First Observation of the Charmless Beauty Decay $\Lambda_b^0 \to pK\eta'$

Timothy Williams

Thesis submitted for the degree of
Doctor of Philosophy

Particle Physics Group,

School of Physics and Astronomy,

University of Birmingham.

11/05/2018

ABSTRACT

A search is performed for the decay $\Lambda_b^0 \to pK\eta'$ using pp collision data collected with the LHCb detector at centre-of-mass energies of $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. The search is performed in two decay channels; the η' is reconstructed through the decays $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$. In the $\eta' \to \pi^+\pi^-\gamma$ decay channel $117 \pm 15(\text{stat.}) \pm 10(\text{sys.})$ signal events are observed and $45 \pm 8(\text{stat.}) \pm 2(\text{sys.})$ signal events are observed in the $\eta' \to \pi^+\pi^-\eta$ decay channel. The combined statistical significance of these signals is 12.0σ, therefore this is the first observation of the decay $\Lambda_b^0 \to pK\eta'$. The branching fraction of the decay $\Lambda_b^0 \to pK\eta'$ is measured relative to the decay $B^+ \to K^+\eta'$. The ratio of branching fractions is measured to be

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\eta')}{\mathcal{B}(B^+ \to K^+\eta')} = 0.120 \pm 0.013(\text{stat.}) \pm 0.013(\text{sys.}).$$

Using the world average value for the branching fraction of the decay $B^+ \to K^+\eta'$ [1], the branching fraction of $\Lambda_b^0 \to pK\eta'$ is measured to be

$$\mathcal{B}(\Lambda_b^0 \to pK\eta') = 8.48 \pm 0.88(\text{stat.}) \pm 0.97(\text{sys.}) \times 10^{-6}.$$

This is the first observation of a b-baryon decaying to an η'.

DECLARATION OF AUTHORS CONTRIBUTION

All of the technical simulation work carried out for this thesis was done by myself. Furthermore, with the exception of the "stripping" selection (the first stage of pre-selection which is run during centralised data processing campaigns) all of the analysis work was also carried out by myself.

I have worked as part of the LHCb simulation group throughout my PhD where I have developed and maintained tests of Geant4 electromagnetic physics, which are presented in Chapter 3. I have also worked with other members of the LHCb simulation group to integrate these tests into the LHCbPR testing framework.

The analysis work was regularly presented to the charmless B decay working group during its development and many helpful suggestions were given which improved the analysis. The analysis work made extensive use of the ROOT analysis framework, including RooFit and RooStats. The TMVA toolkit was also used for the training and application of the BDTs. The analysis has been reviewed within the LHCb collaboration, both within the working group and by a formal review committee. This work is presented in Chapters 4 and 5. At the time of writing, a paper is in preparation.

I also acted as "stripping liaison" for the charmless B decay working group for a period of 15 months. This involved coordinating and testing the working groups "stripping" selections such that they can be executed during centralised data processing campaigns.

ACKNOWLEDGEMENTS

Thanks is given to the LHCb collaboration for providing me with the data and tools to carry out the work presented in this thesis. In particular, I'd like to thank the charmless B decays working group for providing a lot of useful help and suggestions to improve the analysis work. Further thanks goes to the reviewers for further improving the analysis. Thanks is also given to the simulation working group for helping carry out the technical simulation work presented in this thesis.

I'd also like to acknowledge the help provided by the Birmingham LHCb group; I have received many useful suggestions in Wednesday morning meetings. Particular thanks is given to my supervisor, Nigel Watson, for being a great supervisor. You have always been there to help and find solutions to problems as well as fighting my corner (Russian expenses claims springs to mind) and always reducing my stress levels. It's been a pleasure.

Further thanks is given to everyone I have shared West 313 with, in particular Nat for knowing everything about emacs/linux/latex/python, Jimmy for laying the foundations to my analysis and Rhys for his RooFit expertise and ever willingness to go to Yakinori. I'd also like to acknowledge the entertaining lunchtimes I have enjoyed in the ATLAS office. Discussing Dan's outdoorsy nature, throwing Matt's baterangs and snail racing to name a few.

I'd also like to thank the Birmingham Appreciation for Pedalos Society, as they are honestly some of the best friends I have ever had. I can't fully express my appreciation for the continuous support they have given me throughout my PhD, from moaning over coffee to full on crisis meetings. I have also very much enjoyed Burrito O'Clock, Staff house beers, Currys at Dilshad[1], trips to The Goose, Toby

[1]I'm not sure whether to thank Dilshad, or ask for thanks from them for the £140174091 we have spent there

Carvery and pub quiz Wednesdays, slapping various ceilings[2], drinking "beers of the month" in the "Belgian bar", three stag-dos, winning the inaugural Bubble Chambers football tournament twice, winning at FPL and many other great times. It was very comforting to know, particularly in the latter years, that whatever happened with my work I would never regret doing a PhD just because of the people I have met.

I also have a lot to thank my Grandad for. He always supported, encouraged and helped me achieve whatever I wanted. From before primary school, to during my PhD. He taught me not to be defeatist and many life lessons. It is such a shame he is no longer around to see me become "Dr. Tim". Further thanks is also given to my Mum and Nan for supporting me throughout my life to date and always encouraging me to reach my potential. I could not have achieved even a fraction of what I have without the support of my family.

On the subject of family, I'd also like to thank my mother and father-in-law Julie and Eugene Holod for welcoming me into their family, letting me live in their house and putting up with me eating all their bacon.

Last, and by no means least, many thanks are given to my wife Katy Williams, for putting up with me being grumpy when everything seemed to be going wrong (so most of my PhD), listening to me moan about physics things which make absolutely no sense to her and always doing her absolute best to cheer me up (and buying me a Playstation 4[3]). Whilst some parts of my PhD have been the toughest of my life, which I couldn't have got through without you, somehow at the same time I have shared some of the happiest times of my life with you. From shooting AK47s (and the ceiling.....not mentioning who did this) in Budapest to stepping on eggs in glorious Morecombe. For some strange reason you also agreed to marry me, which in itself deserves significant thanks! I appreciate the sacrifices you have made so that I can do physics, mainly leaving your beloved Northampton for 8 years. I'm now very much looking forward to our future together and many more wonderful times together.

[2]The tolerance of my neighbours and the landlord of Bourbon and Black, Didsbury circa Summer 2016 is also acknowledged

[3]The 1TB pro version as well!

Dedicated to George Neil Holloway
1938 - 2017

Contents

List of Tables

List of Figures

Acronyms

HLT1 High Level Trigger 1. 78

HLT2 High Level Trigger 2. 78, 89

L0 Level 0 (trigger). 77, 87, 88

`TIS` Trigger Independent of Signal. 77, 87

`TOS` Trigger on Signal. 77, 87

BDT Boosted Decision Tree. 90–96, 99–102, 121, 164, 194, 196

BNL Brookhaven National Laboratory. 2

BPV Best Primary Vertex. 79

CB Crystal Ball (function). 81, 82, 148, 153

CKM Cabibbo Kobayashi Maskawa. 1, 6–11, 13, 17

CP Charge Parity. 7, 8, 10, 11, 13, 183

CPV Charge Parity Violation. 10, 17, 24, 183

DCB Double Crystal Ball. 81, 148, 162

DIRA Direction Angle. 80, 91, 93, 94

DIS Deep Inelastic Scattering. 2

DLL Delta Log Likelihood. 44, 45

DOCA Distance of Closest Approach. 79

DTF Decay Tree Fitter. 76, 92–94, 110, 145

FCNC Flavour Changing Neutral Currents. 6, 9

GEM Gas Electron Multiplier. 46

GPD General Purpose Detectors. 27, 32

HFLAV Heavy FLavour AVeraging group. 11

IP Impact Parameter. 33, 36, 62, 65, 66, 79

KM Kobayashi Maskawa. 7

LSB Lower Sideband. 139

NLL Negative Log Likelihood. 146, 161

nTracks number of tracks in an event. 125–127, 129

OPE Operator Product Expansion. 18

PDF Probability Density Function. 145

PID Particle Identification. 44, 45, 79, 85, 86, 89, 100, 102–104, 119, 124–130, 132, 164, 182

PL Physics List. 56, 57, 59, 61, 62, 66

pp Proton-Proton. i, 27, 29, 53, 76, 77, 79

PRB Partially Reconstructed Background. 139

PV Primary Vertex. 76, 79, 80

QCD Quantum Chromodynamics. 3, 4, 17, 20, 22

RICH Ring Imaging Cherenkov Detector. 43–45, 102, 121

ROC Receiver Operator Curve. 91

SDP Square Dalitz Plot. 115–117, 197

SF Sampling Fraction. 60, 61, 67, 68

SLAC Stanford Linear Accelerator Centre. 2

SM Standard Model. 1, 3–7, 10, 12, 13, 17, 183

USB Upper Sideband. 139

VELO Vertex Locator. 34, 65

CHAPTER 1

Introduction

This chapter provides an introduction to particle physics and the standard model (SM) in Section 1.1, with a description of the Cabibbo Kobayashi Maskawa (CKM) matrix and flavour physics in Section 1.2. This is followed by a description of the theory and current status of η'– η mixing in Section 1.3. Motivation for the study of charmless beauty hadron decays and a discussion of charmless B meson decays to final states involving an $\eta^{(\prime)}$ is presented in Section 1.4. The motivation behind the search for the decay $\Lambda_b^0 \to pK\eta'$, which is the main topic presented in this thesis, is given in Section 1.5.

1.1 The Standard Model of Particle Physics

The first studies of elementary particles came with the discovery of the electron in 1897 by J.J. Thompson [2]. By applying a high voltage to a low pressure gas it was possible to measure the mass to charge ratio of the electron, m/e, which was found to be lower than any existing measurement at the time. The electron was the first example of a lepton, a fundamental particle which does not interact via the strong force.

Throughout the 20th century a whole plethora of strongly interacting particles, collectively known as hadrons, were also discovered. The existence of neutrons and protons was established by 1932 [3,4]. This was followed by the discoveries of pions, kaons and lambda baryons, amongst others, before the start of the 1960s [5–7]. In order to classify and explain the varying properties of these hadrons, the quark model was proposed by Gell-Mann and Zweig in 1964 [8,9]. This predicted the existence of three new fundamental elementary particles (quarks) and proposed the recently discovered hadrons were compositions of quarks. We now know these three particles as the up, down and strange type quarks. This theory was validated in 1969 when Deep Inelastic Scattering (DIS) experiments, at the Stanford Linear Accelerator Centre (SLAC), showed that the proton was made of more fundamental constituent point-like particles [10].

There had been speculation that there could be a fourth quark, to partner the strange quark, nearly as soon as the quark model was proposed [11]. In 1974 the J/ψ meson, a $c\bar{c}$ bound state, was simultaneously discovered at SLAC and Brookhaven National Laboratory (BNL). This confirmed the existence of charm quarks. Even before the discovery of the charm quark, the existence of a third generation of quarks was predicted by Kobayashi and Maskawa [12]. In 1977 the first $b\bar{b}$ resonance, in the form of the upsilon $\Upsilon(1S)$ particle, was discovered at the Fermi National Accelerator Laboratory. This confirmed the existence of the third generation of quarks and paved

Flavour	Mass	Charge/|e|
u (Up)	$2.3^{+0.7}_{-0.5}$ MeV	$\frac{2}{3}$
d (Down)	$4.7^{+0.5}_{-0.4}$ MeV	$-\frac{1}{3}$
c (Charm)	1.28 ± 0.03 GeV	$\frac{2}{3}$
s (Strange)	96^{+8}_{-4} MeV	$-\frac{1}{3}$
t (Top)	173.1 ± 0.6 GeV	$\frac{2}{3}$
b (Bottom)	$4.18^{+0.04}_{-0.03}$ GeV	$-\frac{1}{3}$

Table 1.1: A summary of the six quark flavours which are present in the SM, separated by their generations [1].

the way to search for the partner of the bottom quark. Due to the high mass of the top quark, 173.1 ± 0.6 GeV [1], it was not discovered until 1995 by the CDF and D0 collaborations [13, 14].

Overall, there are six known quarks, the properties of which are summarised in Table 1.1. Particles composed of three quarks are known as baryons and those composed of two are known as mesons. More recently, bound states composed of four and five quarks have been observed [15, 16]. These are commonly known as tetraquarks and pentaquarks respectively.

To accompany the electron, there are two other types of lepton, the muon and the tau. For each flavour of lepton there is also a corresponding neutrino. The electron, muon and tau leptons each carry $-1|e|$ charge whereas the neutrinos are neutral; a summary of the lepton properties is given in Table 1.2.

The best theory available to describe these fundamental particles and their interactions is the SM. The SM is a quantum field theory which arises from the product of the groups $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. The SM is also a gauge theory, which means particle interactions are mediated by the exchange of gauge bosons.

The strong force, which is described by Quantum Chromodynamics (QCD), is mediated by the exchange of gluons and is embedded in the special unitary group of

Name	**Mass**	**Charge/**|e|
Electron, e	$0.5109989461 \pm 0.0000000031\,\mathrm{MeV}$	-1
Electron Neutrino, ν_e	$< 2\,\mathrm{eV}$	0
Muon, μ^-	$105.6583745 \pm 0.0000024\,\mathrm{MeV}$	-1
Muon Neutrino, ν_μ	$< 0.19\,\mathrm{MeV}$	0
Tau, τ^-	$1776.86 \pm 0.12\,\mathrm{MeV}$	-1
Tau Neutrino, ν_τ	$< 18.2\,\mathrm{MeV}$	0

Table 1.2: A summary of the leptons present in the SM, separated by lepton flavour [1].

degree three, $\mathrm{SU(3)_C}$. The strong force acts on colour charge, which can take one of the values R,G or B for quarks and $\bar{\mathrm{R}}$, $\bar{\mathrm{B}}$ and $\bar{\mathrm{G}}$ for anti-quarks. Gluons carry both colour and anti-colour. The group $\mathrm{SU(3)_C}$ is made of $3^2 - 1 = 8$ linearly independent hermitian 3×3 matrices, which are known as the Gell-Mann matrices. These describe the colour quantum numbers of gluons.

As the gluons themselves carry colour charge, self-interactions are possible. Consequently, at long distances (low energies) the potential between two quarks is linearly dependent on separation distance, meaning an infinite amount of energy would be required for complete separation. This gives rise to colour confinement; quarks and gluons are confined inside the QCD potential, meaning only colour-less objects can be observed as free particles under normal conditions.

The product of the groups $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ describes the electroweak interactions, which is the combined theory of electromagnetic and weak interactions. The gauge bosons of the group $\mathrm{SU(2)_L}$, and mediator of the weak force, are the $W_{(1,2,3)}$ massless vector boson fields. These couple to weak isospin T with coupling strength g_W. The gauge boson of the group $\mathrm{U(1)_Y}$, B, couples to hypercharge Y with strength g'_W, and is also a massless vector boson field. The weak isospin of left handed particles is 1/2 and the weak isospin of right handed particles[1] is 0. Each particle also has

[1]Here handedness is defined as chirality

a third component of weak isospin, T_3, which is $-1/2$ for down type quarks and leptons but $+1/2$ for up type quarks and neutrinos. The hypercharge of a particle is then given by

$$Y = 2(Q - \mathrm{T}_3), \tag{1.1}$$

where Q is the electric charge of the particle.

The observed masses of the weak vector bosons as well as the massless photon arise due to the Higgs mechanism [17, 18]. The introduction of a complex scalar doublet field, which has a non-zero vacuum expectation value, causes a spontaneous symmetry breaking. This leads to the appearance of the $W^{\pm}$, Z and γ as linear combinations of the $W_{(1,2,3)}$ and B fields. Explicitly these are given by:

$$\begin{aligned} \begin{pmatrix} \gamma \\ Z \end{pmatrix} &= \begin{pmatrix} \cos\theta_{\mathrm{W}} & \sin\theta_{\mathrm{W}} \\ -\sin\theta_{\mathrm{W}} & \cos\theta_{\mathrm{W}} \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}, \\ W^{\pm} &= \frac{1}{\sqrt{2}}(W_1 \pm iW_2), \end{aligned} \tag{1.2}$$

where θ_{W} is the Weinberg angle, which is given at lowest order by the relations,

$$\cos\theta_{\mathrm{W}} = \frac{g_{\mathrm{W}}}{\sqrt{g_{\mathrm{W}}^2 + g_{\mathrm{W}}'^2}} = \frac{m_{\mathrm{W}}}{m_{\mathrm{Z}}}, \tag{1.3}$$

where m_{W} and m_{Z} are the masses of the W and Z bosons respectively.

The Higgs mechanism also predicts the existence of a new scalar boson, the Higgs boson. Unfortunately, the theory does not predict the mass of this boson. However after over 50 years of searches it was discovered by the ATLAS and CMS experiments in 2012 [19, 20]. The most recent mass measurement gives $m_{\mathrm{H}} = 125.09 \pm 0.21(stat.) \pm 0.11(syst.)\,\mathrm{GeV}$ [21]. With the discovery of the Higgs boson the SM was completed.

1.2 Flavour Physics and the CKM Matrix

In the SM the only interaction that can change quark flavour is the weak interaction, mediated by a $W^{\pm}$ boson. Flavour Changing Neutral Currents (FCNC), where the weak decay is mediated by a Z boson, are forbidden at tree level in the SM. Examples of flavour changing weak decay vertices are shown in Figure 1.1. Two decays which involve these vertices are the decays $K^+ \to \mu^+\nu_\mu$ and $\pi^+ \to \mu^+\nu_\mu$. It was observed experimentally that the branching fraction of the K^+ decay, relative to the π^+ decay, did not match theoretical predictions. This suggested that the coupling of the $W^{\pm}$ may not be independent of quark flavour. Therefore in 1963 Cabibbo proposed that the weak eigenstates of the down type quarks are admixtures of the mass eigenstates, such that

$$\begin{pmatrix} d_{\mathrm{W}} \\ s_{\mathrm{W}} \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ \text{-}\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{1.4}$$

where θ_c is the Cabibbo angle and $d_{\mathrm{W}}, s_{\mathrm{W}}$ are the weak eigenstates [22]. The Cabibbo angle has been experimentally measured to be $\theta_c \approx 13^\circ$. The weak eigenstates being misaligned to the mass eigenstates leads to the coupling of the W boson, in the case of the K^+ decay, being proportional to $\sin\theta_c$. In the case of the π^+ decay the coupling is proportional to $\cos\theta_c$, which accounts for the differences between the observed branching fractions and predictions which assume the coupling of the $W^{\pm}$ is independent of quark flavour.

This formalism was extended to include three generations of quarks with the use of a 3×3 matrix, the Cabibbo Kobayashi Maskawa (CKM) matrix [23]. With all three generations of quarks included, the weak eigenstates of the down type quarks

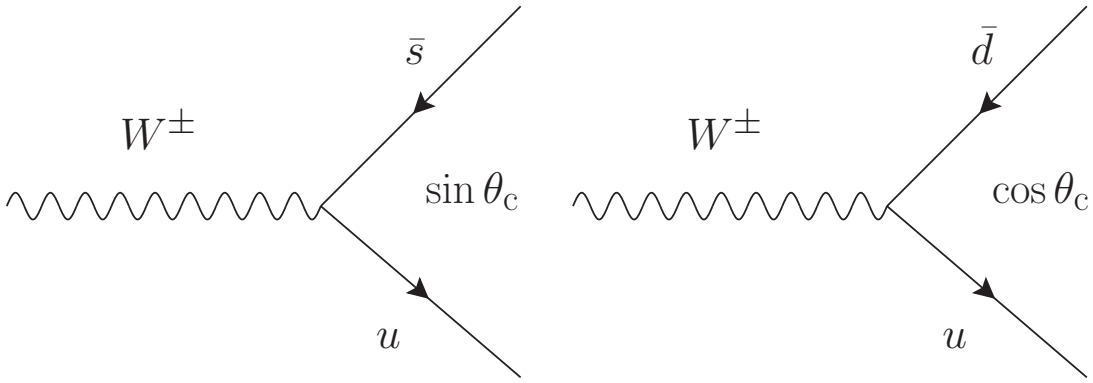


Figure 1.1: Examples of weak decay vertices which involve a change of quark flavour.

are given by

$$\begin{pmatrix} d_{\mathrm{W}} \\ s_{\mathrm{W}} \\ b_{\mathrm{W}} \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.5}$$

$$\begin{pmatrix} d_{\mathrm{W}} \\ s_{\mathrm{W}} \\ b_{\mathrm{W}} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix},$$

where V_{CKM} is the CKM matrix. The CKM matrix can be parameterised by three angles and one complex phase such that

$$V_{\mathrm{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.6}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ and θ_{ij} is the mixing angle between the generation i and j, and δ is the Kobayashi Maskawa (KM) phase. The KM phase accommodates the violation of Charge-Parity (CP) symmetry in SM weak decays. When CP is

conserved, a process is invariant under the combined operations of parity and charge conjugation. However weak decays violate CP symmetry, as first observed in neutral kaon decays in 1964 [24].

Another common way to parameterise the CKM matrix is the Wolfenstein parameterisation [25]. By defining:

$$\lambda = s_{12} = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \tag{1.7}$$

$$A\lambda^2 = s_{23} = \lambda|\frac{V_{cb}}{V_{us}}|, \tag{1.8}$$

$$A\lambda^3(\rho + i\eta) = s_{13}e^{i\delta} = V_{ub}^*, \tag{1.9}$$

the CKM matrix can be defined as

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ \lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right). \tag{1.10}$$

The Wolfenstein parameterisation is useful for showing the hierarchy of the different matrix elements.

