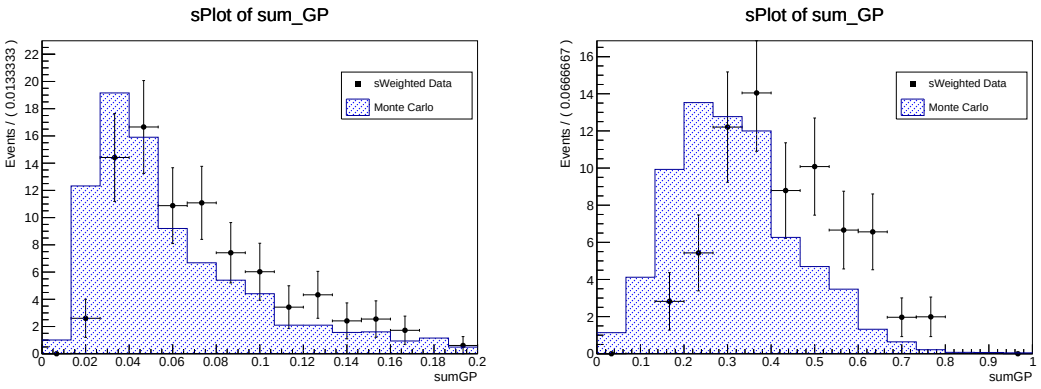

Figure 6.13: sPlots of the sum of the track ghost probability for Long tracks (left) and Downstream tracks (right).

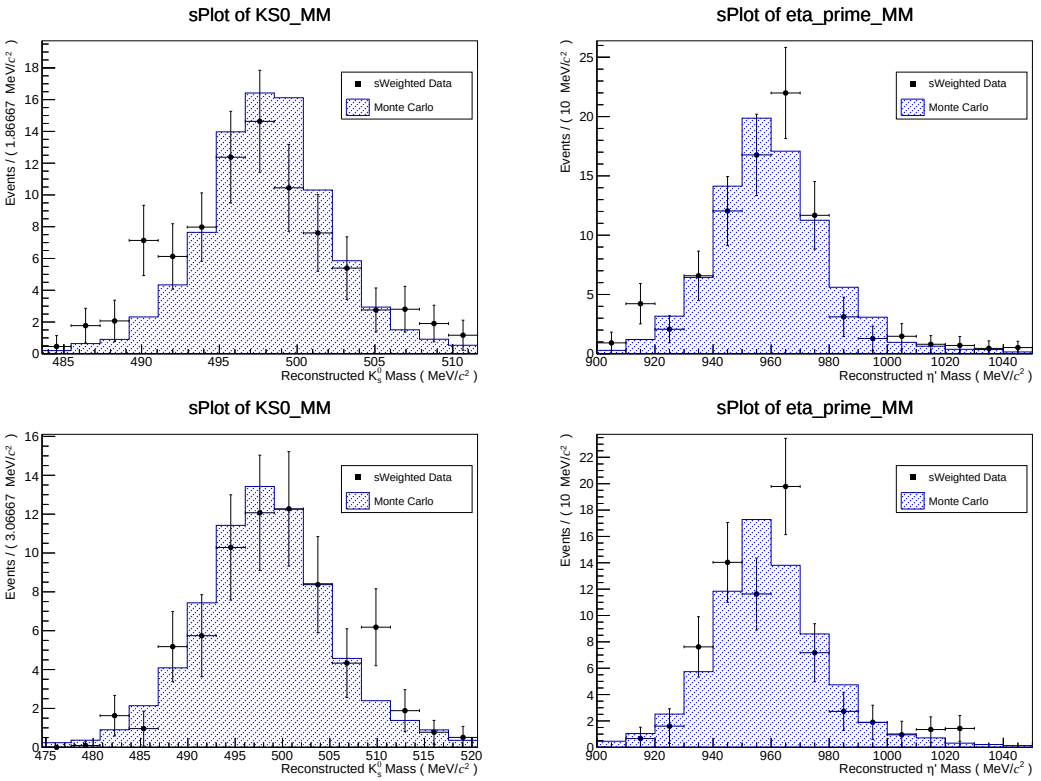

Figure 6.14: sPlots of reconstructed invariant mass of the K^0_S (left) and η' (right) for the $B^0 \to K^0_S \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay reconstructed in 2012 data in the LL (top) and DD (bottom) categories.

used in the BDT, in both the background-subtracted data, and truth-matched MC samples. The data used are only the $B^0 \to K^0_S \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays reconstructed in 2012 data, since this is the channel with the highest statistics. Since the same variables are used in each BDT, it is assumed that any variables which are described well for this channel, will be described well for all channels. It can be seen that most of the variables are described well by the MC, with the exception of the ghost probability of the tracks. This is shown explicitly in Figure 6.13, which shows the distribution of the sum of ghost probabilities of all the tracks, sum_GP. It can be seen that there are some discrepancies in these distributions, more enhanced in the DD category. The impact of this difference on the BDT performance has been seen to be low. By weighting the sum_GP variable by a constant factor of 1.5, the efficiency of the BDT measured with MC changes by less than 0.1%, where the uncertainty on the efficiency is 0.2%. No systematic uncertainty is applied specifically for this. The *sPlot* distributions for the BDT response are shown in Figure 6.12 and show good agreement, with statistical fluctuations.

A final check performed is to plot the reconstructed invariant mass of the K^0_S and η' resonances, to ensure the shape of the mass distribution matches the expected shape in MC. This is shown in Figure 6.14, and indeed a good agreement is observed.

After performing each of these cross checks one can be confident that the selection is performing as expected and is well-described by the MC samples used. For the control channel a clean sample of $B^0 \to K^0_S \eta'$ was obtained with very little contamination from background processes. These cross checks were carried out using the control channel while the signal channels were still blind. The important consequence of these checks is that the selection can be fixed and the signal channels unblinded with confidence that the selection will not bias the results.

Table 6.6: Relative systematic uncertainties, in percent, on the ratio of branching fractions.

	$\Lambda^0\eta'(\pi^+\pi^-\gamma)$		$\Lambda^0\eta'(\pi^+\pi^-\eta)$		$\Lambda^0\eta(\pi^+\pi^-\pi^0)$	
Source	LL (%)	DD (%)	LL (%)	DD (%)	LL (%)	DD (%)
$\mathcal{B}$ (V^0)	0.73		0.73		0.73	
$\mathcal{B}$ $(\eta^{(\prime)})$	0.0		2.9		2.3	
$f_B/f_{\Lambda_b^0}$	8.0		8.0		8.0	
Fit model	1.7	0.3	1.7	0.3	1.7	0.3
Ratio of ϵ_{acc}	2.0		1.8		1.9	
Ratio of ϵ_{sel}	5.9	5.3	8.6	14.5	7.5	5.9
ϵ_{trig}	1.0		1.0		1.0	
C_γ	–		4.2		4.6	
PID	2.1	0.9	2.0	1.0	2.0	1.0
Multiple Cand	1.4	1.9	2.2	1.4	2.0	1.7
Total	10.6	10.1	13.4	17.5	12.7	11.5

6.5 Systematic uncertainties

Various systematic uncertainties are considered when calculating the ratio of branching fractions given in Equation 5.6. Each term will have a systematic uncertainty associated with it.

For the measured branching fractions, the systematic uncertainty is calculated from the uncertainties given in the PDG [7]. The systematic uncertainty of the ratio of branching fractions is therefore 0.73% for the ratio of K_s^0 to Λ^0 branching fractions. For the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) and $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) decays there is an additional uncertainty due to the branching fractions of the different $\eta^{(\prime)}$decay modes compared with the control channel. The uncertainty is 2.9% for the $\eta' \to \pi^+\pi^-\eta$ decay and 2.3% for the $\eta \to \pi^+\pi^-\pi^0$ decay.

The systematic uncertainty on the ratio of production fractions, $f_d/f_{\Lambda_b^0}$, is taken as the uncertainty on the values measured by LHCb in [110]. $f_d/f_{\Lambda_b^0}$ is found by measuring the efficiency corrected ratio of $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $B^0 \to D^+\pi^-$ decays using the data recorded by LHCb in 2011. In order to extract $f_d/f_{\Lambda_b^0}$ from this measurement, it is necessary to know the branching fractions for these decays, which is not well

measured for the Λ_b^0 decay. To reduce the uncertainty, this measurement is combined with another measurement of this quantity by LHCb which uses semileptonic Λ_b^0 decays to extract $f_d/f_{\Lambda_b^0}$ [110]. The ratio of hadronic yields is plotted, and then scaled up to match the semileptonic result, so that knowledge of the absolute values of branching fractions is not necessary. The uncertainty on this measurement is then dominated by knowledge of the semileptonic results and along with the uncertainty on the average p_T for the Λ_b^0; the uncertainty on this quantity is found to be 8%.

One of the dominant systematic uncertainties on the number of B^0 candidates will be due to the model used to fit the data. In order to assign an uncertainty to this, the parameters which are fixed in the fit are varied within the uncertainties on the parameters obtained from MC. The uncertainty is then the relative change in the number of signal events observed by varying the fixed parameters. In addition, for the DD sample, the fit is performed with and without the correction to N_2/N_1. There is no difference in the number of signal events observed in these fits, and so no additional systematic uncertainty is applied.