All of the values of the elements of the CKM matrix have to be determined experimentally and the current world averages[2] are [1]

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{pmatrix}. \tag{1.11}$$

These values show that the CKM matrix is almost a diagonal matrix; the diagonal

[2]It should be noted that at the time of writing the PDG average for V_{ub}, which is quoted here, does not include a recent measurement by LHCb using $\Lambda_b^0 \to p\mu^-\nu_\mu$ decays which is of comparable precision to the world average [26].

elements are all close to or compatible with unity. This means that vertices, and consequently decays, which involve a change of quark flavour between generations are heavily suppressed. This makes the study of b hadron decays to charmless final states experimentally challenging.

The CKM matrix is unitary, otherwise the sum of interaction probabilities would not be conserved. This imposes the relationships:

$$\sum_i V_{ij} V_{ik}^* = \delta_{jk}, \tag{1.12}$$

and

$$\sum_j = V_{ij} V_{kj}^* = \delta_{ik} \tag{1.13}$$

between the matrix elements. Explicitly, this results in the relations:

$$\begin{aligned}
|V_{ud}|^2 + |V_{cd}|^2 + |V_{td}|^2 &= 1, \\
|V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 &= 1, \\
|V_{ub}|^2 + |V_{cb}|^2 + |V_{tb}|^2 &= 1, \\
|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 &= 1, \\
|V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 &= 1, \\
|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 &= 1.
\end{aligned} \tag{1.14}$$

There are also six vanishing relations, of which the most useful for experimental measurements is

$$V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0. \tag{1.15}$$

Physically, this can be interpreted as imposing the non-existence of FCNC. As these relationships are three complex numbers summing to zero, they can be represented in the complex plane as so-called “unitary triangles”. The most commonly used unitary triangle arises from dividing Equation. (1.15) by the most precisely measured

quantity, $V_{cb}^* V_{cd}$, which gives the relationship,

$$\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}} + 1 + \frac{V_{tb}^* V_{td}}{V_{cb}^* V_{cd}} = 0. \tag{1.16}$$

This results in the unitary triangle shown in Figure 1.2, where

$$\overline{\rho} = \rho\left(1 - \frac{1}{2}\lambda^2\right) + \mathcal{O}(\lambda^4) \quad \text{and} \quad \overline{\eta} = \eta\left(1 - \frac{1}{2}\lambda^2\right) + \mathcal{O}(\lambda^4). \tag{1.17}$$

In the SM the unitary triangle is a closed triangle, the sum of the three angles is 180°. However in several extensions to the SM, such as those that introduce a 4th generation of quarks, the triangle does not close. It is therefore of interest to measure precisely the three angles of the triangle and the lengths of each side in order to test the SM. The lengths of the sides are determined by measurements of the CKM elements. The three angles have to be determined by measuring Charge Parity Violation (CPV) parameters and they are given by:

$$\alpha = \arg\left(-\frac{V_{td} V_{tb}^*}{V_{ud} V_{ub}^*}\right), \tag{1.18}$$

$$\beta = \arg\left(-\frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*}\right), \tag{1.19}$$

$$\gamma = \arg\left(-\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*}\right). \tag{1.20}$$

The angle α has been determined by measuring time-dependent CP asymmetries in $b \to u\bar{u}d$ decay modes. These measurements have predominantly been made using the decay modes $B \to \pi\pi$, $B \to \rho\pi$ and $B \to \rho\rho$ [27, 28]. The theoretically cleanest way to determine the angle β is through measurements of time-dependent CP violation in the interference between B^0 decays with and without B^0-$\bar{B}^0$ mixing, in the channel $b \to c\bar{c}s$. The most precise results for β are achieved using the $B^0 \to J/\psi\, K_S^0$ decay mode [29].

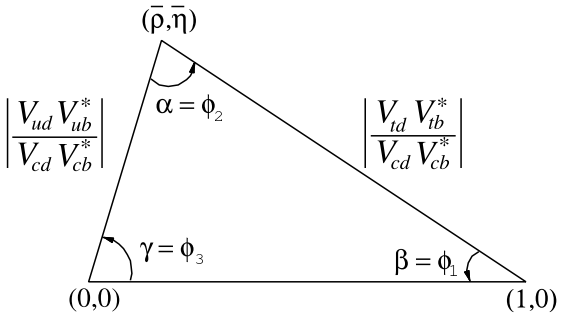


Figure 1.2: The unitary triangle most commonly used to illustrate the CKM matrix [1].

The angle γ is the only phase which does not depend on CKM elements involving the top quark. Therefore, it can be determined from tree level B meson decays. Commonly, this measurement is made using the decay channels $B^+ \to D^0 K^+$ and $B^+ \to \overline{D}^0 K^+$, where both the $\overline{D}^0$ and the D^0 decay to the same final state, such as $D^0 \to K^+ K^-$. The products of weak vertex factors in these decays are $V_{ub}^* V_{cs}$ and $V_{cb}^* V_{us}$ respectively, which are both of the order λ^3. Therefore both of these decays have similar amplitudes, which means there is significant interference between them, giving sensitivity to the CKM angle γ. By measuring the time-independent CP asymmetry for these decays a measurement of γ can be made [30]. The current world averages for direct measurements of the CKM angles, as calculated by the Heavy FLavour AVeraging group (HFLAV) [31], are:

$$\alpha_{\text{dir}} = (84.9^{+5.1}_{-4.5})^\circ,$$

$$\beta_{\text{dir}} = (22.2 \pm 0.7)^\circ,$$

$$\gamma_{\text{dir}} = (73.5^{+4.2}_{-5.1})^\circ.$$

Global fits are performed to the unitary triangle by the CKMFitter group [32], using a frequentist approach. These fits take many experimental measurements as input parameters: direct measurements of the CKM matrix elements; the B^0 and B_s^0 mixing parameters Δm_d and Δm_s; the K^0 CP violation parameter ϵ_{K}; constraints

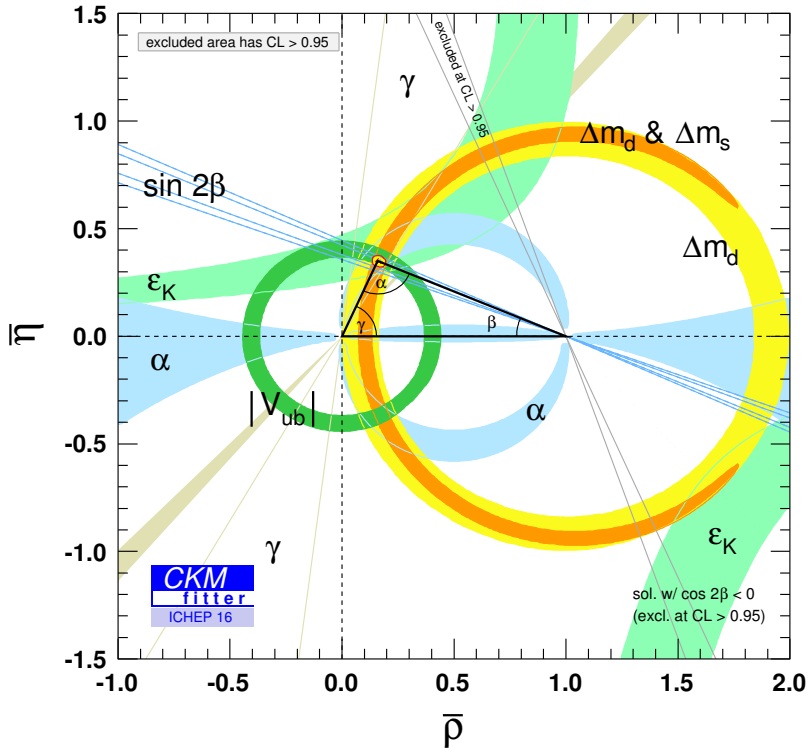


Figure 1.3: The results of a global fit of the unitary triangle, as performed by the CKMFitter group [32]

from direct measurements of the angles α, β and γ; and several branching fraction measurements which provide additional sensitivity. The result of the fit performed by the CKM Fitter group, in the $(\overline{\rho}, \overline{\eta})$ plane, can be seen in Figure 1.3. The best fit values for the angles are [32]:

$$\alpha = (92.0^{+1.3}_{-1.1})^\circ,$$

$$\beta = (22.6^{+0.36}_{-0.35})^\circ,$$

$$\gamma = (65.4^{+0.97}_{-1.16})^\circ,$$

which are consistent with the direct measurements. Furthermore the sum of the angles from the direct measurements are consistent with 180°, therefore no deviations from the SM have been observed.

However, as more data is collected by the LHCb experiment the uncertainties on the the unitary triangle angles will decrease and the SM will be tested more rigorously. Until recent measurements by LHCb, γ was the least constrained angle of the unitary triangle; the precision of the current world average is driven by the most recent LHCb measurement, $\gamma = (76.8^{+5.1}_{-5.7})^\circ$ [33]. It is the physics goal of the LHCb experiment to measure γ to degree level precision.

Another quantity which can be determined from the global fit is the total amount of CP violation in the SM. This is quantified by the Jarlskog invariant which is given by

$$\mathcal{J} = \sum_{m,n=1}^{3} \varepsilon_{ikm}\varepsilon_{jln}\mathrm{Im}\left(V_{ij}V_{kl}V_{il}^{*}V_{kj}^{*}\right), \tag{1.21}$$

where $V_{ij}, V_{kl}, V_{il}^{*}, V_{kj}^{*}$ are the CKM matrix elements and ε_{ikm} is the Levi-Civita tensor[3]. An example of one possible term is $\mathrm{Im}\left(V_{ud}V_{cs}V_{us}^{*}V_{cd}^{*}\right)$. The Jarlskog invariant is equal to twice the area of the unitary triangle, and it is the same for every possible phase convention. The global fit performed by CKMFitter also extracts the Jarlskog invariant and the best fit value is $\mathcal{J} = (3.099^{+0.052}_{-0.063}) \times 10^{-5}$ [32]. This is consistent with the standard model prediction, but it is also a factor of 10^9 too small to explain the matter-antimatter imbalance observed in the universe. Therefore, the level of CP violation seen in the SM is not sufficient and other undiscovered sources must exist.

1.3 η'– η Mixing

The η' and η particles are peculiar. They are light neutral mesons which are members of the pseudoscalar nonet and they break $\mathrm{SU(3)_C}$ flavour symmetry. Both the composition of the $\eta^{(\prime)}$ wave functions and the large mass difference between the η and η' ($M(\eta) = 547.862 \pm 0.017\,\mathrm{MeV}$, $M(\eta') = 957.78 \pm 0.06\,\mathrm{MeV}$) are long stand-

[3] Also known as the totally antisymmetric tensor

ing puzzles [1, 34, 35]. The physical η' and η states can be represented as linear combinations of the $SU(3)_C$ singlet (η_0) and octet(η_8) states such that

$$\begin{pmatrix} |\eta\rangle \\ |\eta'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta_p & -\sin\theta_p \\ \sin\theta_p & \cos\theta_p \end{pmatrix} \begin{pmatrix} |\eta_8\rangle \\ |\eta_0\rangle \end{pmatrix}, \tag{1.22}$$

where the singlet state is given by

$$|\eta_0\rangle = \frac{1}{\sqrt{3}}|u\bar{u} + d\bar{d} + s\bar{s}\rangle \tag{1.23}$$

and octet state is given by

$$|\eta_8\rangle = \frac{1}{\sqrt{6}}|u\bar{u} + d\bar{d} - 2s\bar{s}\rangle. \tag{1.24}$$

The rotation angle, θ_p, is the η'– η mixing angle. An often more convenient basis for the $\eta^{(\prime)}$ wave functions is the quark flavour basis [36]. In this basis

$$\begin{pmatrix} |\eta\rangle \\ |\eta'\rangle \end{pmatrix} = \begin{pmatrix} \cos\phi_p & -\sin\phi_p \\ \sin\phi_p & \cos\phi_p \end{pmatrix} \begin{pmatrix} |\eta_q\rangle \\ |\eta_s\rangle \end{pmatrix}, \tag{1.25}$$

where

$$|\eta_q\rangle = \frac{1}{\sqrt{2}}|u\bar{u} + d\bar{d}\rangle \tag{1.26}$$

and

$$|\eta_s\rangle = |s\bar{s}\rangle. \tag{1.27}$$

The alternative η'– η mixing angle, ϕ_p, is given by,

$$\phi_p = \theta_p + \arctan\sqrt{2} \simeq \theta_p + 54.7^\circ. \tag{1.28}$$

In theory the η' and η wave functions could also contain contributions from $|c\bar{c}\rangle$ and $|b\bar{b}\rangle$ states, but the large masses of these quarks would make them highly suppressed.

However, the massless nature of the gluon means a gluonic component, $|gg\rangle$, to the wave function is not unreasonable. In fact, the size of the $|gg\rangle$ component in the η' wave function is a long debated and unanswered puzzle [37, 38]. A large $|gg\rangle$ component is often proposed as a solution to the problem of the surprisingly large η' mass [39, 40].

It is assumed that the η meson has a zero or negligible $|gg\rangle$ component due to its much smaller mass. The $|gg\rangle$ component can be introduced to the η' wave function with the addition of the gluonic mixing angle ϕ_{G}, which allows the η' wave function to be given by

$$|\eta'\rangle \simeq \cos\phi_{\mathrm{G}} \sin\phi_{\mathrm{p}} |\eta_{\mathrm{q}}\rangle + \cos\phi_{\mathrm{G}} \cos\phi_{\mathrm{p}} |\eta_{\mathrm{s}}\rangle + \sin\phi_{\mathrm{G}} |gg\rangle, \tag{1.29}$$

and the η wave function remains as

$$|\eta\rangle \simeq \cos\phi_{\mathrm{p}} |\eta_{\mathrm{q}}\rangle - \sin\phi_{\mathrm{p}} |\eta_{\mathrm{s}}\rangle. \tag{1.30}$$

Theoretical calculations of the $\eta^{(\prime)}$ mixing angles are challenging. However, lattice QCD has been used to make predictions of the η'– η mixing angle; a calculation giving the result $\theta_{\mathrm{p}} = -14.1 \pm 2.8^{\circ}$ is reported in Ref. [41] and a value of $\phi_{\mathrm{p}} = 42 \pm 1^{\circ}$ is reported in Ref. [42]. These calculations are consistent given Equation. (1.28).

Experimental measurements of ϕ_{p} and ϕ_{G} have been made by the LHCb collaboration using $B^0_{(\mathrm{s})} \to J/\psi\,\eta^{(\prime)}$ decays. The mixing angles are obtained from the expressions:

$$\tan^4\phi_{\mathrm{p}} = \frac{R}{R_{\mathrm{s}}}, \quad \cot^4\phi_{\mathrm{G}} = R R_{\mathrm{s}}, \tag{1.31}$$

where,

$$R_{(\mathrm{s})} = \left(\frac{\Phi^{\eta}_{(\mathrm{s})}}{\Phi^{\eta'}_{(\mathrm{s})}}\right)^3 \frac{\mathcal{B}(B^0_{(\mathrm{s})} \to J/\psi\,\eta')}{\mathcal{B}(B^0_{(\mathrm{s})} \to J/\psi\,\eta)} \tag{1.32}$$

and $\Phi^{\eta^{(\prime)}}_{(\mathrm{s})}$ are the phase space factors for the $B^0_{(\mathrm{s})} \to J/\psi\,\eta^{(\prime)}$ decays [43]. Us-

Reference	$\phi_{\rm p}$	$\phi_{\rm G}$	Measurements Used
A. Bramon et al. [46]	$(37.8 \pm 1.7)^\circ$	assumed 0°	$\mathcal{B}\,(J/\psi \to PV)$
R. Escribano et al. [47]	$(44.6 \pm 4.1)^\circ$	$(32^{+11}_{-22})^\circ$	$\mathcal{B}\,(J/\psi \to PV)$
A. Bramon et al. [48]	$(37.7 \pm 2.4)^\circ$	assumed 0°	$\mathcal{B}\,(V \to P\gamma)$
R. Escribano et al. [49]	$(41.4 \pm 1.3)^\circ$	$(12 \pm 13)^\circ$	$\mathcal{B}\,(V \to P\gamma)$, $\mathcal{B}\,(P \to V\gamma)$
F. Ambrosino et al. [50]	$(40.5 \pm 0.6)^\circ$	$(20.3 \pm 3.5)^\circ$	$\mathcal{B}\,(V \to P\gamma)$, $\mathcal{B}\,(P \to V\gamma)$
V. Anisovich et al. [51]	$(37.7 \pm 2.6)^\circ$	$16.4^\circ \leq \phi_{\rm G} \leq 20.3^\circ$	$\mathcal{B}\,(D_s^+ \to \eta^{(\prime)}\ell\nu)$, $\eta^{(\prime)} \to \gamma\gamma^*$
F. Cao et al. [52]	$(39.8 \pm 1.8)^\circ$	assumed 0°	$\eta^{(\prime)} \to \gamma\gamma^*$
T. Feldman et. al. [53]	$(39.3 \pm 1.0)^\circ$	assumed 0°	Many (see reference)
A. Bramon et. al. [54]	$(39.2 \pm 1.3)^\circ$	assumed 0°	Many (see reference)

Table 1.3: A summary of selected phenomenological results for the two η' mixing angles, $\phi_{\rm p}$ and $\phi_{\rm G}$. V denotes a vector meson and P denotes a pseudoscalar meson.

ing $3\,\text{fb}^{-1}$ of proton-proton collision data, the results are $\phi_{\rm p} = (43.5^{+1.4}_{-2.8})^\circ$ and $\phi_{\rm G} = (0 \pm 24.6)^\circ$ [44]. This result for $\phi_{\rm p}$ is consistent with the theoretical predictions from lattice QCD. The value for $\phi_{\rm G}$ is consistent with no $|gg\rangle$ component to the η' wave function, but it is not excluded.

Using $B^0 \to J/\psi\,\eta^{(\prime)}$ decays, the Belle collaboration has also set a limit on $\phi_{\rm p}$; $\phi_{\rm p} < 42.2^\circ$ at the 90% confidence level [45]. This result is consistent with the LHCb measurement of $\phi_{\rm p}$, but it should be noted that the Belle result neglects phase space factors and assumes the $|gg\rangle$ contribution to the η' wavefunction is negligible.

Many phenomenological results have been reported, where other experimental measurements have been reinterpreted to determine the η'– η mixing angles. A selection of these results are summarised in Table 1.3.

Overall there is agreement between theory, experiment and phenomenological results that $\phi_{\rm p} \approx 40^\circ$, with all results falling in the range $37.7^\circ - 44.6^\circ$. This indicates strong mixing between the η and η'. However, for the case of $\phi_{\rm G}$ there is no such agreement; some phenomenological results suggest that $\phi_{\rm G}$ is consistent with zero whilst others report results which are inconsistent with no $|gg\rangle$ component to the η' wavefunction. Interestingly, the $\phi_{\rm G}$ result which reports the strongest evidence for a $|gg\rangle$ contribution is a matter of contention; the authors of Ref. [49] disagree

with the result presented in Ref. [50] because of two additional parameters included to account for the overlap of the vector and pseudoscalar meson wavefunctions in $P \to V\gamma$ and $V \to P\gamma$ decays. Considering all the results available, there is a suggestion that a small $|gg\rangle$ component to the η' wavefunction could be present. However, the available data are not sufficiently precise to clarify this and therefore further experimental input is required.

1.4 Charmless b Hadron Decays

Charmless b-hadron decays provide fertile ground for searching for physics beyond the standard model. These decays can typically proceed via either $b \to s, d$ gluonic loop level diagrams (gluonic penguins) or $b \to u$ tree level diagrams. The tree level diagrams are heavily suppressed by a factor of V_{ub}, which is the smallest element of the CKM matrix (see Equation. (1.11)). Consequently, the tree and loop level diagrams often have similar amplitudes. This means that the rates of these processes are sensitive to new physics, entering via the loop diagrams, which can be indirectly detected by measuring branching fractions.

Also, the interference between the tree and loop level diagrams provides sensitivity to CPV effects. As we know that the amount of CPV in the SM is insufficient to explain the observed antimatter-matter imbalance in the universe, new physics may also introduce additional sources of CPV. Therefore, measuring CPV observables and making comparisons to SM predictions provides another way to infer the presence of new physics. Furthermore, the study of charmless b-hadron decays can greatly improve our knowledge of low energy QCD. This, in turn, improves the predictions of SM values and makes distinguishing new physics from SM processes possible.