The systematic uncertainties due to the efficiencies measured are also calculated, and the uncertainty on the ratio of selection efficiencies is just the statistical uncertainty on the measured efficiencies, due to the limited statistics of the MC samples.

In addition, there is a systematic uncertainty on the efficiency of applying the multiple candidate selection to signal events. This quantity has been calculated using MC signal samples, but it is unclear how well the underlying event is modelled, and hence a systematic uncertainty is assigned to the fraction of events with multiple candidates, and the photon p_T spectrum of these multiple candidates. This uncertainty is difficult to quantify unambiguously and so a conservative uncertainty is applied: it is assumed that the fraction of events with multiple candidates, and the probability of selecting the correct candidate both have an uncertainty of 10%.

The systematic uncertainty due to the difference between the PID distributions in data and MC is taken as the difference between the PID efficiencies measured by MC

and the calibration data.

The systematic uncertainty due to the measurement of the L0 efficiency in 2012 Monte Carlo samples is 1%. This is from the 5% uncertainty on the measured L0 efficiency of the 2012a sample, which makes up 17% of the total L0 efficiency.

There is a further systematic uncertainty due to the measured value of the photon correction factor, $1/C_\gamma$. For the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, there is no correction applied, and no systematic uncertainty needs to be assigned. For the other signal modes, the uncertainty on this correction comes from the uncertainties presented in Ref [101]. The correction factor is measured by comparing the yields of $B^+ \to J/\psi K^{*+}$ with $K^{*+} \to K^+\pi^0$ and $B^+ \to J/\psi K^+$ decays. The dominant uncertainty is from the measurement of the efficiencies of these decays, in particular the trigger efficiency and detector acceptance efficiencies. The correction factor is calculated as the weighted average of the correction factors in each bin of photon p_T, and the uncertainty on this average is found using the standard formula

$$\sigma_C^2 = \frac{1}{\sum(\frac{1}{\sigma_i^2})}. \tag{6.12}$$

A summary of all systematic uncertainties is shown in Table 6.6. The errors are presented as percentages, and the total systematic error is a quadrature sum of all the systematic uncertainties.

6.6 Yields of signal channels

The Λ_b^0 invariant mass distributions after unblinding are shown in Figures 6.15, 6.16 and 6.17 for each of the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$), $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) and $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) respectively. An extended unbinned maximum likelihood fit is performed to the distributions and the model used is the same as that used for the control channel. The only difference is that, due to the limited statistics,

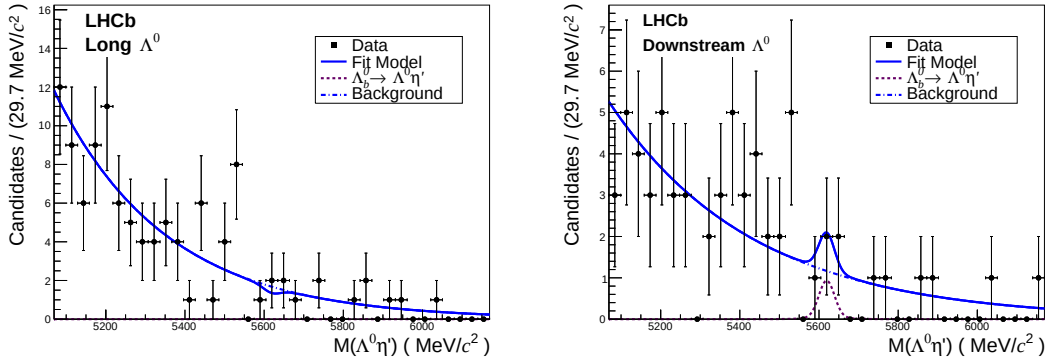

Figure 6.15: Reconstructed $\Lambda^0\eta'$ invariant mass from the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay, combining 2011 and 2012 data in the LL (left) and DD (right) categories.

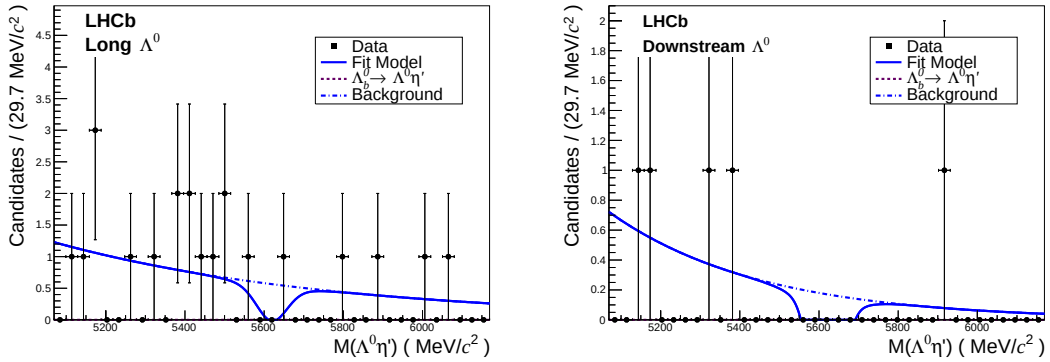

Figure 6.16: Reconstructed $\Lambda^0\eta'$ invariant mass reconstructed from the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) decay, combining 2011 and 2012 data in the LL (left) category and 2012 data in the DD (right) category.

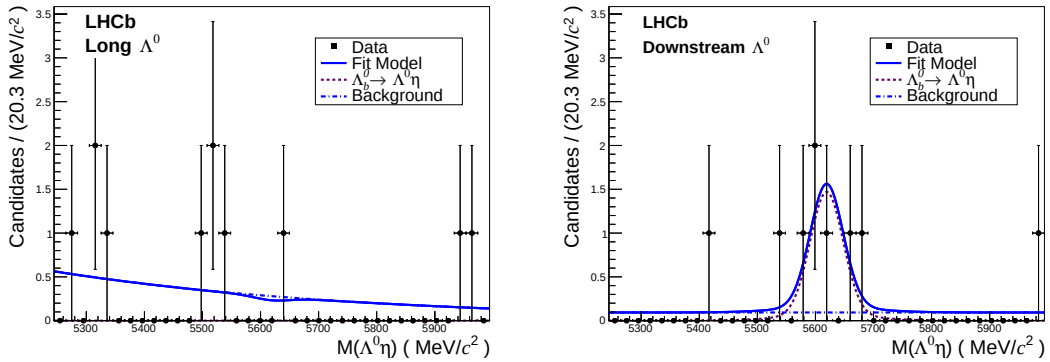

Figure 6.17: Reconstructed $\Lambda^0\eta$ invariant mass from the $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) decay, combining 2011 and 2012 data in the LL (left) category and 2012 data in the DD (right) category.

Table 6.7: Yields obtained from the unbinned maximum likelihood fit to data.

Decay	Cat	$N_{\rm bkg}$	$N_{\rm S}$	Significance
$\Lambda_{\rm b}^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$)	LL	19 ± 3	0 ± 2	0.2
	DD	14 ± 2	1 ± 2	0.7
$\Lambda_{\rm b}^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$)	LL	6 ± 1	-2 ± 1	1.0
	DD	2.2 ± 0.8	-2.1 ± 0.2	2.1
$\Lambda_{\rm b}^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$)	LL	5 ± 1	0 ± 1	0.3
	DD	0.25 ± 0.9	5 ± 3	3.0

all parameters of the fit are fixed to the values found from the signal MC samples, including the resolution of the first Gaussian function. Only the signal and background yields are allowed to float, and these are given in Table 6.7 along with the significance of each signal. No significant signal is seen in any channel except for the $\Lambda_{\rm b}^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) decay reconstructed in the DD category, where evidence is seen for this decay at the level of 3σ. Usually in HEP a significance of 3σ would constitute evidence for a decay. However, due to the lack of any signal in the LL category, the interpretation of this signal is more difficult. Before the selection was fixed it was decided that the Feldman-Cousins (FC) method would be used to find the branching fractions of the signal decays. The advantage of this method is that the result is the same independent of how the signal is interpreted, and so no claim is made on the discovery of this mode. The extraction of the branching fractions is shown in Section 6.7.

6.7 Feldman-Cousins intervals

A common practice in HEP is to construct both two sided and one sided CL intervals and to choose which to use based on the signal observed. For example, if the significance of the signal is greater than 3σ then the two sided limits are presented, else the one sided limits are presented. The problem with this strategy is that the results are inherently biased by the observation. This is known as "flip-flopping", and has been shown by Feldman and Cousins to lead to under-coverage of the

Table 6.8: Components contributing to the scale factor, α.