The study of charmless b hadron decays to final states involving an $\eta^{(\prime)}$ particle has greatly added to our understanding of non-perturbative QCD whilst also intro-

ducing some unexpected puzzles. For example the branching fraction of the decay $B^+ \to K^+\eta'$ is unexpectedly large [1],

$$\mathcal{B}(B^+ \to K^+\eta') = (7.06 \pm 0.25) \times 10^{-5}, \tag{1.33}$$

both compared to other charmless B meson decays and relative to the corresponding η decay mode where

$$\mathcal{B}(B^+ \to K^+\eta) = (2.4 \pm 0.4) \times 10^{-6}. \tag{1.34}$$

These decays can proceed through both gluonic penguin diagrams and tree level diagrams, as shown in Figure 1.4. The same pattern of suppressed branching fractions for the η mode is seen with neutral B meson decays [1] where

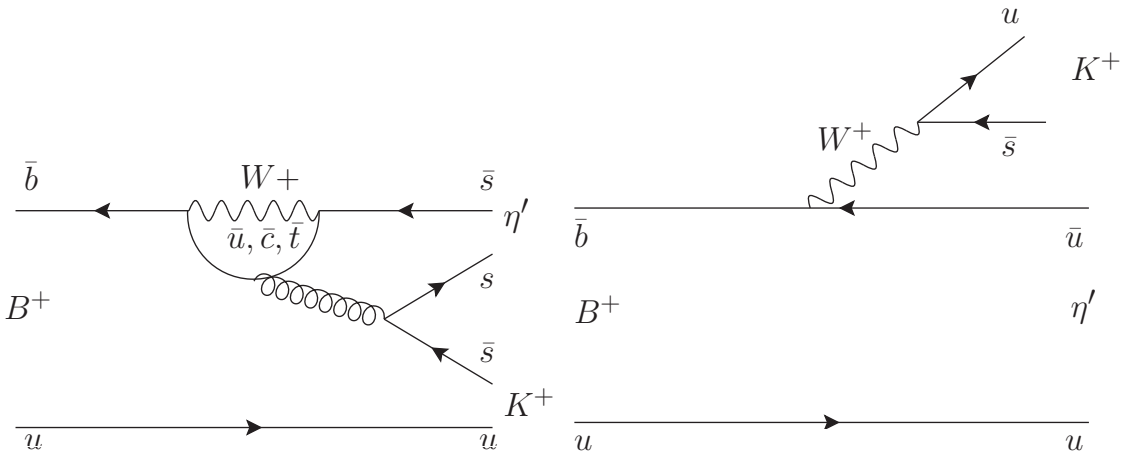
$$\mathcal{B}(B^0 \to K^0_S\eta') = (6.6 \pm 0.4) \times 10^{-5} \gg \mathcal{B}(B^0 \to K^0_S\eta) = 1.23^{+0.27}_{-0.24} \times 10^{-6}. \tag{1.35}$$

Many explanations for this have been proposed [55], one of which is the presence of a gluon component in the η' wavefunction, which is discussed in Section 1.3. This would allow the decay $B^+ \to K^+\eta'$ to proceed through additional Feynman diagrams such as that shown in Figure 1.4c, which could enhance the decay rate to η' mesons [56, 57].

In order to calculate the decay rate of a general b-hadron decay involving a $b \to s$ transition, it is common to use the Operator Product Expansion (OPE) approach. In this approach the amplitude of a two body decay is given by

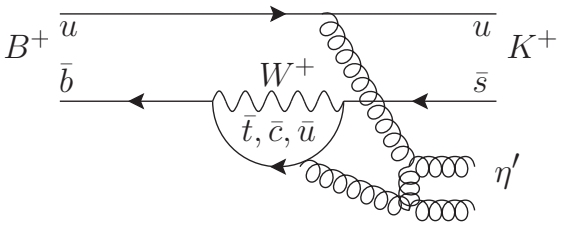
$$A(B \to P_1P_2) = \langle P_1P_2|\mathcal{H}_{\text{eff}}|B\rangle, \tag{1.36}$$

where $\mathcal{H}_{\text{eff}}$ is an effective Hamiltonian. For the decay $B^+ \to K^+\eta'$, the effective



(a) Gluonic loop level diagram

(b) Tree level diagram.



(c) The spectator quark radiates a gluon and the η' is formed through the gluonic component to the wavefunction.

Figure 1.4: Allowed Feynman diagrams for the decay $B^+ \to K^+\eta'$.

Hamiltonian is given by

$$\mathcal{H}_{\text{eff}} = \frac{\mathrm{G_F}}{\sqrt{2}} \left[V_{ub}V_{us}^* \left(c_1 \mathcal{O}_1^u + c_2 \mathcal{O}_2^u \right) + V_{cb}V_{cs}^* \left(c_1 \mathcal{O}_1^c + c_2 \mathcal{O}_2^c \right) - V_{tb}V_{ts}^* \sum_{i=3}^{10} c_i \mathcal{O}_i \right] + h.c., \tag{1.37}$$

where c_i are the Wilson coefficients and $\mathcal{O}_i$ are operators [55]. The Wilson coefficients describe the perturbative short distance effects above a given energy scale, which in B decays is usually chosen to be m_b. As new physics is expected to be present at high energies, any new physics effects are likely to appear in the Wilson coefficients.

The operators, on the other hand, describe long distance non-perturbative strong interaction effects. More specifically: the tree level $b \to u$ transitions are described by $\mathcal{O}_1$ and $\mathcal{O}_2$; gluonic loop level diagrams are described by the operators $\mathcal{O}_{3-6}$; and electroweak penguin diagrams are described by the operators $\mathcal{O}_{7-10}$.

The challenging part of this approach is calculating the hadronic matrix elements, $\langle K^+\eta'|\mathcal{O}_i|B^+\rangle$. One method for doing so is the factorisation approach, within which the hadronic matrix elements can be expressed as

$$\begin{aligned} \langle K^+\eta'|\mathcal{O}_i|B^+\rangle &= \langle \eta'|\bar{q}\gamma_\mu\gamma_5 q|0\rangle \langle K^+|\bar{s}\gamma_\mu \left(1-\gamma_5\right) b|B^+\rangle \\ &= i f_{\eta'}^q \left(m_B^2 - m_K^2\right) F_0^{BK}\left(m_{\eta'}^2\right), \end{aligned} \tag{1.38}$$

or

$$\begin{aligned} \langle K^+\eta'|\mathcal{O}_i|B^+\rangle &= \langle K^+|\bar{s}\gamma_\mu\gamma_5 u|0\rangle \langle \eta'|\bar{u}\gamma_\mu \left(1-\gamma_5\right) b|B^+\rangle \\ &= i f_K \left(m_B^2 - m_{\eta'}^2\right) F_0^{B\eta'}\left(m_K^2\right) \end{aligned} \tag{1.39}$$

depending on which operator is present in the matrix element [55]. $F_0^{BK}, F_0^{B\eta'}$ are hadronic form factors, which can be calculated using QCD sum rules on the light-cone [58–61] or perturbative QCD [62]. The decay constants, $f_{\eta'}, f_K$, have been measured experimentally; for kaons f_K has been determined to be $f_K = 159.8 \pm 1.4\,\text{MeV}$ [1]. The η' decay constants are less straightforward due to η'– η mixing. In

the quark flavour basis they are given by:

$$\begin{aligned} f_{\eta'}^{u} &= f_u \sin\phi_{\rm p}, \\ f_{\eta'}^{s} &= f_s \cos\phi_{\rm p}, \end{aligned} \tag{1.40}$$

where f_u and f_s have been determined phenomenologically to take the values $f_u = 139.8 \pm 2.6\,\text{MeV}$ and $f_s = 175 \pm 8\,\text{MeV}$ [53]. The mixing angle, $\phi_{\rm p}$, has been measured by LHCb to have the value $\phi_{\rm p} = (43.5^{+1.4}_{-2.8})^\circ$ [44] and many phenomenological studies are summarised in Table 1.3. Using these values the decay amplitude, and subsequently the branching fraction, can be calculated.

1.5 The Decay $\Lambda_b^0 \to pK\eta'$

The Λ_b^0 baryon is the lightest baryon containing a bottom quark in the standard model, with a quark content of (udb) and a mass of $5619.60 \pm 0.17\,\text{MeV}$ [1]. Before the advent of the LHC opportunities to study Λ_b^0 baryon decays were very limited. The only colliders which produced Λ_b^0 baryons were the Tevatron and LEP, but in both cases the samples of Λ_b^0 baryons available were limited in statistics. At the LHC copious amounts of $b\bar{b}$ pairs are produced[4], and approximately 15% hadronise to form a Λ_b^0 baryon [64]. This means the LHCb experiment has the first opportunity to study precisely the properties of charmless Λ_b^0 decays.

The decay of a beauty baryon to a final state involving an $\eta^{(\prime)}$ has never been observed, making this a completely unexplored area of charmless b physics. A search for the decays $\Lambda_b^0 \to \Lambda\eta^{(\prime)}$ was previously performed by LHCb using $2\,\text{fb}^{-1}$ of 8 TeV data and $1\,\text{fb}^{-1}$ of 7 TeV data[5]. This resulted in 3σ evidence for the decay $\Lambda_b^0 \to \Lambda\eta$

[4] $\sigma(pp \to b\bar{b}) = 284 \pm 53\,\mu\text{b}$ at 7 TeV [63]
[5] The Λ baryon has a quark content (uds) and mass of $1115.683 \pm 0.006\,\text{MeV}$ [1]

[57], with a measured branching fraction of

$$\mathcal{B}(\Lambda_b^0 \to \Lambda\eta) = 9.3^{+7.3}_{-5.3} \times 10^{-6}. \tag{1.41}$$

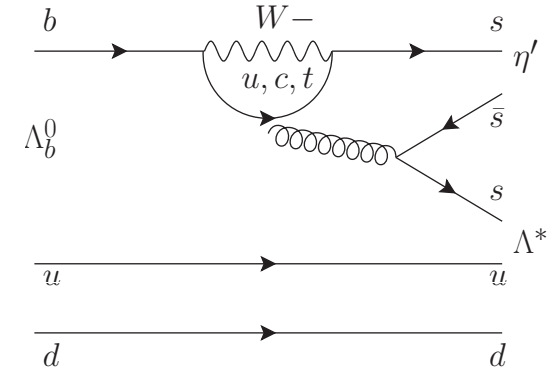
This is consistent with theoretical predictions, but has large uncertainties [65, 66]. The QCD sum rule approach for evaluating the $\Lambda_b^0 \to \Lambda$ form factors is weakly favoured. No evidence for the decay $\Lambda_b^0 \to \Lambda\eta'$ was seen and an upper limit on the branching fraction was set,

$$\mathcal{B}(\Lambda_b^0 \to \Lambda\eta') < 3.1 \times 10^{-6} \tag{1.42}$$

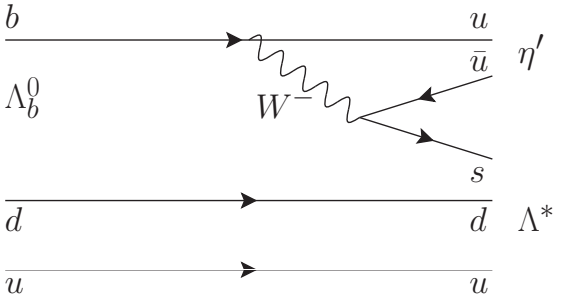
at 90% confidence level. This is also consistent with theoretical predictions [65, 66].

The next step forwards in the search for the decay of a beauty baryon to an $\eta^{(\prime)}$ is searches for the decays $\Lambda_b^0 \to pK^-\eta^{(\prime)}$. In these decays it is highly likely there will be a rich resonant structure of excited Λ^* resonances, such that the decay proceeds $\Lambda_b^0 \to (\Lambda^* \to pK^-)\eta^{(\prime)}$. An amplitude analysis of the decay $\Lambda_b^0 \to pK^- J/\psi$ revealed contributions from 13 Λ^* resonances [16]. The resonant Feynman diagrams, which are shown in Figure 1.5, are identical to those for $\Lambda_b^0 \to \Lambda\eta^{(\prime)}$. However, experimentally these decays are quite different. The presence of the long lived neutral Λ particle caused low trigger efficiencies in the search for $\Lambda_b^0 \to \Lambda\eta^{(\prime)}$, but this is not present in the decays $\Lambda_b^0 \to pK^-\eta^{(\prime)}$. Therefore, a search for the decays $\Lambda_b^0 \to pK^-\eta^{(\prime)}$ should have improved sensitivity.

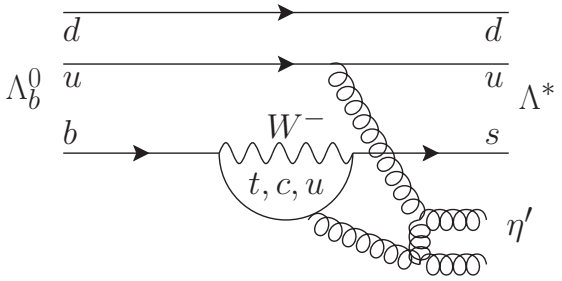
In this thesis a search for the decay $\Lambda_b^0 \to pK\eta'$ is presented, with the decay $B^+ \to K^+\eta'$ used as a normalisation channel. A measurement of the branching fraction could help understand the puzzle of enhanced branching fractions in $B \to K\eta'$ decays; both the gluonic loop level and tree level resonant diagrams are the same except for the presence of an additional spectator quark. Furthermore, it would add information that could be used to understand η'– η mixing and it would be the first



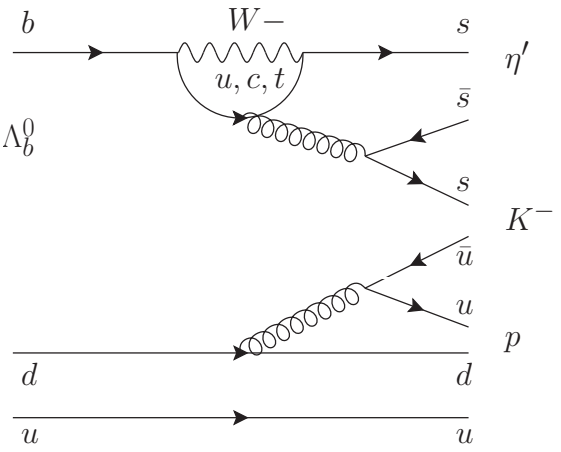
(a) Gluonic loop level diagram



(b) Tree level diagram.



(c) The spectator quark radiates a gluon and the η' is formed through the gluonic component to the wavefunction.



(d) Non resonant decay.

Figure 1.5: Allowed Feynman diagrams for the decay $\Lambda_b^0 \to pK\eta'$.

opportunity to constrain $\Lambda_b^0 \to \eta'$ form factors.

With a sufficiently large signal yield, it could also be possible to perform an amplitude analysis to measure the properties of any intermediate resonances, which have not been investigated before for a $\Lambda_b^0 \to pK^- +$ pseudoscalar decay. Furthermore it would be of significant interest to measure CPV parameters for this decay, given that the first evidence for CPV in the baryon sector was recently reported in charmless Λ_b^0 decays [67]. However it is unlikely a signal yield large enough to perform these measurements will be seen using the data currently available, but as it is part of the LHCb physics programme to collect significantly larger data samples in the future these measurements may become possible.

CHAPTER 2

The LHCb Experiment and the Large Hadron Collider

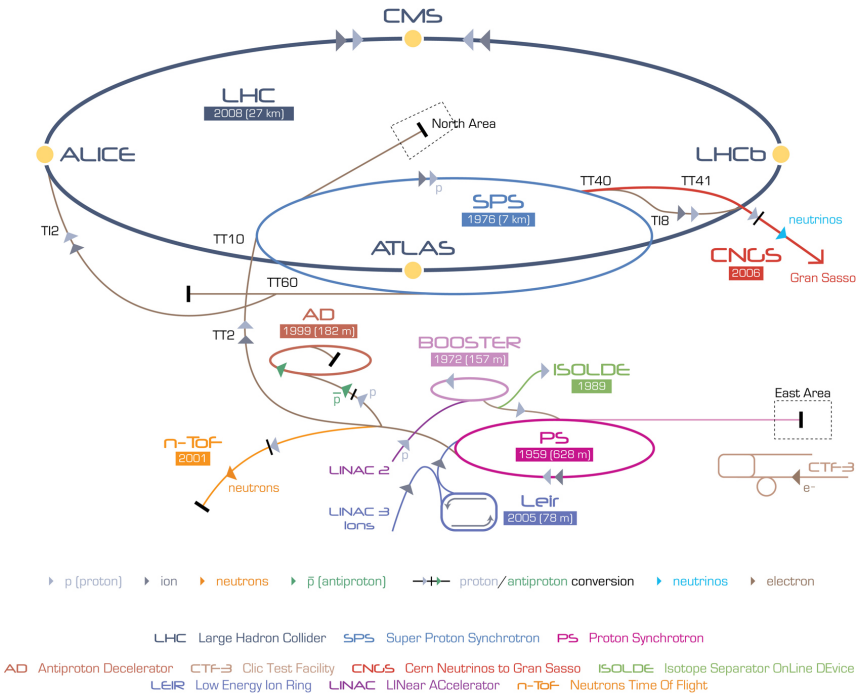


Figure 2.1: The CERN accelerator complex showing the LHC and the chain of injectors used to supply it with protons.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [68] is the highest energy particle accelerator ever built. The 27km circumference ring, which straddles the French-Swiss border at the European Organisation for Nuclear Research (CERN), is currently capable of accelerating protons to a centre of mass energy $\sqrt{s} = 13\,\text{TeV}$. This is achieved whilst delivering an instantaneous luminosity of more than $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ and colliding bunches of protons at a rate of 40 MHz.

The chain of accelerators required to supply the LHC with protons is depicted in Figure 2.1. Protons are isolated by using an electric field, which strips the electrons from hydrogen atoms. These protons are then accelerated to an energy of 50 MeV by a linear accelerator known as "Linac 2" before they are injected into a synchrotron accelerator known as the "Proton Synchrotron Booster" (PSB), where they are accelerated to an energy of 1.4 GeV. The "Proton Synchrotron" (PS) further accelerates the protons to an energy of 25 GeV before they are injected into

the "Super Proton Synchrotron" (SPS) where they are accelerated to an energy of 450 GeV. At this energy, they are injected into the LHC.

The LHC itself is situated 100m underground and uses 1232 superconducting dipole magnets to steer two proton beams around the circumference of the ring in opposite directions. These Niobium-Titanium magnets are supercooled to a temperature of 1.9K($-271.3°$C) and achieve field strengths greater than 8T. The beams nominally consist of 2808 bunches of protons, each of which will contain 1.2×10^{11} protons at the start of a fill. These protons are accelerated to a maximum energy of 6.5 TeV using 8 RF cavities per beam. In order to maximise the number of interactions per bunch crossing, and consequently the instantaneous luminosity, quadrupole magnets are used to keep the proton bunches tightly packed together.

Proton-Proton (pp) collisions take place at four interaction points around the circumference of the LHC. At each of these interaction points there is an experimental hall and particle detector. Two of these interaction points are occupied by General Purpose Detectors (GPD), known as ATLAS and CMS. These are hermetic detectors; they have an angular coverage of 4π steradians. They are used to study a wide variety of physics topics. At another interaction point there is the ALICE experiment, which is a detector specialised for heavy ion physics. The final interaction point is occupied by the LHCb experiment, which is dedicated to studying hadrons containing beauty and charm quarks and is described in further detail in Section 2.2.

The first pp collisions took place at the LHC during 2009, at the lower energy of $\sqrt{s} = 900\,\text{GeV}$. During 2010 the first collisions took place at the energy $\sqrt{s} = 7\,\text{TeV}$, but the corresponding integrated luminosity of the data taken by the LHCb experiment was only $38\,\text{pb}^{-1}$ during this commissioning year [69]. In 2011 the first full year of physics data taking took place at an energy of $\sqrt{s} = 7\,\text{TeV}$; a dataset corresponding to an integrated luminosity of $1.11\,\text{fb}^{-1}$ was collected by LHCb [69].

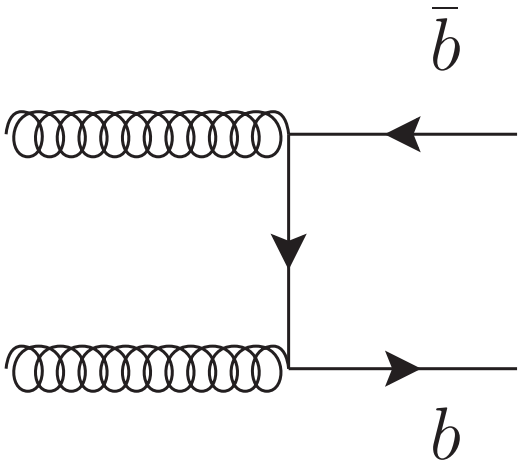


Figure 2.2: The Feynman diagram for the production of a $b\bar{b}$ pair through the gluon-gluon fusion process, which is dominant at the LHC.

This was followed by the collection of a dataset corresponding to an integrated luminosity of $2.08\,\text{fb}^{-1}$ at an energy of $\sqrt{s} = 8\,\text{TeV}$ during 2012 [69]. The data collected by the LHCb experiment during 2011 and 2012 is collectively known as the Run I dataset.

During a long shutdown period that took place in 2013 and 2014, upgrades to the LHC were installed with the aim of increasing the beam energy. In 2015 the LHC started producing proton-proton collisions again with an upgraded collision energy of $\sqrt{s} = 13\,\text{TeV}$, this was the beginning of Run II. However, LHCb was only able to collect a data set corresponding to an integrated luminosity of $0.328\,\text{fb}^{-1}$ during 2015. In 2016 however, the LHCb collaboration was able to collect a dataset corresponding to an integrated luminosity of $1.67\,\text{fb}^{-1}$. With the further $1.8\,\text{fb}^{-1}$ of data collected during 2017 and the planned data taking in 2018, it is expected that a dataset with a total integrated luminosity of $5.4\,\text{fb}^{-1}$ will be collected during Run II.

2.2 The LHCb Experiment

As the primary aims of the LHCb experiment involve studying hadrons containing b quarks, it is crucial the LHCb detector has a high acceptance for b-hadrons. The dominant process for the production of heavy quarks at LHC energies is gluon-gluon fusion, for which the Feynman diagram is shown in Figure 2.2. The values of Bjorken-x[1] for the two virtual gluons contributing to this diagram are given by

$$x_1 = \frac{m_T}{\sqrt{s}}\left(e^{y_Q} + e^{y_{\bar{Q}}}\right) \quad \text{and} \quad x_2 = \frac{m_T}{\sqrt{s}}\left(e^{-y_Q} + e^{-y_{\bar{Q}}}\right), \tag{2.1}$$

where $m_T = \sqrt{m_Q^2 + p_{\mathrm{T}}^2}$, m_Q is the mass of the heavy quark, p_{T} is the transverse momentum of the heavy quark, $\sqrt{s}$ is the centre of mass energy of the pp collision and $y_Q(y_{\bar{Q}})$ are the rapidities of the quark(anti-quark) [70].

The consequences of the expressions given in Equation (2.1) are two fold. Firstly it can be seen in Figure 2.3 that the gluon Parton Distribution Function (PDF) increases significantly at low Bjorken-x, which means it is likely that a gluon contributing to the production of heavy quarks will have a small value of Bjorken-x. In this scenario, given that the mass of a b-quark is 4.18 GeV [1], the transverse momentum and rapidity of the heavy quark must also be small. Secondly, this means that the rapidity of the two heavy quarks produced in the gluon-gluon fusion process are highly correlated.

Figure 2.4 shows the production angles of b and $\bar{b}$ quarks at $\sqrt{s} = 8$ TeV as determined using Pythia 8 [71]. It is clear that, as expected, the production angles of $b\bar{b}$ pairs are highly correlated and that production is concentrated close to the LHC beam pipe. Consequently, despite the LHCb detector only covering the pseudorapidity range $2 < \eta < 5$, 25% of $b\bar{b}$ pairs are produced within the detector acceptance. The advantage of only covering this limited pseudorapidity range is significantly

[1] Bjorken-x is the fraction of a proton's momentum carried by a given parton

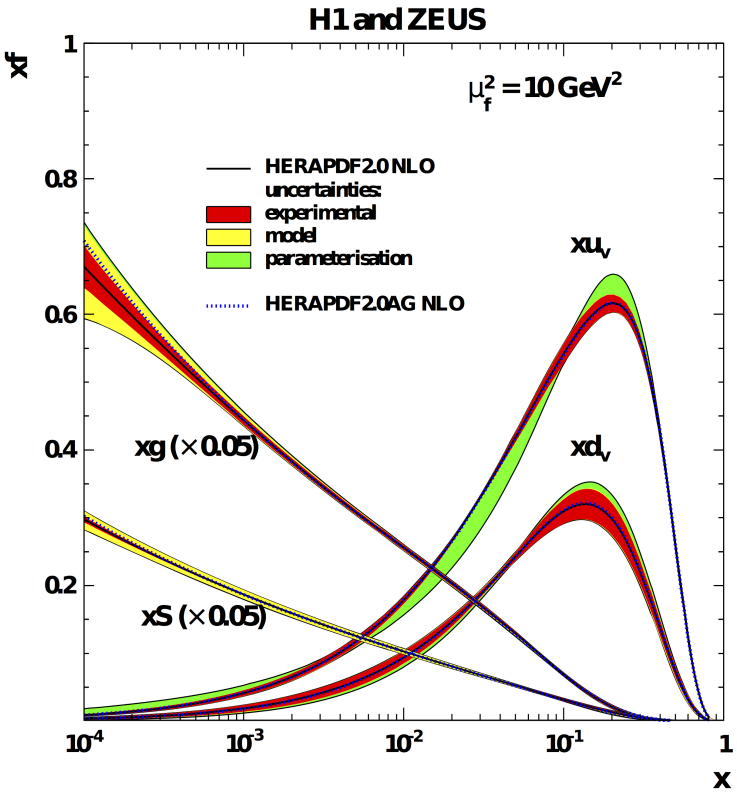


Figure 2.3: Parton Distribution Functions (PDFs), which show the momentum distributions of the partons. A rapid increase in the gluon PDF can be seen at low Bjorken-x. The product of Bjorken-x and the distribution function is plotted on the y-axis because momentum sum rules dictate that $\sum_i \int_0^1 x f_i(x) dx = 1$, where $f_i(x)$ is the distribution function for the parton i [72].

reduced material costs, allowing the use of higher precision instrumentation.

Another key requirement of the LHCb detector is the ability to reconstruct displaced vertices. The mean lifetimes of B^0 mesons, B_s^0 mesons and Λ_b^0 baryons are all approximately 1.5ps [1], which means these neutral hadrons travel $\approx 1\,\mathrm{cm}$ inside the LHCb detector before they decay. Consequently, the decay of such hadrons (to charged particles) leaves a secondary vertex in the LHCb tracking system which is displaced from the initial proton-proton collision. A high reconstruction efficiency for these displaced vertices is crucial to maximise the signal to background ratio of the data collected. The precise reconstruction of displaced vertices is essential for measurements of particle lifetimes and time dependent CP violation, both of which are key physics goals of the LHCb experiment [73].

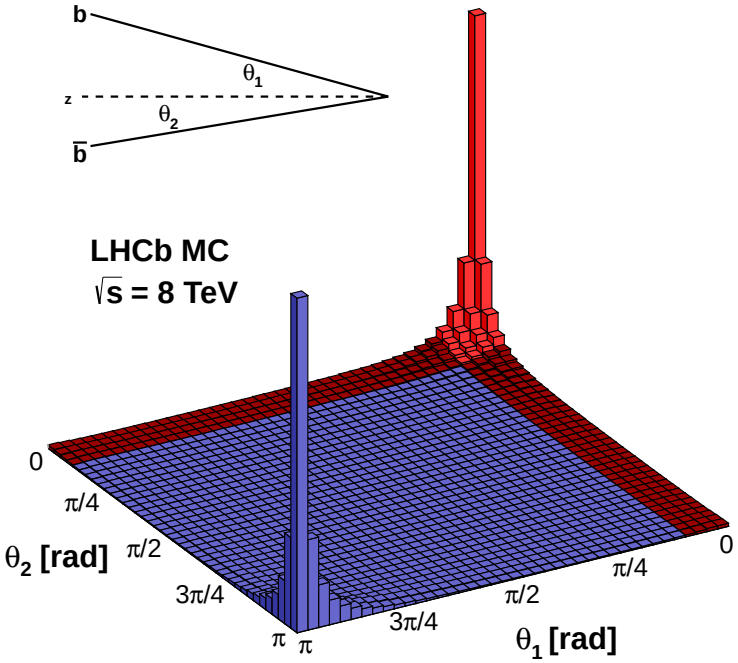


Figure 2.4: The production angles of $b\bar{b}$ quark pairs produced in proton-proton collisions at a centre of mass energy $\sqrt{s} = 8\,\text{TeV}$, as determined with PYTHIA 8 [71]. The region highlighted red corresponds to the LHCb detector acceptance. The production angle is relative to the axis along the centre of the LHC beam pipe and can be translated to pseudorapidity as $\eta = -\ln\tan\frac{\theta}{2}$.

One design choice partially motivated by this requirement is the use of a lower instantaneous luminosity than the GPD. In 2012 the LHCb experiment took data at an instantaneous luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is significantly lower than the $20 - 60 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ luminosity range used by the GPD during the same period. One of the key advantages of using a lower instantaneous luminosity is an increased efficiency for the reconstruction of displaced vertices, because pile-up levels are significantly reduced. It also results in a lower radiation environment which allows precision apparatus to be placed closer to the beam pipe.

The instantaneous luminosity of LHCb, ATLAS and CMS throughout the duration of a single fill of the LHC is shown in Figure 2.5. The instantaneous luminosity of the GPD decreases with time because, as collisions take place, the number of protons in each bunch reduces. Conversely, the instantaneous luminosity at LHCb is deliberately kept constant (within 5%). This is achieved by introducing a transverse offset between the beams, which gradually decreases throughout the duration of a fill. This allows the same trigger conditions to be used throughout a fill and the constant detector occupancy reduces systematic uncertainties [69].

A schematic diagram of the LHCb detector is shown in Figure 2.6; the beam line runs horizontally through the centre of the detector[2] and collisions take place within the vertex locator (VELO), which is described further in Section 2.2.1 The magnet is a warm dipole magnet with an integrated field strength of 4Tm, for tracks 10m in length, which bends the trajectory of charged particles in the $x - z$ plane [74]. The polarity of the magnet is switched periodically to cancel any discrepancies between the detection of particles and anti-particles. The tracking system, which is described further in Section 2.2.2, is composed of a silicon microstrip detector upstream of the magnet (TT), and three further tracking stations (T1, T2 and T3) downstream of the magnet [75]. The two Ring Imaging Cherenkov sub detectors (RICH1 and RICH2)

[2]LHCb uses a right-handed coordinate system, with z defined along the beam axis entering the detector, y vertical and x horizontal towards the inside of the LHC ring. Cylindrical polar coordinates (r,ϕ,z) are also used, as appropriate. The angle θ is defined relative to z.

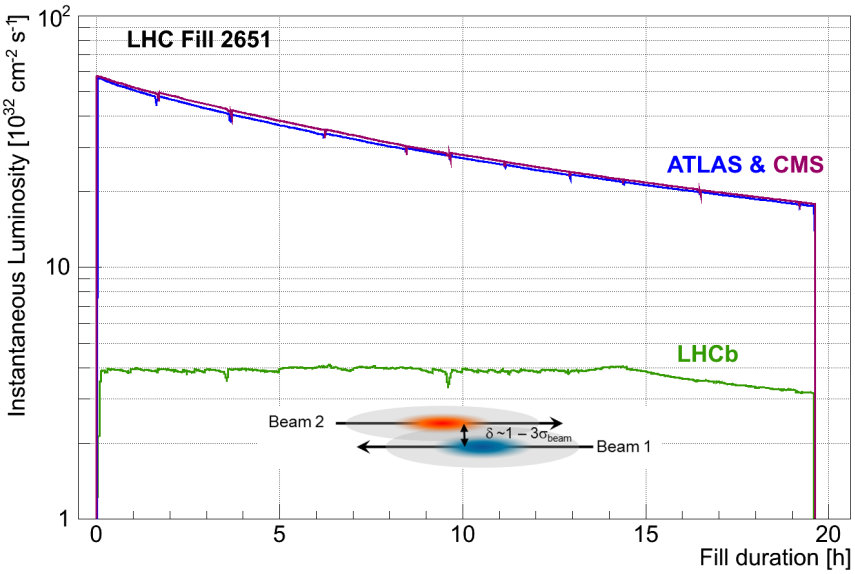


Figure 2.5: The instantaneous luminosity of LHCb, ATLAS and CMS throughout the duration of a single fill of the LHC [69].

give LHCb a unique particle identification ability; these systems are described in Section 2.2.4. The calorimeter system is described in Section 2.2.3 and consists of: a scintillator pad pre-shower detector (SPD/PS); shashlik style electromagnetic sampling calorimeter and a lead-scintillator hadronic calorimeter [73]. The muon system, which is described in Section 2.2.5, has five stations (M1-M5) and is a combination of multi wire proportional chambers and triple-GEM detectors [76]. The trigger system, which is described in Section 2.2.6, consists of a hardware trigger followed by two levels of software triggers [77].

2.2.1 Vertex Locator

The primary purpose of the LHCb vertex locator (VELO) is to identify and reconstruct displaced secondary vertices, which are a distinctive characteristic of many b and c hadron decays. The accurate reconstruction of these displaced vertices is essential for achieving a high signal to background ratio and precise Impact Parameter (IP) and decay time resolution. The VELO consists of 21 circular silicon microstrip

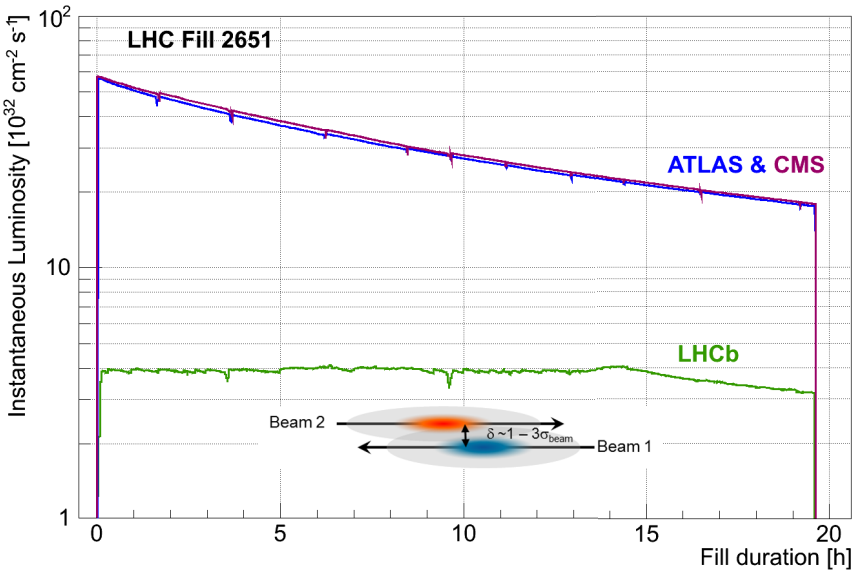


Figure 2.5: The instantaneous luminosity of LHCb, ATLAS and CMS throughout the duration of a single fill of the LHC [69].

give LHCb a unique particle identification ability; these systems are described in Section 2.2.4. The calorimeter system is described in Section 2.2.3 and consists of: a scintillator pad pre-shower detector (SPD/PS); shashlik style electromagnetic sampling calorimeter and a lead-scintillator hadronic calorimeter [73]. The muon system, which is described in Section 2.2.5, has five stations (M1-M5) and is a combination of multi wire proportional chambers and triple-GEM detectors [76]. The trigger system, which is described in Section 2.2.6, consists of a hardware trigger followed by two levels of software triggers [77].

2.2.1 Vertex Locator

The primary purpose of the LHCb vertex locator (VELO) is to identify and reconstruct displaced secondary vertices, which are a distinctive characteristic of many b and c hadron decays. The accurate reconstruction of these displaced vertices is essential for achieving a high signal to background ratio and precise Impact Parameter (IP) and decay time resolution. The VELO consists of 21 circular silicon microstrip

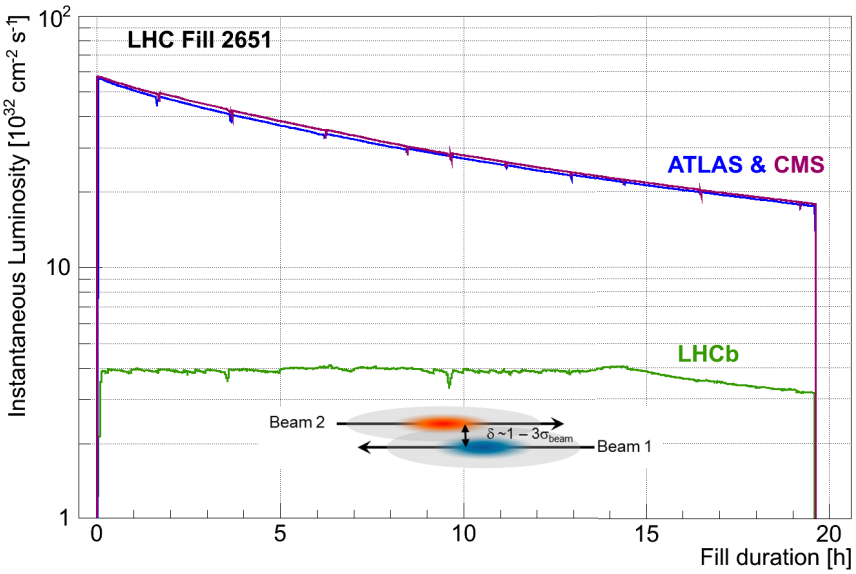


Figure 2.5: The instantaneous luminosity of LHCb, ATLAS and CMS throughout the duration of a single fill of the LHC [69].

give LHCb a unique particle identification ability; these systems are described in Section 2.2.4. The calorimeter system is described in Section 2.2.3 and consists of: a scintillator pad pre-shower detector (SPD/PS); shashlik style electromagnetic sampling calorimeter and a lead-scintillator hadronic calorimeter [73]. The muon system, which is described in Section 2.2.5, has five stations (M1-M5) and is a combination of multi wire proportional chambers and triple-GEM detectors [76]. The trigger system, which is described in Section 2.2.6, consists of a hardware trigger followed by two levels of software triggers [77].

2.2.1 Vertex Locator

The primary purpose of the LHCb vertex locator (VELO) is to identify and reconstruct displaced secondary vertices, which are a distinctive characteristic of many b and c hadron decays. The accurate reconstruction of these displaced vertices is essential for achieving a high signal to background ratio and precise Impact Parameter (IP) and decay time resolution. The VELO consists of 21 circular silicon microstrip

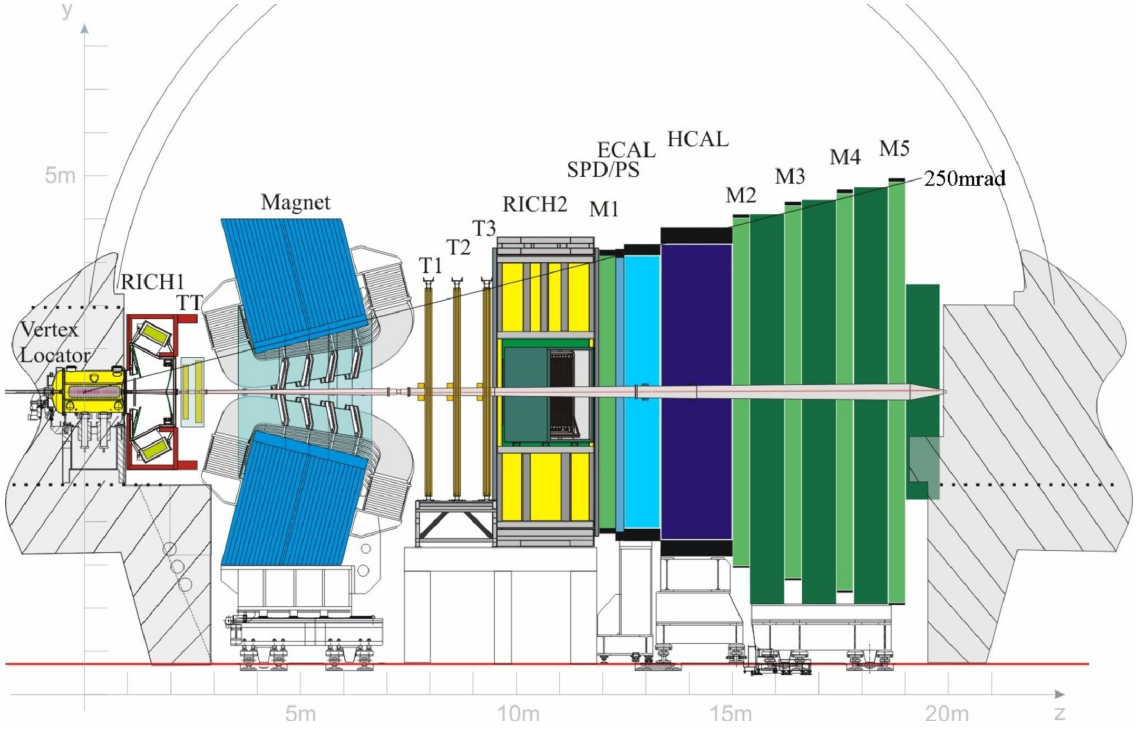


Figure 2.6: A schematic diagram of the LHCb detector [69].

modules positioned along the beamline, a schematic diagram of the VELO positioning relative to the beam line is shown in Figure 2.7. As the VELO is positioned just 7 mm from the LHC beams, it is situated inside the LHC beam pipe. Therefore, it is mounted in a separate vacuum chamber which is separated from the LHC vacuum by an RF box. The close proximity of the VELO to the LHC beams also means they have to be retracted, to the position described as "fully open" in Figure 2.7, until stable beams are declared. This significantly reduces the radiation dose that the VELO receives.

Each VELO module consists of two silicon detectors positioned back to back, one of which is an R sensor and the other a Φ sensor. The R sensors consist of microstrips positioned in concentric semi-circles that provide information on the radial position of tracks. The Φ sensors consist of microstrips which are positioned radially with a pitch that varies linearly between 38μm at the inner and 102μm at the outer edges of each VELO module; these provide information on the azimuthal angle of tracks.