	$\Lambda^0\eta'(\pi^+\pi^-\gamma)$		$\Lambda^0\eta'(\pi^+\pi^-\eta)$		$\Lambda^0\eta(\pi^+\pi^-\pi^0)$	
Factor	LL	DD	LL	DD	LL	DD
$\epsilon_{acc}(B^0)/\epsilon_{acc}(\Lambda_b^0)$	1.00 ± 0.02		1.09 ± 0.02		1.03 ± 0.02	
$\epsilon_{tot}(B^0)/\epsilon_{tot}(\Lambda_b^0)$	2.3 ± 0.1	1.55 ± 0.08	7.4 ± 0.6	9.5 ± 1.4	4.6 ± 0.3	3.4 ± 0.2
$f_B/f_{\Lambda_b^0}$	2.5 ± 0.2		2.5 ± 0.2		2.5 ± 0.2	
$1/C_\gamma$	1 (fixed)		0.95 ± 0.04		1.13 ± 0.04	
$0.5 \cdot \mathcal{B}(K_S^0)/\mathcal{B}(\Lambda^0)$	0.541 ± 0.004		0.541 ± 0.004		0.541 ± 0.004	
$\mathcal{B}(\eta')/\mathcal{B}(\eta^{(\prime)})$	1 (fixed)		1.71 ± 0.05		1.31 ± 0.03	
α	3.1 ± 0.3	2.1 ± 0.2	17.7 ± 2.3	22.8 ± 4.0	9.5 ± 1.2	7.0 ± 0.8

intervals [111]. The same authors have devised a unified approach to combat this problem. This is the method adopted in this analysis to find the branching fractions for the signal decays, and is described in this section.

The FC method constructs confidence intervals using an ordering principle based on the ratio of likelihoods

$$R_L = \frac{P(\text{Nfit}|\mu)}{P(\text{Nfit}|\mu_{\text{best}})} \tag{6.13}$$

where P is the probability of observing Nfit events if the mean number of signal events is μ and μ_{best} is the value of μ which maximises P(Nfit|μ). It is constrained to the physical region (non-negative) by requiring $\mu_{\text{best}} = \max(0,\text{Nfit})$. For a given value of μ, values of Nfit are added into the acceptance region, in decreasing order of R_L, until the sum of probabilities exceeds the required confidence limit, i.e.

$$\int_{n_{lo}}^{n_{hi}} P(\text{Nfit}|\mu)dn \geq 0.9 \tag{6.14}$$

where the integral is calculated numerically.

This is repeated for a range of values of μ to construct an upper and a lower bound for the number of observed events. If the lower bound is zero then an upper limit is placed, and if the number of observed events is large enough, then a two sided limit is applied, along with a central value.

Here, we use the FC method to place a limit on $R = \mu \times \alpha/N_S(B^0)$, where R is the

ratio of branching fractions,

$$
\begin{aligned}
R &= \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta')}{\mathcal{B}(B^0 \to K^0 \eta')} \\
&= \frac{N_S(\Lambda_b^0)}{N_S(B^0)} \times \frac{\epsilon_{acc}(B^0)}{\epsilon_{acc}(\Lambda_b^0)} \times \frac{\epsilon_{tot}(B^0)}{\epsilon_{tot}(\Lambda_b^0)} \times \frac{f_d}{f_{\Lambda^0}} \times \frac{1}{C_\gamma} \times \frac{0.5 \times \mathcal{B}(K_S^0)}{\mathcal{B}(\Lambda^0)} \times \frac{\mathcal{B}(\eta')}{\mathcal{B}(\eta^{(\prime)})} \\
&= \alpha \times \frac{N_S(\Lambda_b^0)}{N_S(B^0)}
\end{aligned}
\tag{6.15}
$$

and the factors determining α are shown in Table 6.8.

In this case the probability, P(Nfit|μ), used in the FC method is defined by a Gaussian function, with a resolution of $\sigma_\mu = \sqrt{\sigma_{syst}^2 + \sigma_{stat}^2}$. The systematic uncertainty, σ_{syst} is the relative systematic uncertainty taken from Table 6.6. The statistical uncertainty on the number of signal events is the uncertainty due to the fit. This has been calculated using toy MC studies. A total of 1000 toy MC samples are generated according to the signal model found by the fit to the signal MC samples, with the background model and number of background events fixed to those found in the data. The model is then allowed to float, and the number of signal events along with its uncertainty is found from a fit to the generated data. This is then repeated for fixed numbers of signal events and Figure 6.18 shows the number of fitted events as a function of the number of generated events. There is some evidence in these plots of a small bias introduced by the fit, where the number of fitted events, Nfit, can be parameterised as a linear function of the number of generated events, Ngen, by the form:

$$\text{Nfit} = p_1 \times \text{Ngen} + p_0. \tag{6.16}$$

This bias will need to be taken into account when setting a limit on the branching fraction based on the number of signal events obtained from the fit, and to do this the values of p_1 and p_0 are given in Table 6.9 for each signal decay. Also shown is the corrected number of observed events (measured n) which is used to obtain the confidence limits.

Figure 6.19 shows the square of the error on the number of signal events as a function

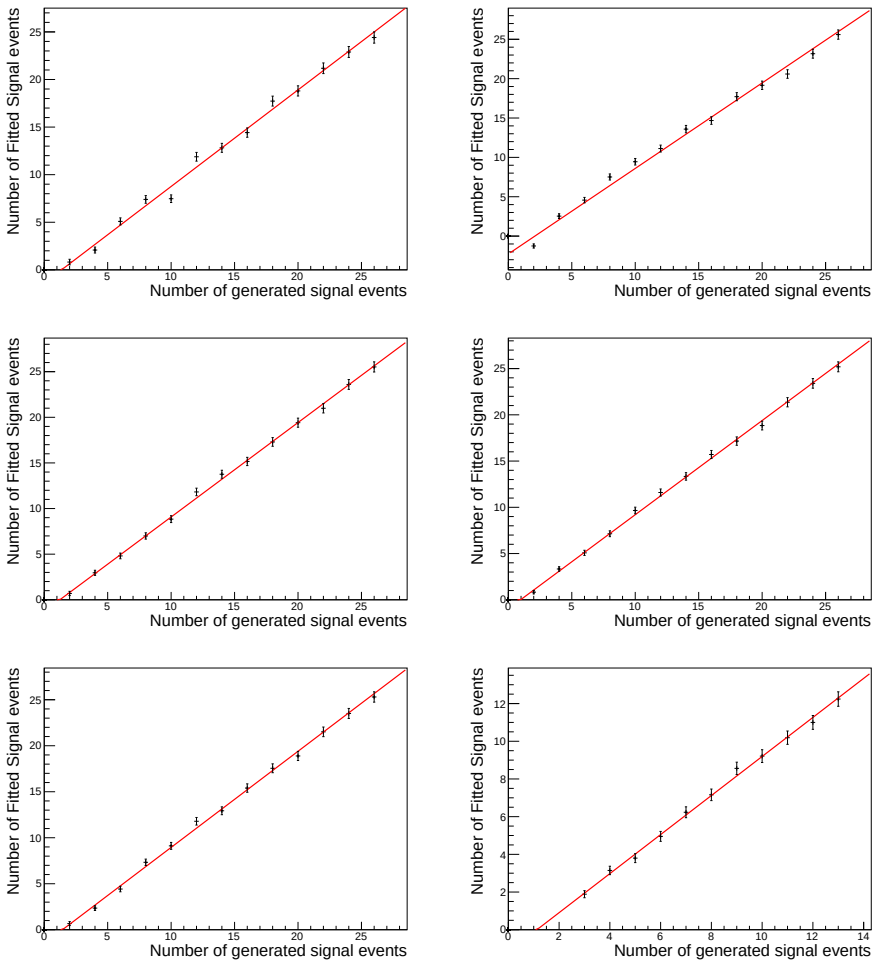

Figure 6.18: Plots to show the number of events obtained from fits against the number of events generated in toy MC studies of each of the signal decay modes, $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) (top), $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\eta$) (middle) and $\Lambda_b^0 \to \Lambda^0 \eta$ ($\eta \to \pi^+\pi^-\pi^0$) (bottom), for LL (left) and DD (right) events.

Table 6.9: Parameters obtained from a linear fit to the number of fit events against number of generated events for each signal decay, and the corrected number of observed events.

Decay	**Cat**	p_1	p_0	measured n
$\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$)	LL	1.002 ± 0.001	-1.01 ± 0.09	-1.01
	DD	1.006 ± 0.001	-1.06 ± 0.09	1.06
$\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\eta$)	LL	1.003 ± 0.001	-0.62 ± 0.08	-2.63
	DD	1.004 ± 0.001	-0.70 ± 0.08	-2.8
$\Lambda_b^0 \to \Lambda^0 \eta$ ($\eta \to \pi^+\pi^-\pi^0$)	LL	1.002 ± 0.001	-0.48 ± 0.08	-0.48
	DD	1.002 ± 0.001	-0.67 ± 0.08	4.34

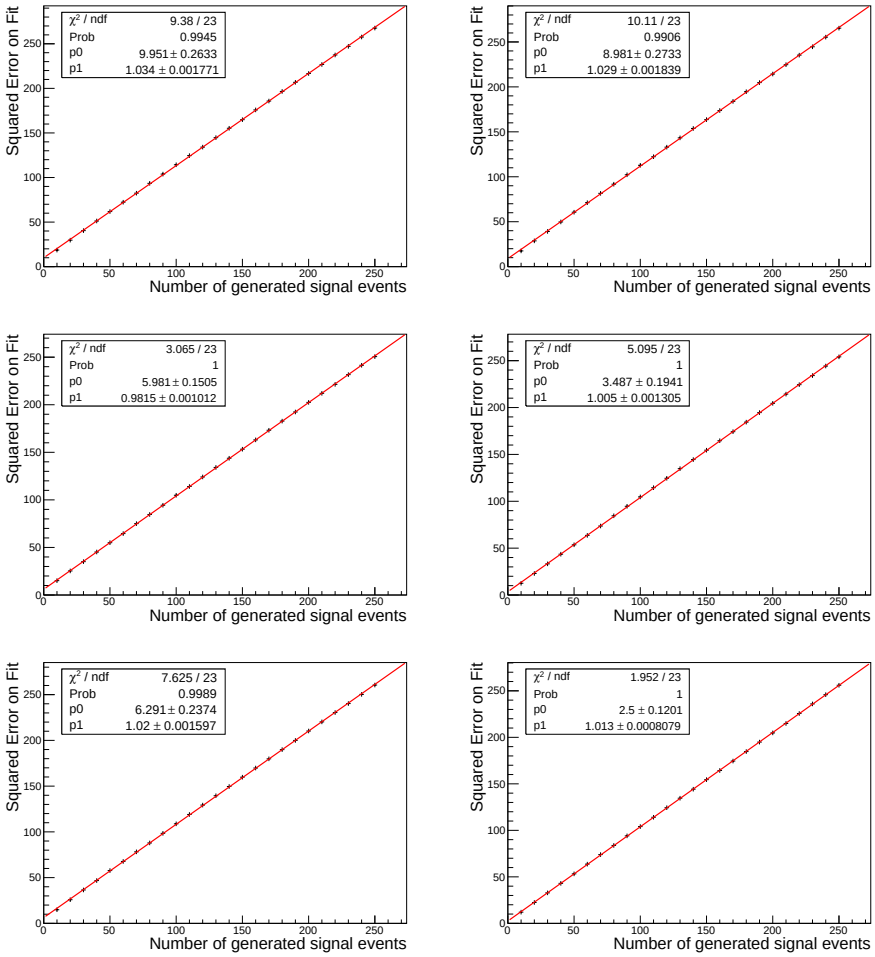

Figure 6.19: Plots to show the squared statistical error on the number of fit events against the number of events generated in toy MC studies of each of the signal decay modes, $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) (top), $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) (middle) and $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) (bottom), for LL (left) and DD (right) events.

of the number of events generated. The uncertainty is proportional to the square root of the number of generated events, and the statistical error on μ is given by:

$$\sigma_{\text{stat}(\mu)} = \sqrt{p_3 \times \mu + p_2}. \tag{6.17}$$

The values of p_3 and p_2 obtained from linear fits to the plots in Figure 6.19 are given in Table 6.10, and will be used to calculate the statistical uncertainty used in constructing the FC intervals.

Table 6.10: Parameters obtained from a linear fit to the square of the error on the fit against number of generated events for each signal decay.

Decay	**Cat**	p_3	p_2
$\Lambda_{\rm b}^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$)	LL	1.034 ± 0.002	9.95 ± 0.26
	DD	1.029 ± 0.002	8.98 ± 0.27
$\Lambda_{\rm b}^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$)	LL	0.982 ± 0.001	5.98 ± 0.15
	DD	1.005 ± 0.001	3.49 ± 0.19
$\Lambda_{\rm b}^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$)	LL	1.020 ± 0.002	6.29 ± 0.24
	DD	1.013 ± 0.001	2.50 ± 0.12

The limits on the branching fractions are calculated for each category separately and the categories are combined. This is done by extending the ordering principle to incorporate the likelihood of each selection. The confidence bands are built by calculating the likelihood ratio, R_L using a Gaussian function with mean of $\hat{\mu}$, which is the weighted average mean of all categories in the combination, given by

$$\hat{\mu} = \frac{\sum \frac{\mu_i \times \alpha_i / N_S(\mathrm{B}^0)}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2}} \times \frac{N_S\hat{(\mathrm{B}^0)}}{\hat{\alpha}}, \tag{6.18}$$

where μ_i is the mean number of events in a particular category, which is multiplied by $\alpha/N_S(\mathrm{B}^0)$ to calculate the average ratio of branching fractions for a given mean, σ_i is the uncertainty on this ratio, and $\hat{\alpha}$ and $N_S\hat{(\mathrm{B}^0)}$ are the average values of α and number of signal events in the normalisation channel. The uncertainty used in this calculation is only the uncorrelated systematic uncertainties. The uncertainty on the ratio of hadronisation fractions is not included in the calculation since it is fully correlated between all decays. The same method is used to calculate the weighted average number of events extracted from the fit, and the confidence intervals are the constructed in the same way as before.

6.7.1 Measured branching fractions

Using the signal yields and the confidence belts constructed, limits on the ratio of branching fractions are obtained for each signal decay.

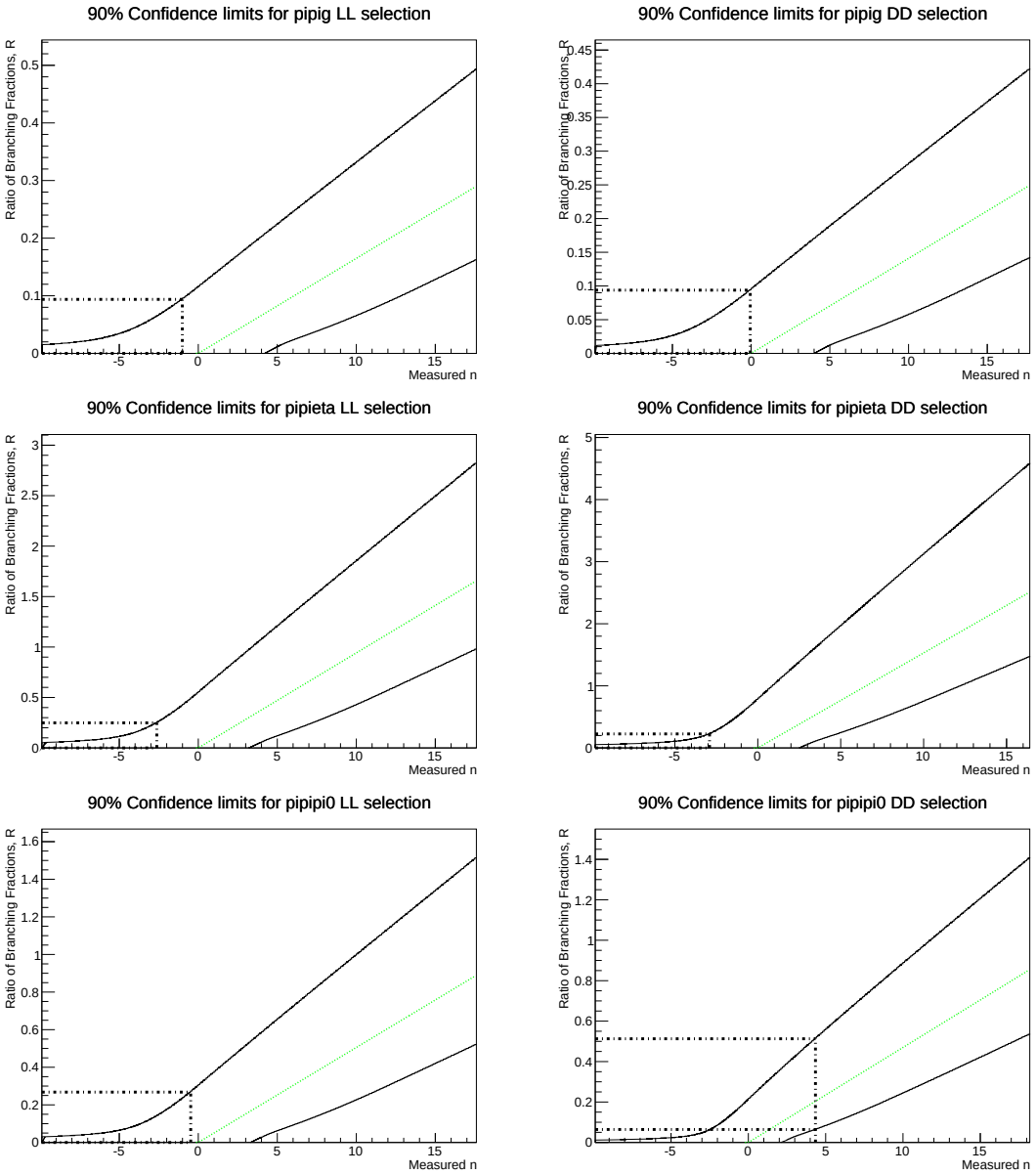

Figure 6.20: 90% confidence belts for each of the signal decay modes, $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) (top), $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) (middle) and $\Lambda_b^0 \to \Lambda^0\eta$ ($\eta \to \pi^+\pi^-\pi^0$) (bottom), for LL (left) and DD (right) events.