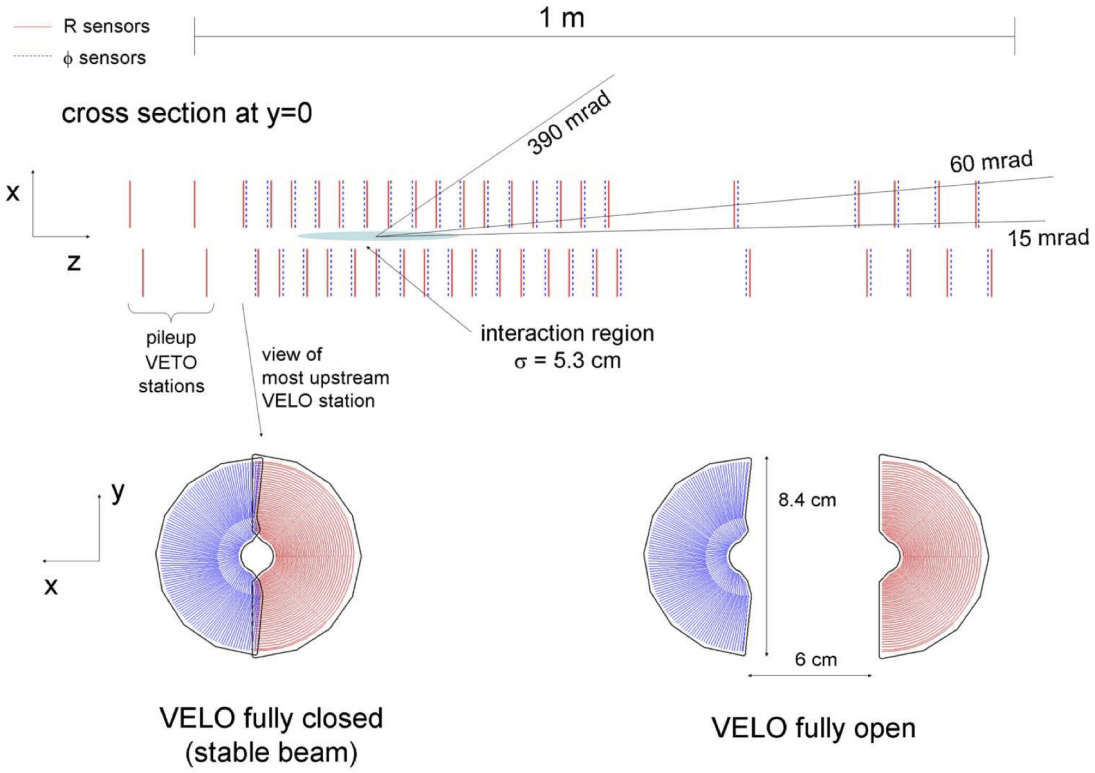


Figure 2.7: Top: Positioning of VELO modules along the beam axis, which runs in the Z direction. Bottom Left: Cross section of the LHCb vertex locator in the fully closed position. Bottom Right: A single VELO module in the open positions [73].

By combining the R and Φ information with the known positions of the modules along the Z axis (beam line) a 3D spatial reconstruction of tracks traversing the VELO can be performed. The reconstructed tracks can then be extrapolated back to their origin vertex and provide precise measurements of impact parameters[3] and particle lifetimes. The use of this circular geometry, rather than a rectilinear detector system, was motivated by the fact that track reconstruction can be performed faster in this coordinate system. The fast reconstruction of VELO tracks is crucial for the efficient triggering of events with displaced vertices.

The IP resolution has been assessed using prompt tracks; prompt tracks originate from the primary vertex, therefore any non-zero values of IP are due to experimental resolution. Figure 2.8 shows the IP resolution as a function of $\frac{1}{p_{\mathrm{T}}}$ for 2012 data and simulation. This demonstrates excellent performance; an IP resolution of $< 35\,\mu\mathrm{m}$ is achieved for tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ [78].

The decay time resolution of the VELO has also been assessed using prompt tracks, specifically events with two muon tracks and two kaon tracks that mimic $B_s^0 \to J/\psi\,\phi$ decays. Figure 2.9 shows the decay time resolution of the VELO for 2011 data and simulated events. This again demonstrates excellent performance with a decay time resolution of 50 fs. Crucially, this is small enough to resolve B_s^0-$\overline{B}_s^0$ oscillations [78].

2.2.2 Tracking

The purpose of the tracking system is to reconstruct the paths of charged particles and measure their momenta. Accurate momentum measurements are crucial to the majority of the physics programme of LHCb because they have a direct impact on the invariant mass resolution of reconstructed hadrons. For example, in order to

[3]The impact parameter of a particle is defined as the transverse distance of closest approach between the trajectory of a particle and the primary vertex.

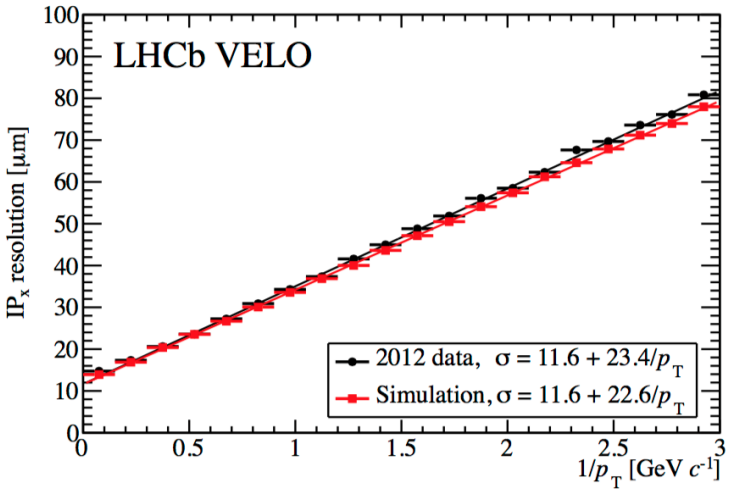


Figure 2.8: Impact parameter resolution of the VELO, as measured using prompt tracks in 2012 data and simulation [78].

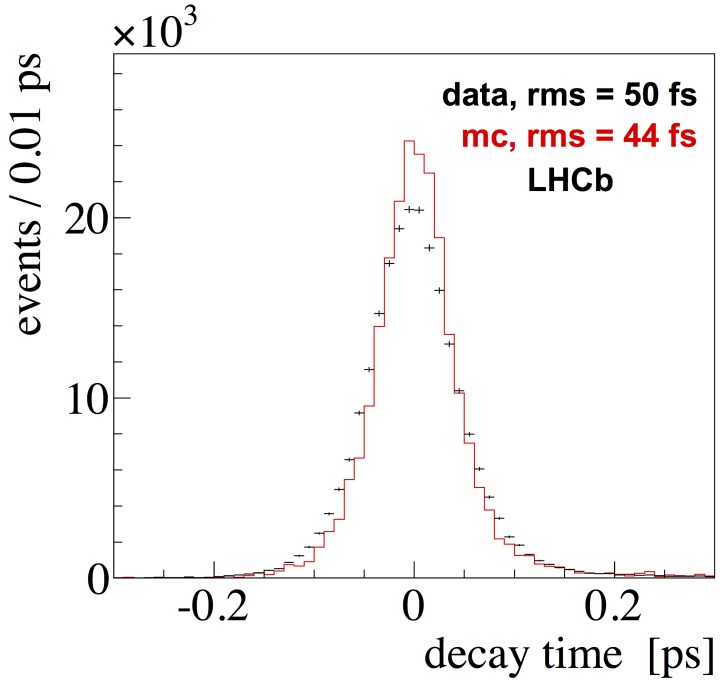


Figure 2.9: Decay time resolution of the VELO. This is measured using prompt events which mimic $B_s^0 \to J/\psi\,\phi \to \mu^+\mu^- K^+ K^-$ decays in 2011 data and simulation [78].

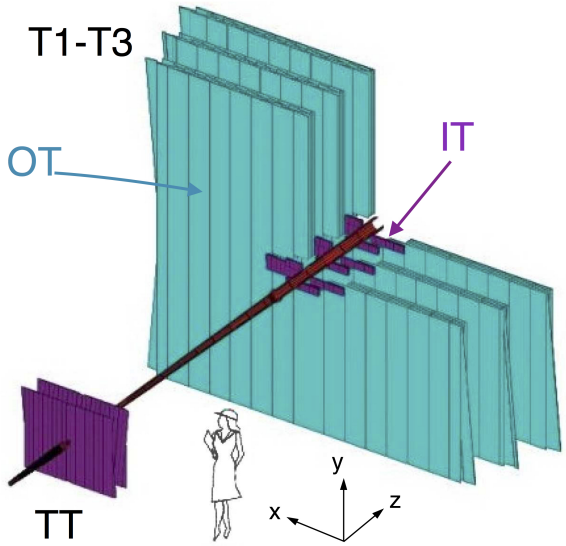


Figure 2.10: A schematic diagram of the LHCb tracking stations. The ST modules are shown in purple; the TT is in the foreground and the cross shaped IT modules are shown at the centre of the three larger tracking stations. The turquoise areas of T1-T3 are the OT. [73]

achieve a resolution of 10 MeV on the reconstructed B_s^0 mass in $B_s^0 \to D_s^- \pi^+$ decays, a momentum resolution of $\frac{\delta p}{p} \approx 0.4\%$ is required [73].

The LHCb tracking system utilises two different technologies: silicon microstrip detectors and straw drift tubes. There is one tracking station upstream of the magnet known as the Tracker Turicensis (TT) which uses solely silicon microstrips. The T1-T3 tracking stations, which are downstream of the magnet (see Figure 2.6), are further divided into the inner tracker (IT) and outer tracker(OT). The IT occupies an area 120cm wide and 40cm high in a cross-shaped arrangement around the beam pipe and also uses silicon microstrips. Together, the IT and TT form the silicon tracker (ST). The OT occupies an area $5 \times 6\text{m}^2$ and uses straw drift tubes; the use of silicon microstrips for an area the size of the OT would have been too expensive, therefore silicon microstrips are reserved for areas with the highest particle densities. A schematic diagram showing the IT, OT and TT can be seen in Figure 2.10. Each tracking station is formed of four layers, with the two outer layers orientated vertically and the inner two layers tilted $\pm 5^\circ$ with respect to the vertical.

The silicon microstrips used in the TT and IT are single sided p^+-on-n sensors and have a pitch of 183 µm and 198 µm respectively. The TT uses 143360 readout strips up to 38 cm in length, whereas the IT uses 129024 readout strips which are between 11cm and 22cm long. The TT and IT have active areas of $8.4m^2$ and $4.0m^2$ respectively. The ST was designed to have a single hit resolution of 50 µm because below this level momentum resolution is dominated by multiple scattering. Another key design requirement was radiation hardness; the TT (IT) needs to survive a fluence of $5 \times 10^{14}\,cm^{-2}(2 \times 10^{12}\,cm^{-2})$ 1 MeV neutron equivalent dose. In order to meet this requirement the ST has to be kept at temperatures below 5°C.

When charged particles traverse a drift tube the gas is ionised and ionisation electrons drift to the anode at the centre of the tube and produce an electrical signal. By measuring the time difference between the signal on the anode and the beam crossing time (the drift time) the distance of the charged particle's trajectory from the centre of a drift tube can be inferred. By combining information from all the drift tubes in the OT the trajectories of charged particles can be reconstructed. The straw drift tubes used in the OT are arranged into two staggered layers, as shown in Figure 2.11. Each drift tube has an inner diameter of 4.9 mm and is filled with a mixture of Argon(70%) and CO_2(30%). This mixture was chosen to ensure a drift time of less than 50ns and a drift-coordinate resolution of 200 µm.

Overall, the LHCb tracking system provides excellent momentum resolution; in Run I the momentum resolution was found to vary from $\frac{\delta p}{p} = 0.5\%$ for tracks with $p = 5\,GeV$ up to $\frac{\delta p}{p} = 1.1\%$ for tracks with $p = 200\,GeV$ [69]. The track reconstruction efficiency depends on momentum, pseudorapidity, total number of tracks in an event and also the number of reconstructed primary vertices in an event. In the kinematic range $5\,GeV < p < 200\,GeV$ and $2 < \eta < 5$ the average track reconstruction efficiency was measured in Run I to be 96% [69].

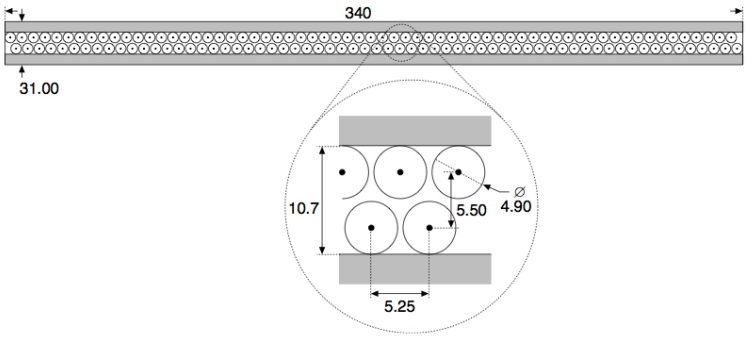


Figure 2.11: The arrangement of straw drift tubes used in the OT

2.2.3 Calorimeters

The main purposes of the calorimeter systems are: to provide discrimination between photons, electrons, neutral pions and hadrons; reconstruct the position and energy of photons, electrons and neutral pions; and provide fast transverse energy measurements for use in the L0 trigger. The first substantial part of the calorimeter system is an electromagnetic sampling calorimeter (ECAL) which makes use of the shashlik design [79]. The role of the ECAL is to provide position and energy measurements of photons, electrons and neutral pions. The next major component is a hadronic sampling calorimeter (HCAL), which provides position and energy measurements of hadrons, and is positioned downstream of the ECAL. These energy measurements are crucial for effective triggering at the L0 hardware trigger level. In order to provide discrimination between charged hadrons, photons and electrons, longitudinal separation is added with the use of pre-shower detectors upstream of the ECAL.

The pre-shower detectors consist of a plane of high granularity scintillator pads (SPD) followed by a $15\,\mathrm{mm}(2.5X_0)$ lead converter. The role of the lead converter is to induce showers from photons which are then detected by the pre-shower detector (PS), which is another plane of scintillator pads positioned between the lead converter and the ECAL. The use of the PS/SPD system, in combination with the calorimeters, provides discrimination between particle species. Photons will typi-

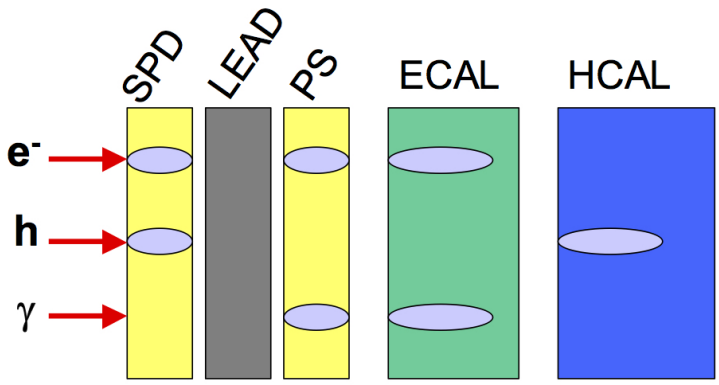


Figure 2.12: The differing signals left in the calorimeter systems and SPD/PS detectors by photons, electrons and charged hadrons. [80]

cally leave no hits in the SPD but hits in the PS and ECAL, whilst electrons will typically leave hits in the SPD, PS and ECAL. Charged hadrons will typically leave hits in the SPD but no hits (or very few) in both the PS and ECAL. A diagram summarising the differing signatures of electrons, photons and charged hadrons is shown in Figure 2.12. This shows the importance of the SPD/PS for particle identification.

The ECAL is segmented laterally into three different cell sizes; 40.4mm square cells are used closest to the beam pipe, 60.6mm cells are used slightly further from the beam pipe and 121.2mm cells are used to cover the outer areas of the calorimeter. The positioning of the different cell sizes is shown in Figure 2.13. Each calorimeter cell consists of 66 alternating layers of lead and scintillator sheets. The lead sheets are 2mm thick and the purpose of this small radiation length material is to induce an electromagnetic shower within a short distance. The scintillator sheets are 4mm thick and their role is to provide light signals that are proportional to the energy of an electromagnetic shower. The light signals are collected by wavelength shifting fibres (WLS), which both transport the light signal to photomultiplier tubes (PMT) at the rear of the cell and shift the wavelength of the light to a range suited to the operating range of the PMT. With knowledge of within which cell the shower was detected and information on the size of the shower, both the position and energy of incident particles can be reconstructed. The ECAL is 42cm and 25 radiation

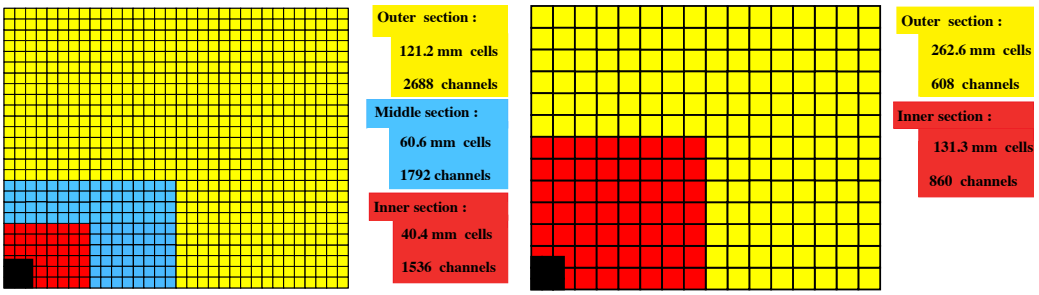


Figure 2.13: Left (Right): The lateral segmentation and different sized cells used in the ECAL (HCAL). [73]

lengths deep; full containment of electromagnetic showers is required to reach the design energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, which translates to a mass resolution of 65 MeV for the important decay $B^0 \to K^{*0}\gamma$.

Conversely, the energy resolution of the HCAL is not as important because, given a mass hypothesis, the energy of a charged hadron is determined from the far more precise momentum measurements made by the tracking system. Therefore, full shower containment is not necessary and the HCAL is only 5.6 hadronic interaction lengths deep. This is also partly due to the limited space available for the HCAL. The HCAL uses a similar shashlik sampling calorimeter design to the ECAL, but in this case iron is used as an absorber material. The cell sizes are also different; only two different cell sizes are used, which are 131.3mm (262.6mm) in the inner(outer) region as shown in Figure 2.13. The energy resolution of the HCAL has been measured in test beams to be $\sigma_E/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$ [73].

2.2.4 Charged Particle Identification

Another key requirement of the LHCb detector is to identify correctly different charged hadron species; the ability to positively identify protons, kaons and pions is crucial to the LHCb physics programme. This is achieved through the use of two ring imaging Cherenkov detectors, RICH1 and RICH2. When a charged particle

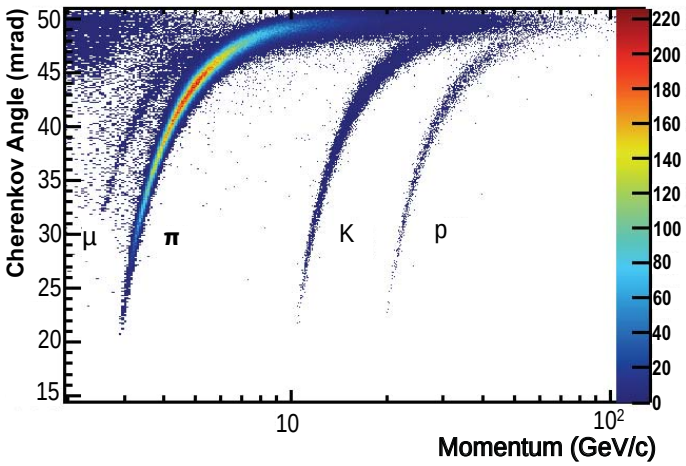


Figure 2.14: Reconstructed Cherenkov angle as a function of momentum. The discrimination between different particle species can very clearly be seen. [81]

traverses a medium at a speed greater than the speed of light in that medium, Cherenkov photons are emitted. The angle, θ_c, at which these Cherenkov photons are emitted is given by

$$\cos\theta_c = \frac{1}{\beta n}, \tag{2.2}$$

where n is the refractive index of the medium and $\beta = v/c$. The refractive index of the medium used in the RICH1(RICH2) sub detector is 1.0014(1.0005). Given that β can be expressed as $\beta = \frac{p}{\sqrt{p^2+m^2}}$, the Cherenkov angle has a dependence on the mass of the particle. Consequently, by measuring the angle at which Cherenkov photons are emitted in a medium with a well known refractive index, different particle species can be positively identified. This principle is illustrated in Figure 2.14, which shows the reconstructed Cherenkov angle for several particle species as a function of track momentum.

The RICH systems both detect Cherenkov photons by using a series of mirrors to focus the photons onto an array of dedicated and novel Hybrid Photon Detectors (HPDs) [82]. The HPDs are vacuum phototubes which accelerate photoelectrons, emitted from the photocathode, onto a silicon pixel detector. The HPDs offer a high granularity ($2.5 \times 2.5\text{mm}^2$) and 25 ns timing resolution, which are properties

necessary for successful operation of the RICH detectors.

The RICH1 sub detector is positioned upstream of the magnet between the VELO and TT (see Figure 2.6) and covers the low to intermediate momentum range of $2 - 40\,\text{GeV}$. The RICH2 detector is positioned downstream of the magnet after the tracking stations but before the calorimeters and covers the high momentum range of $15 - 100\,\text{GeV}$. The RICH1 sub-detector uses C_4F_{10} gas as a radiator, along with a 50mm thick sheet of silica aerogel at the entrance to the sub-detector [83]. The sheet of aerogel is used to ensure kaon identification can still be performed at low momentum; the threshold for a kaon to radiate Cherenkov light in C_4F_{10} is $9.3\,\text{GeV}$ [81]. The RICH2 sub-detector uses a gas mixture of 95% CF_4 and 5% CO_2, the CO_2 is added to quench scintillation.

The information from the RICH detectors is used by performing a fit to the distribution of hits in the RICH sub-detectors under various particle hypotheses for each track; the Particle Identification (PID) hypothesis which maximises the global likelihood is assigned to the track. The quality of the PID assigned is quantified using the difference in Log-Likelihood when the fit is performed with a given PID hypothesis, and when the fit is performed with the pion ID hypothesis. These likelihood differences are known as the Delta Log Likelihood (DLL) variables; for example $\text{DLL}_{p\pi}$ is given by

$$\text{DLL}_{p\pi} = \log(\mathcal{L}_p) - \log(\mathcal{L}_\pi), \tag{2.3}$$

where $\mathcal{L}_p$ is the likelihood of the fit under the proton hypothesis and $\mathcal{L}_\pi$ is the likelihood of the fit under the pion hypothesis.