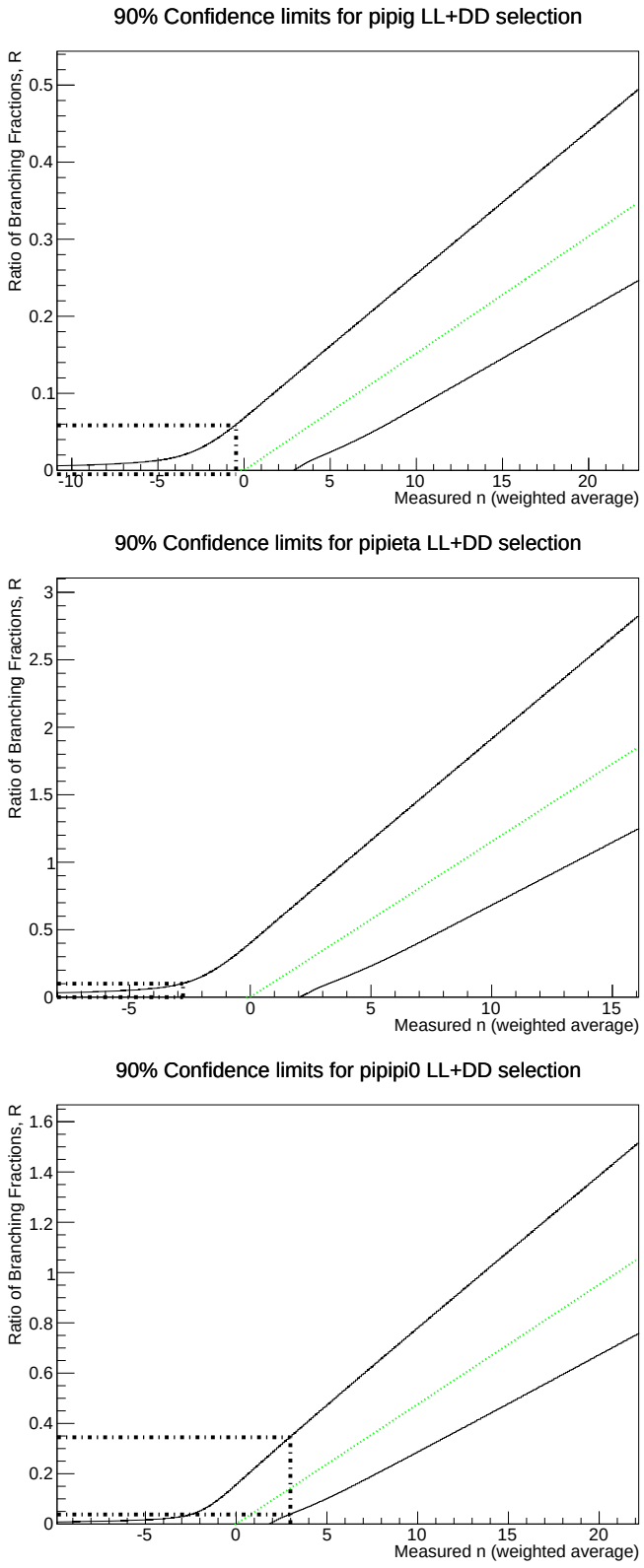

Figure 6.21: 90% confidence belts for each of the signal decay modes, combining the LL and DD selections, $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) (top), $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\eta$) (middle) and $\Lambda_b^0 \to \Lambda^0 \eta$ ($\eta \to \pi^+\pi^-\pi^0$) (bottom).

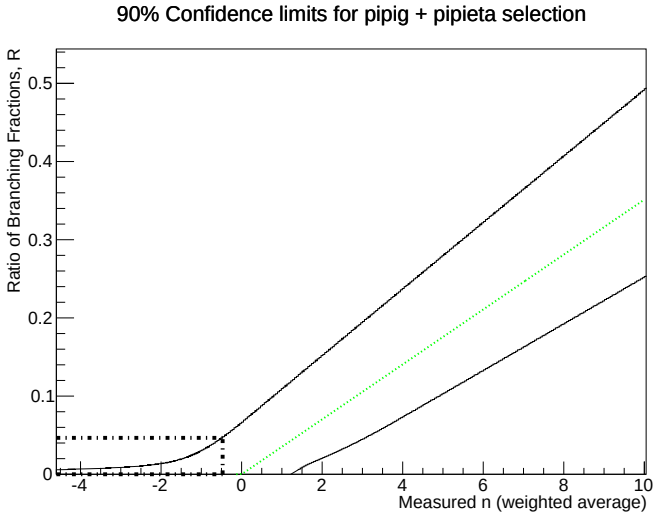

Figure 6.22: 90% confidence belts for each of the signal decay modes, combining the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) and $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) selections.

Figure 6.20 shows the CL intervals obtained using the FC method and the dashed lines show the limits obtained using the observed signal yields. Figure 6.21 shows the CL intervals obtained when the LL and DD categories are combined, and Figure 6.22 shows the CL intervals obtained by combining the two η' decay modes.

The limit on the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decay is found to be

$$R < 0.0938 \text{ at } 90\% \text{ CL}$$

$$R < 0.0938 \text{ at } 90\% \text{ CL}$$

in the LL and DD category respectively. Combining the two categories gives a limit of

$$R < 0.0584 \text{ at } 90\% \text{ CL}.$$

Using the $\Lambda_b^0 \to \Lambda^0\eta'$ ($\eta' \to \pi^+\pi^-\eta$) decay, the limit is found to be

$$R < 0.248 \text{ at } 90\% \text{ CL}$$

$$R < 0.226 \text{ at } 90\% \text{ CL}$$

in the LL and DD category respectively. Combining the two categories gives a limit

of

$$R < 0.100 \text{ at } 90\% \text{ CL.}$$

Combining the $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) and $\Lambda_b^0 \to \Lambda^0 \eta'$ ($\eta' \to \pi^+\pi^-\eta$) limits gives a limit on the ratio of $\mathcal{B}$ ($\Lambda_b^0 \to \Lambda^0 \eta'$)/$\mathcal{B}$ ($B^0 \to K_s^0 \eta'$) of

$$R < 0.0466 \text{ at } 90\% \text{ CL.}$$

Multiplying by the branching fraction of $B^0 \to K^0 \eta'$ gives a limit on the absolute branching fraction of

$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta') < 3.1 \times 10^{-6} \text{ at } 90\% \text{ CL.}$$

The limit on the $\Lambda_b^0 \to \Lambda^0 \eta$ ($\eta \to \pi^+\pi^-\pi^0$) decay is found to be

$$R < 0.268 \text{ at } 90\% \text{ CL}$$
$$R \in [0.0643, 0.513] \text{ at } 90\% \text{ CL}$$

in the LL and DD category resectively. The signal yield observed in the DD category gives a two sided limit since the significance is 3σ. Combining the two categories gives a limit of

$$R \in [0.038, 0.345] \text{ at } 90\% \text{ CL.}$$

The absolute branching fraction then has a limit of

$$\mathcal{B}(\Lambda_b^0 \to \Lambda^0 \eta) \in [2.5, 22.8] \times 10^{-6} \text{ at } 90\% \text{ CL.}$$

CHAPTER

7

CONCLUSIONS AND OUTLOOK

A search is performed for the $\Lambda_b^0 \to \Lambda^0\eta'$ and $\Lambda_b^0 \to \Lambda^0\eta$ decays in the full dataset recorded by the LHCb experiment during 2011 and 2012, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. A multivariate analysis is used to select signal events. No significant signal is observed above the background events for the $\Lambda_b^0 \to \Lambda^0\eta'$ decay, and some evidence is seen for the $\Lambda_b^0 \to \Lambda^0\eta$ at the level of 3σ. The $B^0 \to K^0\eta'$ decay is observed and used as a control channel, so that a limit is placed on the ratio of $\Lambda_b^0 \to \Lambda^0\eta^{(\prime)}$ branching fractions with respect to the $B^0 \to K^0\eta'$ branching fraction using the Unified Approach of Feldman and Cousins. Using the known value of the $B^0 \to K^0\eta'$ branching fraction, the upper limit on the absolute branching fraction of the $\Lambda_b^0 \to \Lambda^0\eta'$ decay is $\mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta') < 4.0 \times 10^{-6}$ at 90% CL. The two sided limit on the $\Lambda_b^0 \to \Lambda^0\eta$ branching fraction is $\mathcal{B}(\Lambda_b^0 \to \Lambda^0\eta) \in [2.5, 22.8] \times 10^{-6}$ at 90% CL.

These limits can be compared with the branching fractions calculated in Ref [63].

The predicted branching fraction depends strongly on the method used to calculate the hadronic form factors, and on the parameters used in the calculation. Using the QCD sum rules approach, the branching fractions are predicted to be in the range $(6.0-19.0)\times10^{-6}$ where the branching fraction for the $\Lambda_b^0\to\Lambda^0\eta'$ decay is expected to be similar to the branching fraction for the $\Lambda_b^0\to\Lambda^0\eta$ decay. Using the pole model approach to calculate the hadronic form factors predicts smaller branching fractions, in the range $(1.8-4.5)\times10^{-6}$. Our results favour the branching fractions calculated using the pole model to estimate the hadronic form factors. In addition, our results are inconsistent with the branching fraction predictions neglecting the anomalous contribution, which would lead to a high predicted branching fraction.

The main limitation of this analysis is the lack of statistics available to observe the signal channels. There are many Λ_b^0 baryons produced in pp collisions in the LHCb experiment, with half as many Λ_b^0 baryons as b-mesons produced. This is a much higher fraction than expected, and makes LHCb the first experiment able to observe and measure Λ_b^0 decays with precision. However, due to the purely hadronic final states of the decays of interest for this analysis, the efficiency for reconstructing and observing these decays is quite low. The limiting part of the selection in this respect is the low efficiency for triggering events containing the interesting decays because the pions from the signal decays have relatively low momenta. Since the trigger can only accept high energy hadrons to reduce the contamination from background events, the efficiency for selecting signal events is low. During the LHC run II, starting in 2015 with pp collisions at $\sqrt{s}=13\,\text{TeV}$, the cross section for background events will be much higher, and so the trigger will need to be more restrictive, meaning the trigger efficiency is likely to be even lower for signal events. Although the cross section for Λ_b^0 baryon production will also increase, this may be offset by the lower efficiency meaning more data will need to be collected in order to observe the signal decays. However, the limits can be improved by reducing the systematic uncertainty on the ratio of branching fractions. The dominant uncertainty is due to the measurement of the ratio of production fractions, $f_{\Lambda_b^0}/f_{B^0}$. This measurement has been made by the LHCb collaboration using only the data collected in 2011.

The uncertainty could therefore be reduced by taking advantage of the full statistics available including the 2012 dataset, and also by making an improved measurement of the branching fraction of the $\Lambda_b^0 \to \Lambda_c^+ \pi^-$ decay.

In order to get a true picture of η–η' mixing from Λ_b^0 baryon decays, as many different decays as possible need to be studied. The next logical decay to search for would be the $\Lambda_b^0 \to pK\eta^{(\prime)}$ decays, which are produced through the excited resonances of the Λ^0 baryons and so proceed through identical Feynman diagrams to the $\Lambda_b^0 \to \Lambda^0\eta^{(\prime)}$ decays. The advantage of these decays is that they do not contain the long-lived Λ^0 baryon, and so the reconstruction efficiency is likely to be higher. Work is already under way to search for these decays, and the search for the $\Lambda_b^0 \to \Lambda^0\eta^{(\prime)}$ decays will be updated when the full run II data set is available.

REFERENCES

[1] R. Aaij *et al.*, "Search for the $\Lambda_b^0 \to \Lambda\eta$ and $\Lambda_b \to \Lambda\eta'$ decays with the LHCb detector," 2015. submitted to JHEP.

[2] P. A. Dirac, "The Quantum theory of electron," *Proc.Roy.Soc.Lond.*, vol. A117, pp. 610–624, 1928.

[3] P. Dirac, "The Quantum theory of electron. 2.," *Proc.Roy.Soc.Lond.*, vol. A118, p. 351, 1928.

[4] C. Anderson, "The Positive Electron," *Phys.Rev.*, vol. 43, pp. 491–494, 1933.

[5] M. Gell-Mann, "A Schematic Model of Baryons and Mesons," *Phys.Lett.*, vol. 8, pp. 214–215, 1964.

[6] E. D. Bloom, D. Coward, H. DeStaebler, J. Drees, G. Miller, *et al.*, "High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees," *Phys.Rev.Lett.*, vol. 23, pp. 930–934, 1969.

[7] K. A. Olive *et al.*, "Review of particle physics," *Chin. Phys.*, vol. C38, p. 090001, 2014.

[8] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys.Rev.Lett.*, vol. 13, pp. 321–323, 1964.

[9] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys.Lett.*, vol. 12, pp. 132–133, 1964.

[10] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys.Rev.Lett.*, vol. 13, pp. 508–509, 1964.

[11] G. Guralnik, C. Hagen, and T. Kibble, "Global Conservation Laws and Massless Particles," *Phys.Rev.Lett.*, vol. 13, pp. 585–587, 1964.

[12] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons," *Phys.Rev.*, vol. 145, pp. 1156–1163, 1966.

[13] T. Kibble, "Symmetry breaking in nonAbelian gauge theories," *Phys.Rev.*, vol. 155, pp. 1554–1561, 1967.

[14] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012.

[15] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys.Lett.*, vol. B716, pp. 1–29, 2012.

[16] L. E. W. Group, "Precision Electroweak Measurements and Constraints on the Standard Model," 2010.

[17] A. H. Peter, "Dark matter," *PoS*, vol. BASH11, p. 014, 2011.

[18] A. Oliveira, "Gravity particles from Warped Extra Dimensions, a review. Part I - KK Graviton," 2014.

[19] A. Pestana Morais, "Grand unification phenomenology at the LHC and beyond,"

[20] K. Petraki and R. Volkas, "Review of asymmetric dark matter," *Int.J.Mod.Phys*, vol. A28, p. 1330028, 2013.

[21] H. Yukawa, "On the interaction of elementary particles," *Proc.Phys.Math.Soc.Jap.*, vol. 17, pp. 48–57, 1935.

[22] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys.Rev.Lett.*, vol. 10, pp. 531–533, 1963.

[23] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog.Theor.Phys.*, vol. 49, pp. 652–657, 1973.

[24] L.-L. Chau and W.-Y. Keung, "Comments on the Parametrization of the Kobayashi-Maskawa Matrix," *Phys.Rev.Lett.*, vol. 53, p. 1802, 1984.

[25] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys.Rev.Lett.*, vol. 51, p. 1945, 1983.

[26] C. Jarlskog, "Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation," *Phys.Rev.Lett.*, vol. 55, p. 1039, 1985.

[27] J. Hardy and I. Towner, "Superallowed $0+ \to 0+$ nuclear beta decays: A New survey with precision tests of the conserved vector current hypothesis and the standard model," *Phys.Rev.*, vol. C79, p. 055502, 2009.

[28] D. Besson *et al.*, "Improved measurements of D meson semileptonic decays to pi and K mesons," *Phys.Rev.*, vol. D80, p. 032005, 2009.

[29] L. Widhalm *et al.*, "Measurement of D0 $\rightarrow$ pi l nu (Kl nu) form factors and absolute branching fractions," *Phys.Rev.Lett.*, vol. 97, p. 061804, 2006.

[30] S. Aoki, Y. Aoki, C. Bernard, T. Blum, G. Colangelo, *et al.*, "Review of lattice results concerning low-energy particle physics," *Eur.Phys.J.*, vol. C74, no. 9, p. 2890, 2014.

[31] D. Acosta *et al.*, "Measurement of $B(t \rightarrow Wb)/B(t \rightarrow Wq)$ at the Collider Detector at Fermilab," *Phys.Rev.Lett.*, vol. 95, p. 102002, 2005.

[32] V. Abazov *et al.*, "Precision measurement of the ratio B($t \rightarrow Wb$)/B($t \rightarrow Wq$) and Extraction of V_{tb}," *Phys.Rev.Lett.*, vol. 107, p. 121802, 2011.

[33] C. Collaboration, "First measurement of B(t $\rightarrow$ Wb)/B(t $\rightarrow$ Wq) in the dilepton channel in pp collisions at sqrt(s)=7 TeV," 2012.

[34] V. M. Abazov *et al.*, "Evidence for s-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," *Phys.Lett.*, vol. B726, pp. 656–664, 2013.

[35] S. Chatrchyan *et al.*, "Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 1212, p. 035, 2012.

[36] G. Aad *et al.*, "Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys.Lett.*, vol. B717, pp. 330–350, 2012.

[37] ATLAS and C. Collaborations, "Combination of single top-quark cross-sections measurements in the t-channel at sqrt[s]=8 TeV with the ATLAS and CMS experiments," 2013.

[38] A. Bornheim *et al.*, "Improved measurement of |V(ub)| with inclusive semileptonic B decays," *Phys.Rev.Lett.*, vol. 88, p. 231803, 2002.

[39] B. Aubert *et al.*, "Measurement of the inclusive electron spectrum in charmless semileptonic B decays near the kinematic endpoint and determination of |V(ub)|," *Phys.Rev.*, vol. D73, p. 012006, 2006.

[40] R. Aaij *et al.*, "Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays," 2015. to appear in Nature Physics.

[41] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2 pi Decay of the k(2)0 Meson," *Phys. Rev. Lett.*, vol. 13, pp. 138–140, 1964.

[42] H. Albrecht *et al.*, "Observation of B0 - anti-B0 Mixing," *Phys.Lett.*, vol. B192, pp. 245–252, 1987.

[43] R. Aaij *et al.*, "Measurement of CP violation in $B^0 \rightarrow J/\psi K^0_S$ decays," 2015. to appear in Phys. Rev. Lett.

[44] A. J. Bevan *et al.*, "The Physics of the B Factories," *Eur. Phys. J.*, vol. C74, p. 3026, 2014.

[45] J. Charles, O. Deschamps, S. Descotes-Genon, H. Lacker, A. Menzel, *et al.*, "Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics," 2015.

[46] Y. Amhis *et al.*, "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014," 2014. updated results and plots available at TT http://www.slac.stanford.edu/xorg/hfag/.

[47] C. Di Donato, G. Ricciardi, and I. Bigi, "$\eta - \eta'$ Mixing - From electromagnetic transitions to weak decays of charm and beauty hadrons," *Phys.Rev.*, vol. D85, p. 013016, 2012.

[48] A. Bramon, R. Escribano, and M. Scadron, "Mixing of $\eta - \eta'$ mesons in J/ψ decays into a vector and a pseudoscalar meson," *Phys.Lett.*, vol. B403, pp. 339–343, 1997.