To further improve the use of the PID information available, a neural network is trained to create global particle ID variables. The neural network makes use of the DLL variables as an input, and also uses information from the calorimeter and tracking systems to make best use of all the information from the LHCb detector. The output of the neural network is then normalised to give the `ProbNN` set of

variables. These variables are simply the probability of a given particle being of the type assigned in the PID hypothesis, for example `Kaon_ProbNNk` is the probability of a track which has been assigned a kaon hypothesis truly being a kaon.

However, neither the DLL nor `ProbNN` variables are well modelled in MC. This is because the interaction of the particles with the RICH detector involves several complex second order processes which are not well simulated. Consequently, the MC cannot be used to calculate efficiencies; data calibration samples have to be used to determine the efficiency of cuts placed on PID variables. The calibration samples are from decays where the PID of the tracks involved can be determined through kinematics alone. These samples can then be used to determine PID efficiencies as a function of kinematic variables (see Section. 4.3.3), allowing per-event efficiencies to be assigned to the MC based on the signal kinematics.

Overall, the RICH detectors provide LHCb with unprecedented particle identification ability; averaged over the momentum range $2-100\,\text{GeV}$ kaons can be positively identified with a $\approx 85\%$ efficiency with just a 3% pion mis-identification rate. The impact of this performance on physics analyses is profound. For example, Figure 2.15 shows the reconstructed invariant mass of $B \to h^+h^-$ decays with (right) and without (left) PID requirements applied. The channel $B^0 \to \pi^+\pi^-$ is a key decay channel for measurements of time dependent CP violation but, as shown in Figure 2.15, without PID information it would be exceptionally difficult if not impossible to study due to mis-ID backgrounds [73, 84].

2.2.5 Muon System

The efficient and accurate reconstruction of muons is imperative for many key physics results. These include, but are not limited to: precision measurements of CP asymmetries in decays such as $B^0 \to J/\psi\,(\mu^+\mu^-)K^0_S$, searches for very rare

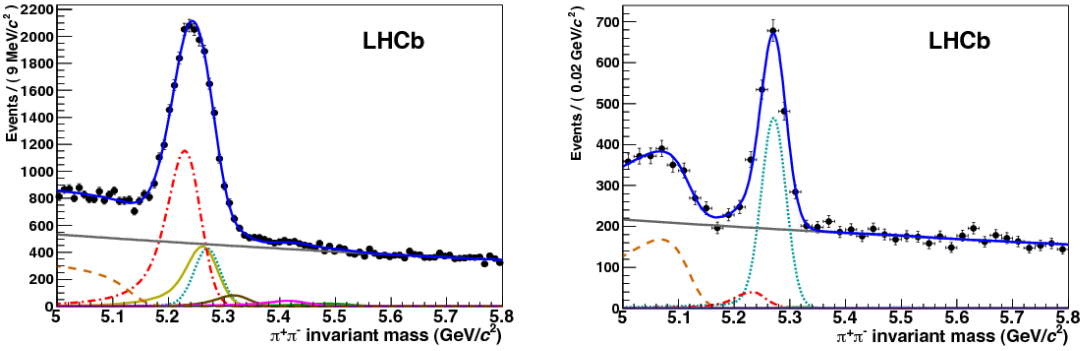


Figure 2.15: A comparison of the $B \to h^+h^-$ invariant mass distribution without (left) and with (right) the use of PID information. The signal decay under study, $B^0 \to \pi^+\pi^-$, is represented by the turquoise dotted line. Background contributions from: $B^0 \to K^+\pi^-$ *red dashed-dotted line*; $B^0 \to 3-$body *orange dashed-dotted line*; $B_s^0 \to K^+K^-$ *yellow/gold solid line*; $B_s^0 \to K^+\pi^-$ *brown line*; $\Lambda_b^0 \to pK^-$ *purple line* and $\Lambda_b^0 \to p\pi^-$ *green line* are also shown. The full fit function is represented by the *solid blue line* [81].

decays such as $B \to \mu^+\mu^-$ and lepton universality tests such as $\mathrm{R}(K^{*0})$ [85–87]. The LHCb muon system is composed of five detector stations (M1-M5). As can be seen in Figure 2.6, the first muon station (M1) is positioned upstream of the calorimeters whilst stations M2-M5 are positioned downstream of the calorimeters. The muon station positioned upstream of the calorimeter improves the resolution of the p_{T} measurement used in the trigger by $\approx 25\% - 35\%$[4] [76].

With the exception of the inner part of M1, each muon station is equipped with 276 Multi-Wire Proportional Chambers (MWPC). The inner region of M1 uses 12 triple Gas Electron Multiplier (GEM) gaseous detectors because this region receives the highest radiation dose, and it could not be guaranteed that MWPCs in this region would survive the required 10 years of operation. The muon stations are separated into logical regions in order to provide spatial information. Each muon station is separated by 80cm of iron, meaning that only muons with $p > 6\,\mathrm{GeV}$ will penetrate all five muon stations.

[4]M1 is not used in the offline reconstruction of muons because tracks from M2-M5 can be extrapolated back to the T1-T3 tracking stations which provides a more precise momentum measurement

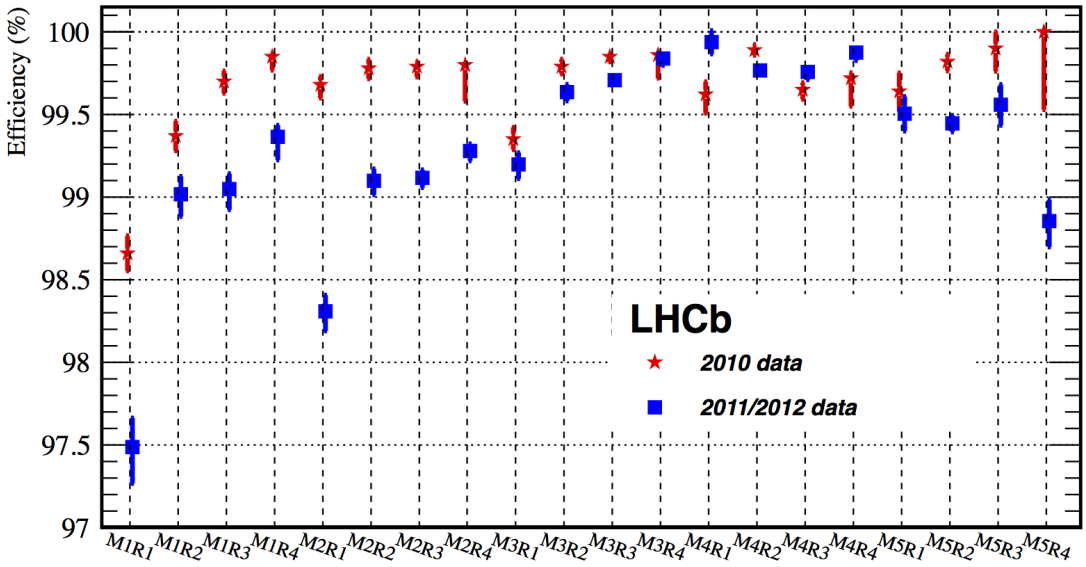


Figure 2.16: Muon reconstruction efficiency for each region of the muon stations, as measured in 2010 and 2011/12 data [69]. The efficiency for a given region is estimated by searching for hits around the predicted trajectory of a track based on extrapolating from hits in every other muon station. The efficiency is taken as the fraction of regions for which hits are found.

The design muon reconstruction efficiency requirement was 95%, which means each muon station individually must have an efficiency of $> 99\%$. The actual efficiency of the muon reconstruction has been measured in 2010, 2011 and 2012 data separately; a plot showing this efficiency for each region of the muon stations is shown in Figure 2.16. The lower efficiencies in 2011/2012 data are caused by dead time in the detector readout, which was a consequence of running at twice the detector design luminosity in these years. However, the combined response for muon identification and reconstruction was still performed with an efficiency $> 95\%$ for all three years of detector operation [69].

2.2.6 Trigger

The LHC collides proton bunches at a rate of 40 MHz, but the maximum rate at which data from the LHCb detector can be read out is ≈ 5 kHz [73]. Therefore,

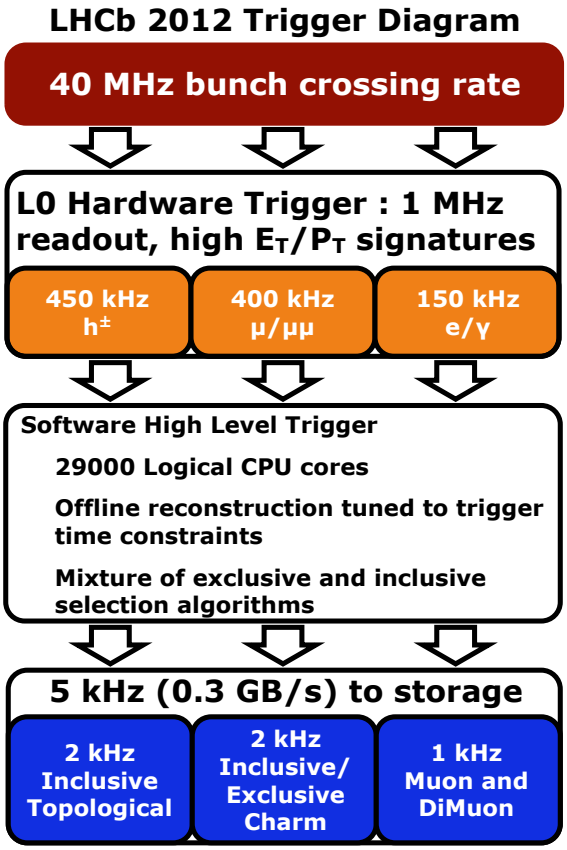


Figure 2.17: The LHCb trigger system in 2012 data taking [77]. A bunch crossing rate of 40 MHz is reduced to 5 kHz.

a trigger system is required to select which events are stored. An overview of the LHCb trigger system is shown in Figure 2.17. There is firstly a hardware trigger, known as the L0 trigger, which reduces the event rate from 40 MHz to 1 MHz. There are then two levels of software trigger known as the High Level Triggers (HLT); the HLT1 reduces the event rate to $\approx$ 80 kHz and the HLT2 reduces the event rate to 5 kHz. These three trigger systems are described further in the following sections.

2.2.6.1 L0 Trigger

In order to make decisions at a rate of 40 MHz, the L0 trigger is implemented in dedicated electronics. The L0 trigger can be further divided into three separate systems: the L0-Calorimeter trigger, L0-Muon trigger and L0-PileUp trigger.

The L0-Calorimeter trigger makes use of information from the ECAL, HCAL, SPD and PS to select hadrons, electrons and photons. The transverse energy, E_T, of particles which leave deposits in the calorimeters is computed in clusters of 2×2 cells. The highest E_T HCAL cluster forms a `L0Hadron` candidate, and if the E_T of the candidate is above a fixed threshold it is selected. In 2011 the `L0Hadron` threshold was 3.5 GeV and in 2012 the threshold was increased to 3.62 GeV; the increase was necessary due to the increase in beam energy [69]. As described in Section 4.2.1, the majority of candidates used in this analysis are selected by the `L0Hadron` trigger. The efficiency of the `L0Hadron` trigger as a function of p_{T} for various hadronic signatures is shown in Figure 2.18.

The highest E_T ECAL cluster is also reconstructed and if there is either one or two PS hits in front of the cluster but no SPD hits then it forms a `L0Photon` candidate. In the event there are both PS and SPD hits in front of the ECAL cluster, the cluster is considered a `L0Electron` candidate. Both `L0Photon` and `L0Electron` candidates were required to meet an E_T threshold of 2.5 GeV during 2011 data taking and 3.0 GeV during 2012 data taking. The `L0Photon` trigger is not used in this analysis, despite the presence of a photon in the final state, because the `L0Photon` threshold is tuned to select events where high p_{T} photons from radiative decays are a defining characteristic.

2.2.6.2 HLT1

The HLT uses the same software framework as the offline reconstruction but with some simplifications due to CPU time constraints. At HLT1, the VELO reconstruction, track reconstruction and the matching of hits in the VELO to hits in the tracking stations are all performed. However, VELO tracks pointing a long way from the beamline are not reconstructed and the Kalman filter, used to perform track fits, operates with a simplified geometry and fewer iterations [88]. Further-

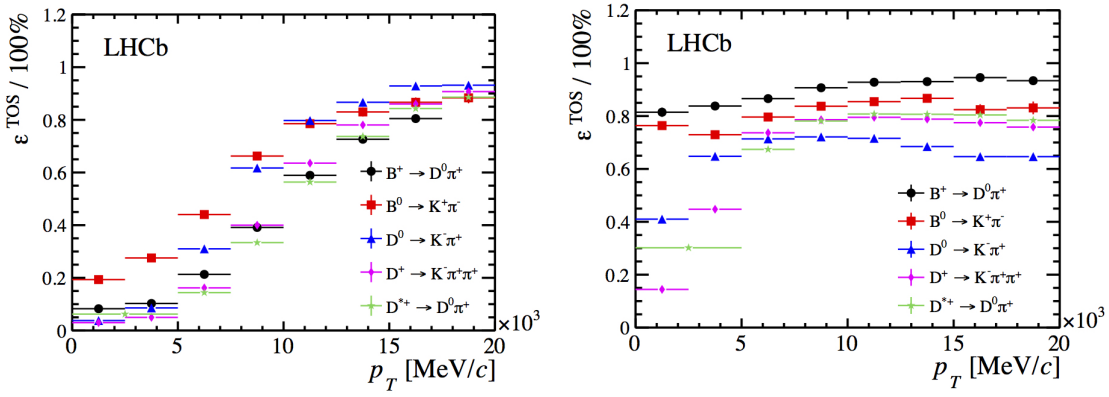


Figure 2.18: The efficiency of the `L0Hadron` trigger (left) and the HLT1 inclusive track trigger (right) for various hadronic decays as a function of p_T [69].

more, only tracks with a significant IP and with a $p_T > 0.5\,\text{GeV}$ or $p_T > 1.25\,\text{GeV}$[5] are matched to the tracking stations. As described in Section 4.2.1, this analysis makes use of an inclusive beauty and charm track trigger at the HLT1 level (`Hlt1TrackAllL0Decision`). This trigger line requires one good quality track with $p_T > 1.6 - 1.7\,\text{GeV}$ (depending on year of data taking) and IP > 0.1mm. The bandwidth used by this trigger is around 58 kHz, which is the largest of any HLT1 line but it is also the most efficient for non-leptonic physics channels. The efficiency of the inclusive HLT1 track trigger for various hadronic decays is shown in Figure 2.18.

2.2.6.3 HLT2

At the HLT2 level the event rate is sufficiently low to allow the use of the "forward tracking" algorithm [89] for all VELO tracks and the calculation of many higher level topological variables. In order to create an efficient generic beauty trigger, these variables are combined using a Bonsai Boosted Decision Tree (BBDT) [90]. One potential problem with the use of a Boosted Decision Tree (BDT) is instabilities due to statistical fluctuations in the training samples used or finite detector resolutions. This would be highly undesirable because it could lead to reduced performance and/or selection efficiencies that may be unreliable. To avoid this a BBDT uses

[5] there are two separate selection requirements which define two types of reconstructed track

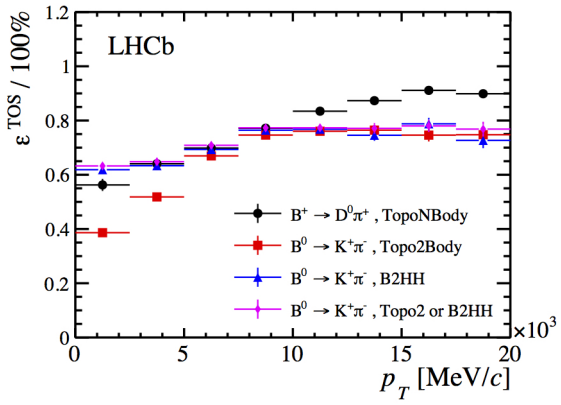


Figure 2.19: Performance of the BBDT based HLT2 generic beauty trigger (TopoN-Body) for $B^0 \to K^+\pi^-$ and $B^0 \to D^+\pi^-$ decays. The performance of the exclusive $B^0 \to hh$ trigger is also shown [69].

discretised data, where the width of the binning for each variable is greater than the detector resolution. This mitigates detector resolution effects and also lowers the impact of instabilities arising from statistical fluctuations, because it reduces the number of ways in which the data can be split. This also means the different splits can be saved in memory, which greatly improves the CPU performance making the use of this algorithm in the trigger feasible. This approach has been shown to have comparable stability and efficiency to a cut based approach whilst providing reduced background retention rates.

There are separate BBDT based generic beauty triggers trained to select two, three and four body decays because the observables characterising a decay vertex depend on the number of particle tracks associated with it. However, as all of these triggers are trained to select partially reconstructed beauty decays these triggers are not mutually exclusive. In this analysis, all three n-body triggers are used. The performance of the two-body trigger for $B^0 \to K^+\pi^-$ and $B^0 \to D^+\pi^-$ decays is shown in Figure 2.19; a signal efficiency greater than 60% for tracks with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ is achieved, which is good performance given the reduction in bandwidth required.

CHAPTER 3

Validation of Simulation of Electromagnetic Processes

In any HEP experiment the accurate simulation of both the physics processes and the subsequent response of the detector is crucial for optimising selections, evaluating their efficiencies and carrying out detector performance studies. Almost every physics measurement makes use of simulated samples, therefore the validation of the simulation process is of paramount importance. As well as ensuring the simulated reconstructed quantities are consistent with data, e.g calorimeter energies or impact parameters of tracks, any change in results between different versions of the simulation software also needs to be fully understood. This chapter begins with a brief introduction to the simulation framework and methods used by LHCb. This is followed by a description of the validation work that has been performed for electromagnetic interactions of particles with the LHCb detector. More specifically, two benchmarking tests of the GEANT4 toolkit [91] have been developed; a simple test that emulates the LHCb electromagnetic calorimeter is described in Section 3.3 and a test of multiple scattering in thin sheets of silicon is discussed in Section 3.4.

3.1 The Simulation of the LHCb detector

The complexity of the calculations required to perform simulations in HEP and the naturally repetitive nature of the experiments mean a Monte Carlo approach to simulation is by far the most efficient, most natural and often the only viable method. The LHCb simulation package, known as GAUSS [92], is no exception. The full simulation of an event within GAUSS comprises three stages: the generation of the underlying pp collision; the decay of hadrons and leptons; and the interaction of final state particles with the LHCb detector.

The generation of the pp collisions and the outgoing particles is performed using PYTHIA [71]. This is an external generic event generator which has an LHCb specific tuning applied; variable hyperparameters are adjusted such that the generated kine-

matic distributions match those seen in data as closely as possible. The next stage is the decay of unstable particles, which is simulated with the EvtGen package [93]. This package was originally designed for the BaBar experiment and can simulate a wide variety of decay amplitudes. Unlike at BaBar, B^0 and B_s^0 meson production is incoherent at LHCb, therefore some modifications to EvtGen are implemented to account for this. The EvtGen package is also responsible for the time evolution of particles.

The final stage of the simulation is the tracking and interaction of particles with the LHCb detector. This task is handled by the Geant4 toolkit [91], and it is this stage of the simulation process that is the subject of the validation work presented in this chapter. Information about the particles created by the previous stages of the simulation is passed to Geant4, along with a detailed description of the detector geometry. Geant4 then simulates the passage of a particle through the detector by stepping it through the geometry being modelled.

Before any step is taken, the step size is calculated for every possible physical process considered, in the material in question, and the smallest is chosen. For any given physical processes the step length is determined purely by cross sections; the distribution of distance traveled before the next interaction takes place is calculated and sampled for a specific scenario. On the other hand, for some physical processes (such as multiple scattering) not every interaction is simulated, instead statistical effects are applied after a step has been taken. In this case the step length is chosen as a balance between accuracy and CPU time. After a step is taken the probabilities of all physics processes having taken place are calculated, for the step size used, and random numbers are generated to decide if they take place. This includes, but is not limited to: energy loss, changes of direction and production of secondary particles. This process is then repeated until the particle exits the geometry, loses all its energy or is annihilated.

The full simulation process is then finalised by passing information about energy deposits in sensitive detector volumes from GEANT4 to a separate package known as BOOLE. This package emulates the response of the detector readout electronics to the energy deposited in the sensitive material of the detector. The simulated electronic signals can then be processed in the same manner as data to produce the same reconstructed quantities.

The full simulation process with GAUSS consumes a significant amount of CPU time, which is not an unlimited resource. When potential sources of background are being considered, some can be ruled out based on variables which are calculable without a full detector simulation. In order to understand which backgrounds need to be considered in the analysis presented in this thesis, the `RapidSim` fast simulation package is use to understand the kinematics of potential backgrounds [94]. This package uses a phase-space model to simulate the kinematics of heavy hadron decay chains, and fixed-order next-to leading-log (FONLL) calculations to reproduce the correct kinematics of the heavy hadron (the boost of the heavy hadron). Smearing effects are then applied to the final state particles to simulate the finite momentum resolution of the LHCb detector.

The `RapidSim` package allows the kinematics of a potential background decay to be studied quickly, and then ruled out if it is removed by kinematic selection requirements. This saves a significant amount of CPU time compared to performing a full simulation with GAUSS. However, `RapidSim` can not be used to estimate signal efficiencies or determine fit models. Therefore, only fully simulated (with GAUSS) samples are used for these purposes.