[49] T. Feldmann, P. Kroll, and B. Stech, "Mixing and decay constants of pseudoscalar mesons," *Phys.Rev.*, vol. D58, p. 114006, 1998.

[50] A. Bramon, R. Escribano, and M. Scadron, "The $\eta-\eta'$ mixing angle revisited," *Eur.Phys.J.*, vol. C7, pp. 271–278, 1999.

[51] F.-G. Cao and A. Signal, "Two analytical constraints on the $\eta - \eta'$ mixing," *Phys.Rev.*, vol. D60, p. 114012, 1999.

[52] A. Bramon, R. Escribano, and M. Scadron, "Radiative V P gamma transitions and $\eta - \eta'$ mixing," *Phys.Lett.*, vol. B503, pp. 271–276, 2001.

[53] B.-W. Xiao and B.-Q. Ma, "Photon-meson transition form-factors of light pseudoscalar mesons," *Phys.Rev.*, vol. D71, p. 014034, 2005.

[54] R. Escribano and J.-M. Frere, "Study of the $\eta - \eta'$ system in the two mixing angle scheme," *JHEP*, vol. 06, p. 029, 2005.

[55] R. Escribano, "Short study of the $\eta - \eta'$ system in the two mixing angle scheme," *PoS*, vol. HEP2005, p. 418, 2006.

[56] T. Huang and X.-G. Wu, "Determination of the $\eta - \eta'$ mixing angle from the pseudoscalar transition form factors," *Eur.Phys.J.*, vol. C50, pp. 771–779, 2007.

[57] R. Aaij *et al.*, "Evidence for the decay $B^0 \to J/\psi\omega$ and measurement of the relative branching fractions of B_s^0 meson decays to $J/\psi\eta$ and $J/\psi\eta'$," *Nucl. Phys.*, vol. B867, pp. 547–566, 2013.

[58] F. Ambrosino, A. Antonelli, M. Antonelli, F. Archilli, P. Beltrame, *et al.*, "A global fit to determine the pseudoscalar mixing angle and the gluonium content of the η' meson," *JHEP*, vol. 07, p. 105, 2009.

[59] R. Escribano and J. Nadal, "On the gluon content of the η and η' mesons," *JHEP*, vol. 05, p. 006, 2007.

[60] R. Escribano, "$J/\psi \to$ VP decays and the quark and gluon content of the eta and eta-prime," *Eur.Phys.J.*, vol. C65, pp. 467–473, 2010.

[61] B. Aubert *et al.*, "B meson decays to charmless meson pairs containing η or η' mesons," *Phys.Rev.*, vol. D80, p. 112002, 2009.

[62] D. Atwood and A. Soni, "B $\to \eta'$ + X and the QCD anomaly," *Phys.Lett.*, vol. B405, pp. 150–156, 1997.

[63] M. Ahmady, C. Kim, S. Oh, and C. Yu, "Heavy baryonic decays of $\Lambda_b \to \Lambda\eta'$," *Phys.Lett.*, vol. B598, pp. 203–210, 2004.

[64] T. Feldmann, P. Kroll, and B. Stech, "Mixing and decay constants of pseudoscalar mesons," *Phys.Rev.*, vol. D58, p. 114006, 1998.

[65] T. Feldmann, P. Kroll, and B. Stech, "Mixing and decay constants of pseudoscalar mesons: The Sequel," *Phys.Lett.*, vol. B449, pp. 339–346, 1999.

[66] Y.-H. Chen, H.-Y. Cheng, B. Tseng, and K.-C. Yang, "Charmless hadronic two-body decays of B(u) and B(d) mesons," *Phys.Rev.*, vol. D60, p. 094014, 1999.

[67] T. Mannel, W. Roberts, and Z. Ryzak, "Baryons in the heavy quark effective theory," *Nucl.Phys.*, vol. B355, pp. 38–53, 1991.

[68] C.-S. Huang and H.-G. Yan, "Exclusive rare decays of heavy baryons to light baryons: Lambda(b) $\to$ Lambda gamma and Lambda(b) $\to$ Lambda l+ l-," *Phys.Rev.*, vol. D59, p. 114022, 1999.

[69] C.-H. Chen and C. Geng, "Lepton asymmetries in heavy baryon decays of Lambda(b) $\to$ Lambda lepton+ lepton-," *Phys.Lett.*, vol. B516, pp. 327–336, 2001.

[70] A. A. Alves Jr. *et al.*, "The LHCb detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[71] R. Aaij *et al.*, "Study of beauty hadron decays into pairs of charm hadrons," *Phys. Rev. Lett.*, vol. 112, p. 202001, 2014.

[72] C. Gwenlan, "Combined HERA Deep Inelastic Scattering Data and NLO QCD Fits," *Nucl.Phys.Proc.Suppl.*, vol. 191, pp. 5–15, 2009.

[73] J. M. Campbell, J. Huston, and W. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics," *Rept.Prog.Phys.*, vol. 70, p. 89, 2007.

[74] R. Aaij *et al.*, "Performance of the LHCb Vertex Locator," *JINST*, vol. 9, p. P09007, 2014.

[75] K. Kleinkneicht, *Detectors for particle radiation.* Cambridge University Press, 1987.

[76] R. Arink *et al.*, "Performance of the LHCb Outer Tracker," *JINST*, vol. 9, p. P01002, 2014.

[77] M. Adinolfi *et al.*, "Performance of the LHCb RICH detector at the LHC," *Eur. Phys. J.*, vol. C73, p. 2431, 2013.

[78] P. Perret, "First Years of Running for the LHCb Calorimeter system," Jul 2014.

[79] R. Aaij *et al.*, "Performance of the LHCb calorimeters," 2014. in preparation.

[80] A. A. Alves Jr. *et al.*, "Performance of the LHCb muon system," *JINST*, vol. 8, p. P02022, 2013.

[81] R. Aaij *et al.*, "Measurement of the track reconstruction efficiency at LHCb," *JINST*, vol. 10, p. P02007, 2015.

[82] A. Puig, "The LHCb trigger in 2011 and 2012."

[83] R. Aaij *et al.*, "The LHCb trigger and its performance in 2011," *JINST*, vol. 8, p. P04022, 2013.

[84] V. V. Gligorov, "A single track HLT1 trigger."

[85] V. V. Gligorov and M. Williams, "Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree," *JINST*, vol. 8, p. P02013, 2013.

[86] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to PYTHIA 8.1," *Comput.Phys.Commun.*, vol. 178, pp. 852–867, 2008.

[87] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth.*, vol. A462, pp. 152–155, 2001.

[88] I. Belyaev *et al.*, "Handling of the generation of primary events in Gauss, the LHCb simulation framework," *J. Phys. Conf. Ser.*, vol. 331, p. 032047, 2011.

[89] N. Amelin, E. Staubo, L. Csernai, V. Toneev, K. Gudima, *et al.*, "Transverse flow and collectivity in ultrarelativistic heavy ion collisions," *Phys.Rev.Lett.*, vol. 67, pp. 1523–1526, 1991.

[90] N. Amelin, L. Csernai, E. Staubo, and D. Strottman, "Collectivity in ultrarelativistic heavy ion collisions," *Nucl.Phys.*, vol. A544, pp. 463–466, 1992.

[91] L. Bravina, L. Csernai, P. Levai, N. Amelin, and D. Strottman, "Fluid dynamics and quark gluon string model: What we can expect for Au + Au collisions at 11.6-A/GeV/c," *Nucl.Phys.*, vol. A566, pp. 461C–464C, 1994.

[92] L. Bravina, "Scaling violation of transverse flow in heavy ion collisions at AGS energies," *Phys.Lett.*, vol. B344, pp. 49–54, 1995.

[93] B. Andersson, G. Gustafson, and B. Nilsson-Almqvist, "A Model for Low p(t) Hadronic Reactions, with Generalizations to Hadron - Nucleus and Nucleus-Nucleus Collisions," *Nucl.Phys.*, vol. B281, p. 289, 1987.

[94] A. Heikkinen, N. Stepanov, and J. P. Wellisch, "Bertini intranuclear cascade implementation in GEANT4," *eConf*, vol. C0303241, p. MOMT008, 2003.

[95] B. Ketzer, "Physics with hadronic probes at COMPASS," *Int.J.Mod.Phys.*, vol. A24, pp. 245–254, 2009.

[96] W. D. Hulsbergen, "Decay chain fitting with a Kalman filter," *Nucl.Instrum.Meth.*, vol. A552, pp. 566–575, 2005.

[97] R. Aaij *et al.*, "Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region," *Phys. Lett.*, vol. B694, p. 209, 2010.

[98] R. Aaij *et al.*, "Production of J/ψ and Υ mesons in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP*, vol. 06, p. 064, 2013.

[99] R. Aaij *et al.*, "Study of the kinematic dependences of Λ_b^0 production in pp collisions and a measurement of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ branching fraction," *JHEP*, vol. 08, p. 143, 2014.

[100] R. Aaij *et al.*, "Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics," *JHEP*, vol. 04, p. 001, 2013.

[101] E. Govorkova, "Study of photon reconstruction efficiency using $B^+ \to J/\psi K^{(*)+}$ decays in the LHCb experiment," no. LHCb-PROC-2015-009.

[102] Y. Xie, "sFit: a method for background subtraction in maximum likelihood fit,"

[103] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Belmont, California, USA: Wadsworth international group, 1984.

[104] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, "Boosted decision trees as an alternative to artificial neural networks for particle identification," *Nucl.Instrum.Meth.*, vol. A543, pp. 577–584, 2005.

[105] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[106] R. E. Schapire and Y. Freund, "A decision-theoretic generalization of on-line learning and an application to boosting," *Jour. Comp. and Syst. Sc.*, vol. 55, p. 119, 1997.

[107] 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.

[108] 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.

[109] M. Pivk and F. R. Le Diberder, "sPlot: A statistical tool to unfold data distributions," *Nucl.Instrum.Meth.*, vol. A555, pp. 356–369, 2005.

[110] R. Aaij *et al.*, "Measurement of b hadron production fractions in 7 TeV pp collisions," *Phys. Rev.*, vol. D85, p. 032008, 2012.

[111] G. J. Feldman and R. D. Cousins, "A unified approach to the classical statistical analysis of small signals," *Phys.Rev.*, vol. D57, pp. 3873–3889, 1998.

APPENDIX

A

APPENDIX

A.1 sWeighted Data

Comparing distributions of sWeighted signal events with Monte Carlo for different variables used in training the BDT. Only events from $B^0 \to K^0_S \eta'$ ($\eta' \to \pi^+\pi^-\gamma$) decays are used with 2012 data, since this is the category with the highest statistics.

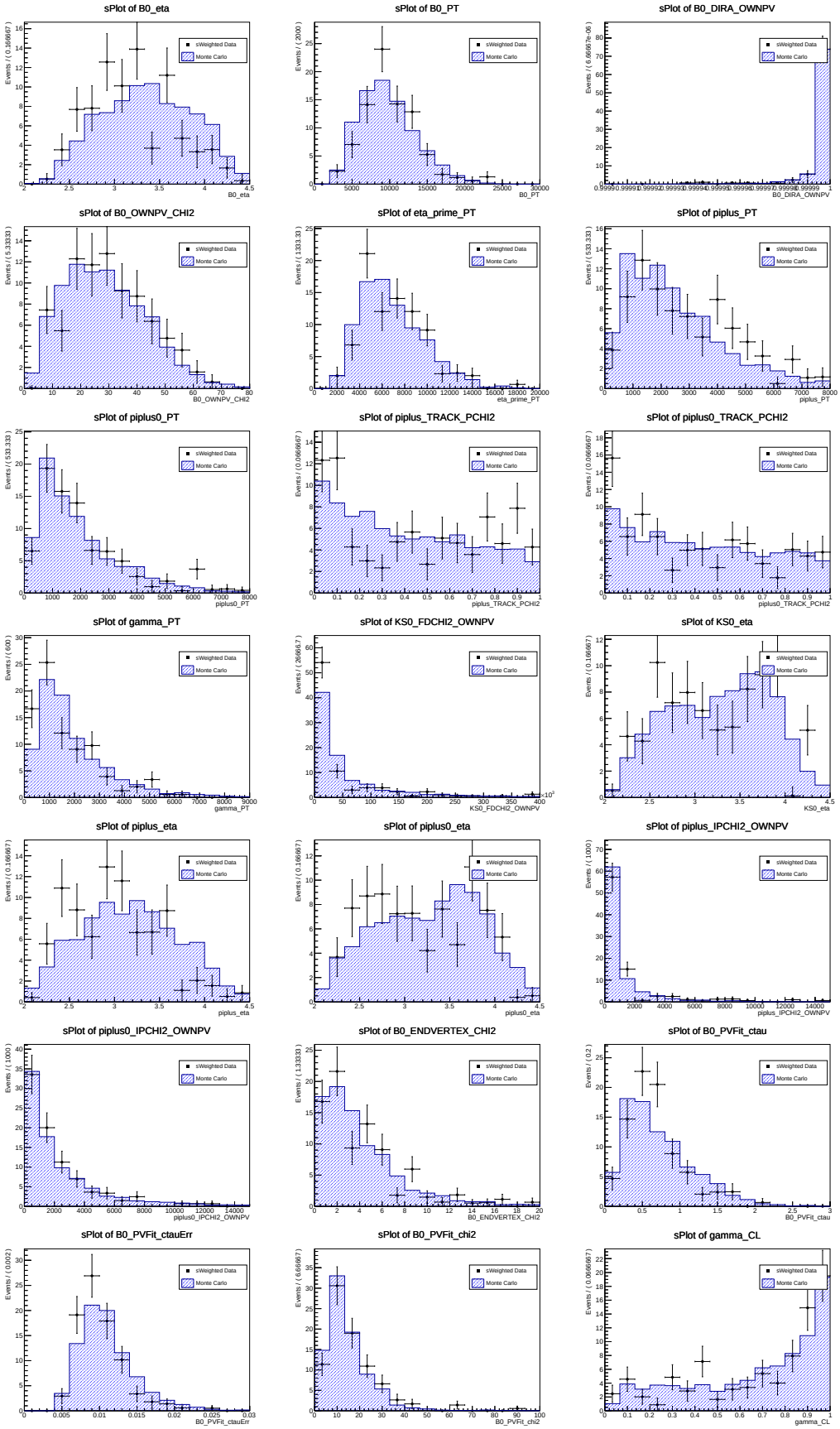

Figure A.1: sPlots for variables used in BDT training for the LL category.

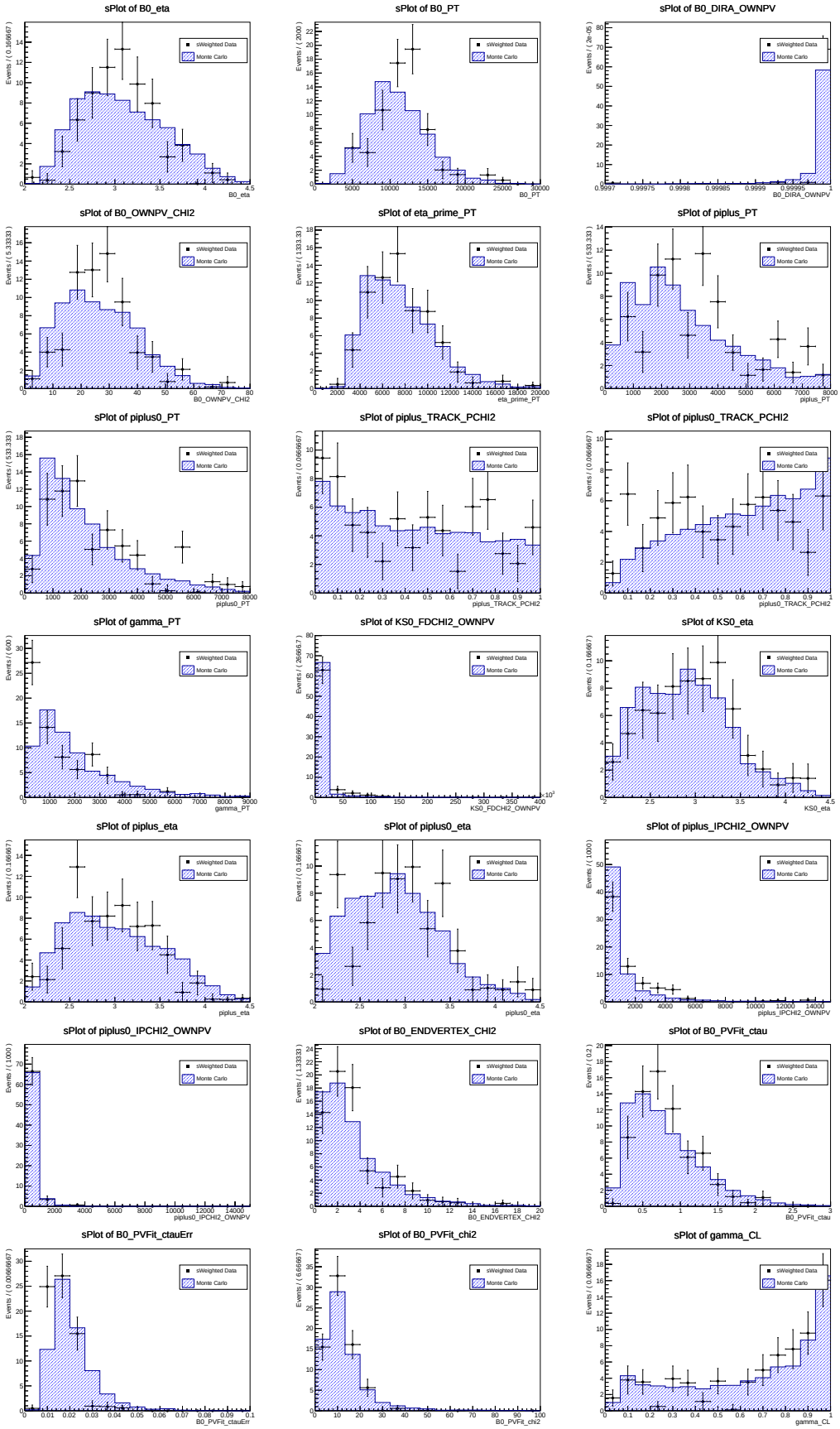

Figure A.2: sPlots for variables used in BDT training for the DD category.