3.2 GEANT4 Physics Lists

For any interaction process there is often a variety of physics models available within GEANT4 that can be used to perform the simulation. The model which offers best performance, in terms of accuracy and CPU time, is dependent on the geometry and particles being simulated. With the large variety of physical processes that are simulated and several models for each, there is a plethora of possible ways to simulate the same scenario. Therefore, GEANT4 provides standard combinations of models known as "physics lists". Each Physics List (PL) is a complete set of models chosen to give optimal simulation performance for a common physics scenario e.g. LHC physics, medical physics or low energy simulations.

The electromagnetic PL used by LHCb, `EMLHCb` (the hadronic models are specified separately) is a custom version of a standard PL designed for HEP applications, known as `EM standard option1`. A key feature of this PL is the use of the `Minimal` option for multiple scattering step limitation, which is often the limiting factor for the length of step taken. The multiple scattering step limit, L, in GEANT4 is defined as

$$L = \min\left[F_r \times \max\left(R\left(T\right), \lambda\left(T\right)\right), F_s \times s, D/F_G\right], \tag{3.1}$$

where F_r is a tunable range factor, $R(T)$ is the particle range, $\lambda(T)$ is the inverse transport cross section, T is the kinetic energy of the particle, F_s and s are geometrical safety factors, D is the distance to a geometrical boundary and F_G is a tunable geometrical factor [95]. With the use of the `Minimal` step limitation option the value of F_r is changed from the default value of 0.04 to 0.2. The advantage of this is that far fewer steps are taken by GEANT4, which leads to a significant reduction in CPU time. The LHCb experiment simulates of the order of 10^{10} events per year, therefore this reduction in CPU time is crucial. The disadvantage of using the `Minimal` option, however, is a reduction in accuracy and a lack of stability against production cuts [96].

During the simulation process many secondary particles can be created, especially in electromagnetic showers. Without production cuts these would all be fully tracked and simulated, despite many low energy particles having negligible effect on the physics results. This would be a very computationally expensive process. Therefore production cuts are applied, which require any secondary particles to have enough energy to travel a minimum distance in order for them to be tracked. In LHCb simulations a production cut of 5 mm is used, which vastly reduces the computation time required to simulate each event. The main customisation made to the `EM standard option1` PL for use by LHCb is the removal of production cuts for secondary particles produced in the photoelectric effect, Compton scattering, gamma conversion and positron annihilation electromagnetic processes. This customisation is made in order to increase the accuracy of the simulation of the RICH sub detectors. A detailed study of the effect of production cuts on the number of hits simulated in several LHCb sub-detectors can be found in Ref. [97].

3.3 Calorimeter Test

The simulation of the LHCb electromagnetic calorimeter (ECAL) is particularly important for any physics results with electrons or photons in the final state. With the advent of the anomalies seen in lepton flavour universality tests, where the ratio of decay rates of b mesons to muons and electrons is measured, the simulation of the ECAL is going to become increasingly important [86, 98]. This calorimeter test benchmarks the simulation of the ECAL, as performed by GEANT4 without introducing the full complexity of the detector as modelled in GAUSS, allowing early comparisons to be made when a new version of GEANT4 is adopted. By making these comparisons any unexpected changes to the simulation results can be accounted for at an early stage.

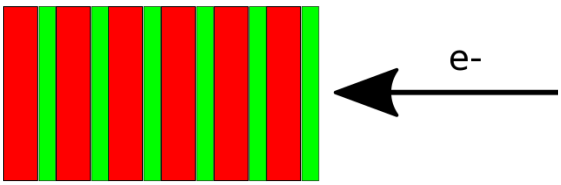


Figure 3.1: The geometry of the model calorimeter simulated by the calorimeter test. The red represents active plastic scintillator layers and the green represents passive lead layers.

The scenario simulated by the calorimeter test is electrons fired into the front face of a model calorimeter at normal incidence. The model calorimeter is a sampling calorimeter consisting of 66 alternating layers of both lead and plastic scintillator; the lead layers are 2 mm thick and the scintillator layers are 4 mm thick. A diagram showing this geometry is given in Figure 3.1. This geometry is deliberately chosen to model the design of the LHCb ECAL as closely as possible [73].

The main aim of this test is to use calorimeter resolution as a metric for comparison. The fractional resolution of an electromagnetic sampling calorimeter can be parameterised, in the absence of electronic noise, as

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} \oplus C, \tag{3.2}$$

where $\frac{\sigma}{E}$ is the fractional resolution of the calorimeter, A and C are free parameters of the model and $\oplus$ represents addition in quadrature. By determining the values of A and C for this model calorimeter, quantitative comparisons of different GEANT4 versions can be made.

The A term in Equation (3.2) arises from statistical fluctuations in the electromagnetic shower induced by the sampling calorimeter. In an electromagnetic shower, many particles are produced and the energy measured by the calorimeter is the sum of the energies deposited by each particle. However in a sampling calorimeter only a fraction of the shower takes place in active regions, therefore only a fraction of the shower particles are actually measured. Consequently the number of shower

particles that are measured is subject to Poisson sampling fluctuations with a standard deviation of $\sqrt{N}$, where N is the number of shower particles in the active region. Assuming the calorimeter response is linear (which it should be) the number of shower particles produced is proportional to energy, therefore $\frac{\sigma}{E} \propto \frac{\sqrt{N}}{N} \propto \frac{1}{\sqrt{E}}$. In a sampling calorimeter it is sampling fluctuations that typically dominate the resolution except at very high energies, thus the A term is usually by far the largest term in Equation (3.2).

The C term in Equation (3.2) arises due to shower leakage from the calorimeter. At a given energy the amount of energy lost from the calorimeter is subject to event-by-event fluctuations, which leads to broadening of the resolution. However, the total amount of energy lost from the calorimeter is proportional to the energy of the incident particle and this dominates over the event-by-event effects. This leads to the standard deviation of measured energies due to shower leakage being proportional to incident energy, and consequently energy independent for the fractional resolution.

In order to extract the fractional resolution as a function of energy, electrons are fired into the calorimeter at a range of energies. At each energy, the distribution of the total energy deposited in scintillator layers is stored and a Gaussian fit to this distribution is performed. An example of one of these fits is shown in Figure 3.2. The fractional resolution at each incident energy is then taken as $\frac{\sigma}{\mu}$, where σ and μ are the standard deviation and mean, respectively, extracted from the Gaussian fit. A minimum χ^2 fit of the function shown in Equation. (3.2) is then performed, with the values of A and C left free to vary.

3.3.1 Calorimeter Test Results

Three versions of GEANT4 have been compared whilst using the `EMLHCb` PL, namely GEANT4 v9.5.2, v9.6.4 and v10.3.3. The fractional resolution as a function of energy

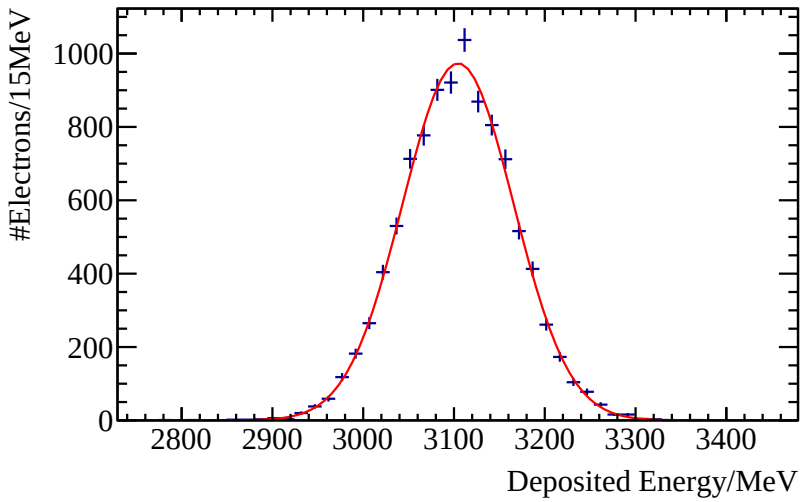


Figure 3.2: The distribution of total energy deposited in the scintillator layers of the ECAL at an incident electron energy of 25 GeV. The Gaussian fit performed to extract the standard deviation of the distribution is overlayed.

GEANT4 version	A	C	SF @ 25 GeV
v9.5.2	$(9.50 \pm 0.02)\%$	$(0.474 \pm 0.007)\%$	0.12417 ± 0.00003
v9.6.4	$(9.10 \pm 0.02)\%$	$(0.478 \pm 0.006)\%$	0.13149 ± 0.00003
v10.3.3	$(9.10 \pm 0.02)\%$	$(0.477 \pm 0.006)\%$	0.13988 ± 0.00003

Table 3.1: Results from fits of fractional resolution against energy. A and C are the free parameters of the model given in Equation (3.2).

for all three versions can be seen in Figure 3.3, the numerical results of the fits are shown in Table 3.1. It is clear there are significant differences between GEANT4 versions v9.5.2 and v9.6.4; there was a step change in the stochastic (A) term between these two versions. However, there was no change in the stochastic (A) term between GEANT4 v9.6.4 and v10.3.3. The statistical significance of the discrepancy between v9.5.2 and v9.6.4 is 20.8σ; there is no doubt this discrepancy is due to changes to the modelling of the calorimeter within GEANT4.

This has been investigated further by studying the average Sampling Fraction (SF) of the model calorimeter for each version of GEANT4. In this case, the SF is defined as the fraction of the incident electron's energy which is deposited in scintillator layers. This has very little dependence on the energy of the incident electron, therefore

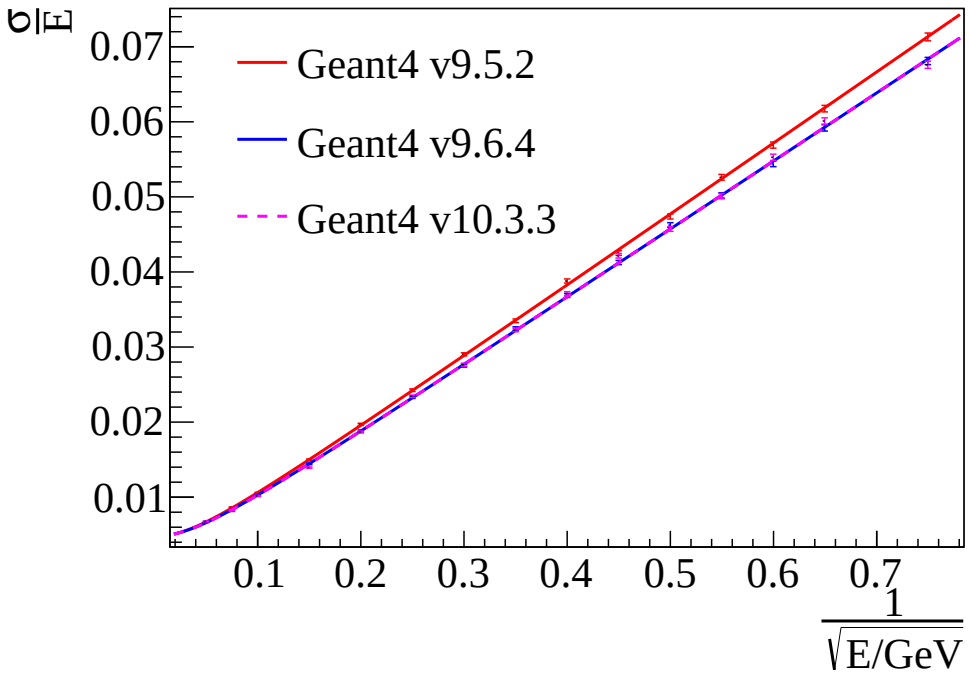


Figure 3.3: Plot of fractional resolution against $1/\sqrt{E}$ for GEANT4 v9.5.2, v9.6.4 and v10.3.3.

the SF at an electron energy of 25 GeV is chosen for comparison, these results are also shown in Table 3.1. The SF results have progressively increased with each new version of GEANT4.

The best way to assess whether these changes are an improvement is to compare the simulated resolution to that obtained in data. However, the test beam data available for the LHCb ECAL only gives a range of values for the A term of the fractional resolution, $8.5\% < A < 9.5\%$ [73]. Furthermore, only a simplified scenario and reconstruction is simulated in this test. For these reasons, a meaningful comparison with data is not possible. Instead, a comparison with simulations performed with the `emstandard_opt0` PL is made. This PL provides optimal accuracy and stability against production cuts, but cannot be used in production due to the significantly larger amount of CPU time it requires. When this test is run with the GEANT4 v10.3.3 and the `emstandard_opt0` PL the fractional resolution results obtained are:

$A = 7.84 \pm 0.02\%$ and $C = 0.459 \pm 0.005\%$. Very similar results are found when GEANT4 v9.5.2 is used with the same PL. Therefore, the change in calorimeter resolution observed is progress towards optimal accuracy with the `EMLHCb` PL whilst keeping CPU time roughly the same.

The consequence of these changes is that the calorimeter calibration in GAUSS has to be reviewed and potentially updated every time a new version of GEANT4 is introduced. Even though the resolution has not changed between v9.6 and v10.3, the change in SF means a re calibration is necessary. This ensures that the energy deposits simulated by GEANT4 can still be used to reconstruct the energy of the incident particle.

3.4 Multiple Scattering Test

After consultation with the GEANT4 authors, it was suggested the most likely reason for the observed resolution changes is alterations to the multiple scattering models. Therefore, a dedicated test of multiple scattering was needed. It has been seen that the modelling of the multiple scattering process has a strong influence on the accuracy of the Impact Parameter (IP) resolution. As the IP is used extensively to isolate secondary vertices in the LHCb detector, it is crucial the modelling of the IP is not degraded.

When a charged particle traverses material there is a non-zero probability that it will undergo elastic Coulomb scattering from a nucleus within the material. The differential cross section for this process is given by

$$\frac{d\sigma}{d\Omega} = \left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{z^2 e^4}{M^2 c^4 \beta^4} \frac{1}{\sin^4(\theta/2)}, \tag{3.3}$$

where Ω is solid angle, z is the atomic number of the material, M is the mass

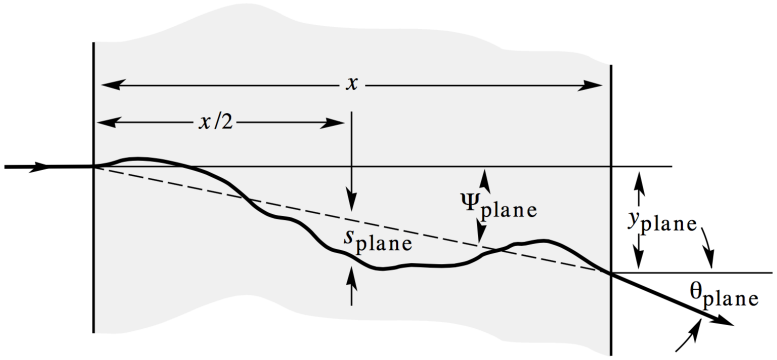


Figure 3.4: The lateral displacement and angular dispersion when a charged particle traverses a medium [1].

of the charged particle and θ is the angle through which the charged particle is scattered [99].

Except for cases where the scattering material is a very thin film, the charged particle will scatter multiple times before exiting the material. Hence, multiple coulomb scattering, which is more commonly known as just multiple scattering, occurs. The net effect is a lateral displacement as well as a scattering angle, as depicted in Figure 3.4. In this case a statistical treatment has to be used to obtain a distribution for the scattering angle, which is defined as θ in Figure 3.4. One such statistical treatment is Molière theory, which has been shown to give very good agreement with data over a wide range of particles, materials and energies [100, 101]. Several other theories have been shown to produce consistent results; Lewis theory also provides moments for the spatial displacement distribution [102]. Both the Moliere and Lewis theories give a scattering angle distribution that is Gaussian for the central 98% of scattering angle values, but the tails of the distribution fall off more slowly than a Gaussian function due to the $1/\sin^4(\theta/2)$ term in Equation 3.3. The width of the central Gaussian is defined as θ_0 which can be approximated by the Highland

formula

$$\theta_0 = \frac{14.1\,\text{MeV}}{pv} z \sqrt{\frac{L}{L_R}} \left[1 + \frac{1}{9} \log_{10} \left(\frac{L}{L_R}\right)\right], \tag{3.4}$$

where p is the incident particle's momentum, v is the incident particle's velocity, L is the length of the material and L_r is the radiation length of the material. This formula is an empirical formula that arises from fits to Molière theory [103].

When one simulates multiple scattering, in a similar way to the theoretical models, it is rarely possible to simulate every individual collision. It is only possible if the number of scatters is small and a large amount of CPU time is available. For the latter reason, this type of simulation based on simulating every single scatter was not implemented until 2005. Unfortunately, there is still a limited number of applications where this is a viable option. The `UrbanMsc` models get around this by using a "condensed" simulation of multiple scattering, which involves simulating one step of the particle's path at a time and applying net effects at the end of each step [95]. More specifically, the angle through which the particle has been scattered and the lateral displacement are applied at the end of each step as part of the multiple scattering simulation. The scattering angle is sampled from distributions calculated using Lewis theory but no theory of a full displacement distribution exists. Therefore, GEANT4 uses its own, approximate, algorithms to calculate the lateral displacement after each step [104].

Another approach to simulating multiple scattering, which has been implemented more recently, is to use a "mixed" approach. This involves sampling scatters, where the scattering angle θ is below a threshold θ_{max}, in a similar way to the "condensed" approach discussed previously. However, if the scattering angle is above θ_{max} then a single scattering approach is used. This is implemented in GEANT4 as the `WentzelVI` model.

To carry out a direct investigation of multiple scattering, a test based on an example provided by the authors of GEANT4 was setup to fire particles into a square sheet

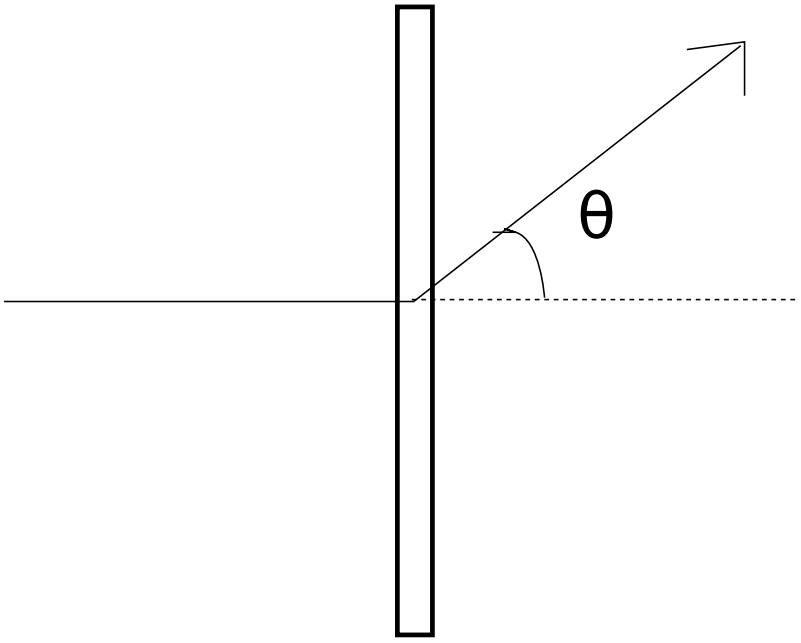


Figure 3.5: The setup of the multiple scattering test. θ is the angle under investigation by this test.

of material at normal incidence and study the angle of the scattered particle to the normal as it exits the material on the opposite side. A diagram of this is shown in Figure 3.5. The type of material used, the width of the material and the thickness of the material can all be specified. In this case the setup used was a 300μm thick sheet of silicon designed to model the LHCb VELO as closely as possible. This is the area of the detector where precise tracking measurements sensitive to multiple scattering take place; the IP is largely dictated by measurements in the VELO.

The aim of this test is to measure the parameter θ_0 of the scattered particles' angular distribution for electrons at a range of energies and use it as a metric with which to compare GEANT4 versions. The θ_0 parameter is then estimated at each energy by calculating the standard deviation of the central 98% of scattering angles, effectively measuring the width of the central Gaussian component of the scattering angle distribution. In order to estimate an uncertainty on the θ_0 parameter the test is re-run 1000 times at each energy, the mean of the resulting θ_0 values is used for comparison and the uncertainty on the mean is assigned as the RMS/$\sqrt{1000}$.

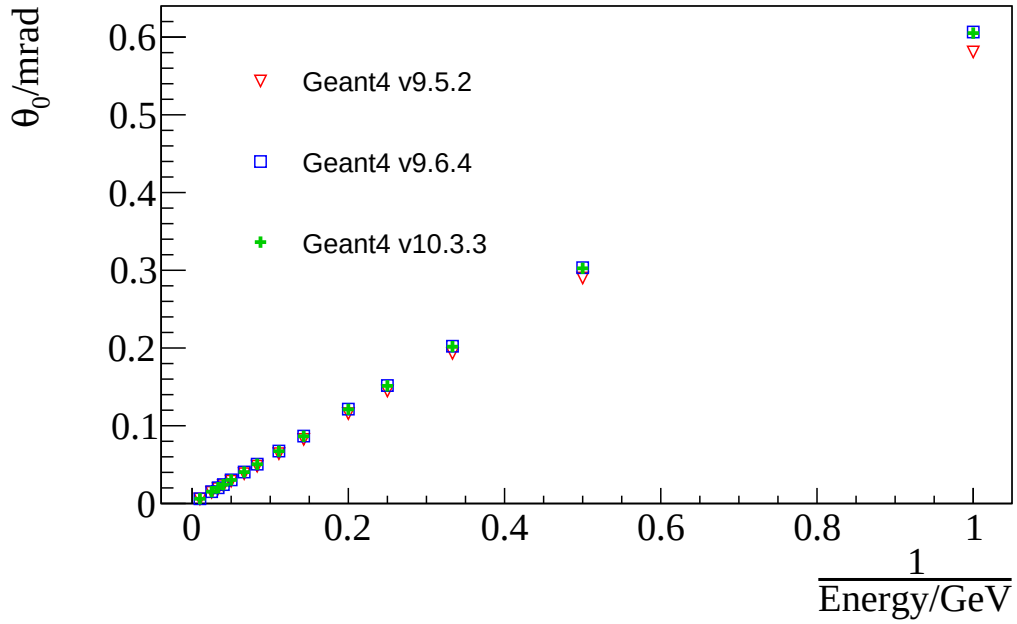


Figure 3.6: θ_0 as a function of the inverse of the incident electron energy.

3.4.1 Multiple Scattering Test Results

This test has been performed at 14 energies in the range 1 – 100 GeV for the three versions of GEANT4 introduced in Section 3.3 (with the `EMLHCb` PL): v9.5.2, v9.6.4 and v10.3.3. The θ_0 parameter as a function of the inverse of the incident electron energy can be seen in Figure 3.6, and the numerical results can be found in Table 3.2. These results show, firstly, that there is no change in the scattering angle between GEANT4 v9.6.4 and v10.3.3. However, it does show that there was a difference in scattering angle between GEANT4 v9.5.2 and v9.6.4. This change is prevalent at low energies. The most likely reason for this change is the fact that the multiple scattering model used for electrons and positrons above 100 MeV changed from the `UrbanMsc95` model in v9.5.2 to the `WentzelVI` model in v9.6.4. This observation meant that particularly close attention was paid to reconstructed physics quantities which are sensitive to multiple scattering, such as the IP resolution, when the validation of the full simulation package was carried out.

Energy/ GeV	θ_0/mrad		
	v 9.5.2	v9.6.4	v10.3.3
1	0.5809 ± 0.0031	0.6065 ± 0.0031	0.6051 ± 0.0029
2	0.2901 ± 0.0015	0.3033 ± 0.0015	0.3025 ± 0.0015
3	0.1931 ± 0.0009	0.2023 ± 0.0010	0.2017 ± 0.0010
4	0.1447 ± 0.0008	0.1517 ± 0.0008	0.1512 ± 0.0007
5	0.1156 ± 0.0006	0.1214 ± 0.0006	0.1210 ± 0.0006
7	0.0824 ± 0.0004	0.0867 ± 0.0004	0.0865 ± 0.0004
9	0.0641 ± 0.0003	0.0674 ± 0.0004	0.0672 ± 0.0003
12	0.0480 ± 0.0003	0.0506 ± 0.00025	0.0504 ± 0.0003
15	0.0383 ± 0.0002	0.0405 ± 0.00020	0.0404 ± 0.0002
20	0.0287 ± 0.0002	0.0304 ± 0.00015	0.0303 ± 0.0002
25	0.0229 ± 0.0001	0.0243 ± 0.00012	0.0242 ± 0.0001
30	0.0191 ± 0.0001	0.0202 ± 0.00010	0.0202 ± 0.0001
40	0.01431 ± 0.00008	0.01524 ± 0.00008	0.0151 ± 0.0001
100	0.00577 ± 0.00003	0.00607 ± 0.00002	0.00605 ± 0.00003

Table 3.2: Results from the multiple scattering test, showing θ_0 at various energies for the three GEANT4 versions tested.

3.5 Conclusions

The standalone benchmarking of GEANT4 that has been described in this Chapter has highlighted several differences between the three GEANT4 versions tested. The sampling calorimeter test has seen a change of around 4% in the fractional resolution of a simplified model of the LHCb calorimeter between GEANT4 v9.5.2 and v9.6.4. A change in SF has also been observed between all three GEANT4 versions. The change between v9.5.2 and v9.6.4 is strongly believed to be caused by the change from a "condensed" multiple scattering model to a "mixed" model for electrons and positrons with an energy above 100 MeV. The consequence of these changes is that the calorimeter calibration has to be re-performed in MC with each release of a new simulation package. Changes to the multiple scattering models have been probed directly with the test described in Section 3.4. This showed a change in the scattering angles of electrons in thin sheets of silicon, particularly at low energies, between GEANT4 v9.5.2 and v9.6.4. However, no change was observed between v9.6.4 and v10.3.3. The observation of this change in the behaviour of GEANT4

meant that extra care was taken to validate reconstructed physics quantities which are sensitive to multiple scattering.

The objective of these tests is to perform early sanity checks when a new version of GEANT4 is released (rather than tuning the MC), and the discovery of the necessity to re-calibrate the calorimeter in MC shows that these tests are capable of achieving this objective. Without these tests the change in SF would have either gone undetected or, if reconstructed quantities behaved differently, it would have taken significant effort to pin down the cause of any discrepancy. In the latter case, a significant amount of CPU time could also be wasted regenerating MC samples which have been simulated with the wrong calorimeter calibration. To ensure that these tests are always run efficiently in the future, they have been integrated into the LHCb Performance and Regression testing platform (LHCbPR) [105]. This means the tests are automatically run periodically on a dedicated server and the results can be viewed on a web page, allowing any changes or discrepancies to be flagged up at the earliest opportunity.

CHAPTER 4

Analysis Strategy in the Search for $\Lambda_b^0 \to pK\eta'$

This chapter describes the strategy used to search for the decay[1] $\Lambda_b^0 \to pK\eta'$, and either measure or set a limit on its branching fraction. Firstly, an overview of the strategy is outlined. This is followed by a description of the event selection, efficiency calculations and background studies.

4.1 Introduction

The aim of this analysis is to perform a "blind" search for the rare and as-yet unobserved decay $\Lambda_b^0 \to pK\eta'$ in a combination of two channels; where the η' is reconstructed through the decay $\eta' \to \pi^+\pi^-\gamma$ and where the η' is reconstructed through the decay $\eta' \to \pi^+\pi^-\eta$ $(\eta \to \gamma\gamma)$. These two channels are used because they

[1]The search is also performed for the charge conjugate decay $\overline{\Lambda}_b^0 \to \overline{p}K^+\eta'$; the inclusion of charge conjugate decays is implied throughout this thesis

are the decay modes with the highest branching fractions that can be reconstructed at LHCb. The only other decay mode with a comparable branching fraction is $\eta' \to \pi^0\pi^0\eta$, but reconstruction of the η' through purely neutral particles is not feasible with the LHCb detector. The branching fraction of the decay $\Lambda_b^0 \to pK\eta'$ will be measured, or a limit set, relative to the well-known decay $B^+ \to K^+\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) [1]. The ratio of branching fractions is measured, as opposed to an absolute branching fraction, because many systematic uncertainties cancel when the ratio is taken. For example, the $b\bar{b}$ production cross section has a $\sim 10\%$ systematic uncertainty which cancels in the ratio [106].

The $B^+ \to K^+\eta'$ decay has a branching fraction of $(7.06 \pm 0.25) \times 10^{-5}$ [1], which means a signal yield of $>$ 10K events is expected. Furthermore only minimal combinatorial background has been seen in other LHCb analyses of this decay, making it an ideal control channel for this search [107, 108]. Combinatorial background occurs when random combinations of tracks and photons are combined to create fake signal candidates. It can usually be described by smooth, non-peaking, functions which makes it simple to model. The decay $B^0 \to K^{*0}\eta'$ was originally investigated as a potential control channel, motivated by its greater similarity to the number of tracks in the rare decay signal final state, but it was found to suffer from significant peaking backgrounds in the signal region. Furthermore, the lower branching fraction ($\mathcal{B} = (3.1 \pm 0.9) \times 10^{-6}$) leads to an estimated signal yield of $<$200 events [1].

The data samples used in this search are the full Run I dataset. This consists of data taken at $\sqrt{s} = 8\,\text{TeV}$ in 2012 corresponding to an integrated luminosity of $2.0\,\text{fb}^{-1}$ and data taken at $\sqrt{s} = 7\,\text{TeV}$ in 2011 corresponding to an integrated luminosity of $1.0\,\text{fb}^{-1}$. The data samples for both the rare channels and the control channels are processed during the same processing campaign, meaning the same reconstruction software is used in both cases. This is important to ensure any mis-modelling of reconstructed variables in simulation cancels in the branching fraction ratio.

An event selection is developed to maximise sensitivity to the rare Λ_b^0 decays, and as similar selection as possible is applied to the control channel. This is developed with the analysis blind; events with an invariant mass in the range[2] $5494\,\text{MeV} < M(\Lambda_b^0) < 5744\,\text{MeV}$ ($5444\,\text{MeV} < M(\Lambda_b^0) < 5794\,\text{MeV}$) in the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) rare channel are concealed until the selection and fit models are finalised. The selection is described in detail in Section 4.2, but consists of: using hadronic and multi-body topological trigger lines; requirements on the quality of reconstructed photons; the use of a multivariate binary classifier; requirements which make use of the LHCb PID system and mass vetoes for specific (mainly charm) backgrounds.

The signal yields are then extracted with a simultaneous extended maximum likelihood fit to the invariant mass of Λ_b^0 candidates and B^+ candidates. In the latter case (control channel), the fit is also performed to the mass of the η' candidates. A simultaneous fit across all three decay channels is used so that parameters can be shared between decay channels, which improves the stability of the fit. This is described in detail in Section 5.1.

The ratio of branching fractions, which it is the aim of this analysis to determine, is given by

$$R = \frac{\mathcal{B}(\Lambda_b^0 \to pK\eta')}{\mathcal{B}(B^+ \to K^+\eta')} = \left(\frac{\epsilon_c N_\gamma}{\epsilon_\gamma N_c} \left(\frac{f_u}{f_{\Lambda_b^0}} \right)_\gamma + \frac{\epsilon_c N_\eta}{\epsilon_\eta N_c} \left(\frac{f_u}{f_{\Lambda_b^0}} \right)_\eta \right) \times \frac{\mathcal{B}_\gamma}{\mathcal{B}_\gamma + \mathcal{B}_\eta}, \qquad (4.1)$$

where:

- ϵ_c and N_c are the total efficiency and the signal yield of the control channel,
- ϵ_γ and N_γ are the efficiency and signal yield for the rare channel where $\eta' \to \pi^+\pi^-\gamma$,
- ϵ_η and N_η are the efficiency and signal yield for the rare channel where $\eta' \to$

[2]The invariant mass resolution of Λ_b^0 candidates in the $\eta' \to \pi^+\pi^-\eta$ channel is worse, hence the different blind regions

$\pi^+\pi^-\eta$,

- $\mathcal{B}_\gamma = \mathcal{B}\ (\eta' \to \pi^+\pi^-\gamma\ ) = 0.291 \pm 0.005$ [1],
- $\mathcal{B}_\eta = \mathcal{B}\ (\eta' \to \pi^+\pi^-\eta\ ) \times \mathcal{B}(\eta \to \gamma\gamma\ ) = 0.169 \pm 0.007$ [1],
- $\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_{\gamma(\eta)}$ is the B^+/Λ_b^0 fragmentation fraction in the $\eta' \to \pi^+\pi^-\gamma(\eta' \to \pi^+\pi^-\eta)$ channel as measured by LHCb but determined separately for each channel (see Section 4.1.1) [109]. This is the fraction of b-quarks, produced in the proton-proton collision, which hadronise to produce a B^+/Λ_b^0 hadron.

As shown in Equation. (4.1), the efficiencies of each decay channel are required to measure the ratio of branching fractions. In the rare channels, the M(pK^-) spectrum is expected to contain a rich resonant structure which is not *a priori* understood [16], but needs to be taken into account when the efficiencies are determined. As no amplitude model for these resonances exists at the time of writing, the simulated samples used in this analysis are generated uniformly across the phase space of the decays(also refered to as just "phase space" for brevity). The use of phase space Monte Carlo to determine efficiencies is not guaranteed to be accurate as there are likely to be variations in efficiency across the phase space of the decay, and intermediate M(pK^-) resonances would lead to a non-uniform population of the phase space. Corrections for the variation of efficiency across the phase space of the decay need to be applied, which leads to a two step unblinding procedure. Prior to unblinding phase space dependent efficiencies will be determined as a function of the variables m'' and $\cos(\theta_{\eta'p})$, which are justified in Section 4.3. m'' is given by

$$m'' = \frac{m_{\eta'p} - m_{\eta'p}^{min}}{m_{\eta'p}^{max} - m_{\eta'p}^{min}}, \tag{4.2}$$

where $m_{\eta'p}$ is the invariant mass of the p and η' system, $m_{\eta'p}^{min} = m_{\eta'} + m_p$ and $m_{\eta'p}^{max} = m_{\Lambda_b^0} - m_{K^-}$. $\cos(\theta_{\eta'p})$ is the cosine of the helicity angle of the $\eta'p$ system, which is the angle between the η' and the K^- in the rest frame of the $\eta'p$ system.

The simultaneous mass fit (combined fit to both the Λ_b^0 decay channels and the B^+ decay channel) will first be performed to extract raw signal yields and Wilks' theorem [110] will be used to decide if the yield in either channel separately has a significance greater than 3σ.

In any channel where this is the case, a background subtraction will be performed using the *sPlot* method and the resulting *sWeights* will be used to calculate a phase space corrected efficiency [111].

When there is more than one source of events contributing to the distribution of a discriminating variable m, the *sPlot* method allows these different sources to be statistically separated. After a maximum likelihood fit to the discriminating variable is performed, the *sWeight* for a source s and an event e is given by

$$sW_s(e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{sj} f_j\left(m_e\right)}{\sum_{k=1}^{N_s} N_k f_k\left(m_e\right)}, \tag{4.3}$$

where N_s is the number of sources of events, $\mathbf{V}$ is the covariance matrix of the fit, f_x is the Probability Density Function (PDF) of source x evaluated at the value of m for event e (m_e) and N_k is the event yield of source k. *sWeights* are constructed such that

$$N_n M_n(\bar{y})\delta y \equiv \sum_{e \in \delta y} sW_s(e), \tag{4.4}$$

where N_n is the total number of events in the data sample, M_n is the true distribution of a variable of interest y for the source s, $\bar{y}$ is the value of y at the centre of a bin with width δy and $\sum_{e\in\delta y} sW_s(e)$ is the sum of *sWeights* for events falling in a given bin of y. In practice this means plotting the distribution of the variable y weighted by the *sWeights* for a source s will produce, on average, the true distribution of the variable y for the source s. In this analysis the reconstructed masses of Λ_b^0 candidates will be used as discriminating variables to determine the true distribution of the phase space variables m'' and $\cos(\theta_{\eta' p})$ for signal decays.

One caveat that comes with the use of *sWeights* is that all yields must be free in the fit from which the covariance matrix $\mathbf{V}$ is extracted. This means that the yields of background components can not be fixed or constrained. Consequently, the complexity of fit possible whilst still maintaining stability is significantly limited.

In any channel where the significance is $< 3\sigma$, the signal yield is unlikely to be large enough to perform a correction. Therefore, the phase space integrated efficiency will be calculated and a systematic uncertainty will be assigned for the variation of the efficiency over the phase space of the decay.

A second simultaneous mass fit will then be performed with an identical fit model but with the efficiency and fragmentation fraction information included such that the ratio of BFs, as given in equation 4.1, is a parameter of the fit which is free to vary. This allows the ratio of BFs, R, and the corresponding likelihood function to be extracted directly from the fit. In the case of observing a combined significance $< 3\sigma$, upper limits will be set on the BF by convolving the likelihood function with a Gaussian of width equal to the total systematic uncertainty and integrating in the physical region.

4.1.1 Λ_b^0/B^+ fragmentation fraction

The ratio of fragmentation fractions $\frac{f_{\Lambda_b^0}}{f_u}$ used in Equation (4.1) is assumed to be identical to $\frac{f_{\Lambda_b^0}}{f_d}$ under isospin symmetry, which has been measured by LHCb [109]. The LHCb measurement shows that $\frac{f_{\Lambda_b^0}}{f_d}$ has a pseudorapidity dependence that can be described by the equation [109],

$$\frac{f_{\Lambda_b^0}}{f_d} = (0.387 \pm 0.033) + (0.067 \pm 0.013)(\eta - \bar{\eta}), \tag{4.5}$$

and $\bar{\eta} = 3.198$. Although this only depends on Λ_b^0 production kinematics, the different decay products mean the η distribution measured by the LHCb detector

	$\eta' \to \pi^+\pi^-\gamma$	$\eta' \to \pi^+\pi^-\eta$
$< \eta >$	3.6273 ± 0.0004	3.6697 ± 0.0003
$\frac{f_{\Lambda_b^0}}{f_d}$	0.416 ± 0.033	0.419 ± 0.033

Table 4.1: The mean pseudorapidity and fragmentation fraction used for each Λ_b^0 decay channel, averaged over both years of data taking. See text for a description of how these are calculated.

will be slightly different in each channel. Therefore, separate fragmentation fractions are used in each channel.

In order to calculate these fragmentation fractions simulation of the signal decays without the effects of the detector ("generator level"), but with the requirement that the decay products are within angular acceptance of the LHCb detector[3] is used to calculate a mean pseudorapidity, $< \eta >$, in each decay channel. The mean pseudorapidity is averaged over both years of data taking and then used in Equation (4.5) to calculate a fragmentation fraction for each channel; these values are shown in Table 4.1. It should be noted that the approximately 8% uncertainty on these fragmentation fractions is expected to be one of the dominant systematic uncertainties on the branching fraction measurement.

4.1.2 Global Decay Chain Fitting

The default approach to determining quantities such as the B^+/Λ_b^0 kinematics and decay time is to start from the measurements of the final state particles and propagate the kinematics up the decay chain to the mother particles. Explicitly, in the case of the $\Lambda_b^0 \to pK\eta'$ decay, the kinematics of the η' resonance are determined by summing the four vectors of the daughter particles and then the Λ_b^0 kinematics are calculated by summing the four vectors of the p, K^- and η'. However, this means that none of the information from upstream of the final state particles is

[3]Subsequently, this acceptance requirement will be referred to as satisfying the `DaughtersInLHCb` criteria.

used when calculating the kinematics of any intermediate resonances. In order to improve the resolution on quantities of interest, such as the masses and lifetimes of parent particles, the entire decay chain is parameterised using quantities such as vertex positions, track kinematics and calorimeter clusters. A least squares fit to the entire decay chain is then performed using a Kalman filter [88], known as Decay Tree Fitter (DTF). This approach also allows additional physics constraints to be incorporated, such as constraining the mass of any intermediate resonances to the PDG value and the requirement that the mother particle originated at the primary vertex.

In all channels the decay chain is refitted with the DTF tool, with the η' mass constrained to the PDG value and the mother particle constrained to originate from the Primary Vertex (PV) of the Proton-Proton (pp) collision. In the case of the $\eta' \to \pi^+\pi^-\eta$ channel, it is also possible to constrain the mass of the η to the known value [1], but this additional constraint is not used because it causes a high fraction of fits to fail. The main benefit of this is a significantly improved B^+/Λ_b^0 mass resolution, but the χ^2 of the fit is also a powerful variable for discriminating between signal and background.

When the phase space variables (m'' and $\cos(\theta_{\eta' p})$) are calculated, DTF is further utilised. When the variables are determined using kinematics which are calculated without constraints on the decay chain, there are candidates which are reconstructed outside the allowed kinematic boundaries. This can lead to m'' outside the allowed range and erroneous values of $\cos(\theta_{\eta' p})$. Therefore, the entire decay chain is fitted with both the Λ_b^0 mass and η' mass constrained in addition to requiring the Λ_b^0 to originate from the PV. The daughter kinematics which result from this fit are then used to calculate m'' and $\cos(\theta_{\eta' p})$; the constraint on the Λ_b^0 mass ensures the kinematics stay within allowed regions of the phase space.

4.2 Selection

When pp collisions take place the large majority of the candidates reconstructed are background candidates, where other processes mimic the signal decays of interest to this analysis. It is the goal of the event selection to achieve the best expected signal significance by optimising the relative number of signal and background candidates selected.

4.2.1 Trigger Selection

In all channels, including the control channel, candidates are required to pass trigger decisions at all levels. As the primary purpose of the trigger system is to select which pp collision data to record, it is an event which passes a trigger selection rather than a candidate. Consequently, candidates can be split into two categories, Trigger Independent of Signal(`TIS`) and Trigger on Signal(`TOS`). When it is the physics objects (tracks, calorimeter clusters etc.) associated with a given candidate which caused an event to pass a trigger selection, the trigger decision is defined as `TOS`. When an event passes a trigger selection based on physics objects not associated with a candidate, the trigger decision is defined as `TIS`. A more detailed description of the LHCb trigger system and a description of the trigger lines used in this analysis can be found in Section 2.2.6.

At L0 candidates are required to pass either the `L0HadronDecision_TOS` or the `L0Global_TIS` trigger decision. The `L0HadronDecision` trigger line makes use of the hadronic calorimeter information to select events which contain at least one high transverse energy (E_T) hadron. The E_T threshold used is 3620 MeV in 2012 data and 3500 MeV in 2011 data. The `L0Global_TIS` requirement simply requires the candidate to have passed any other L0 trigger decision as TIS.

Candidates are also required to pass the `Hlt1TrackAllL0Decision_TOS` decision at the HLT1 level. This decision ensures there is at least one high p_{T} track in the event. More specifically, there has to be at least one track with $p_{\mathrm{T}} > 1.6\,\mathrm{GeV}$ present. At the HLT2 level candidates are required to pass either the `Hlt2Topo2BodyBBDTDecision_TOS`, `Hlt2Topo3BodyBBDTDecision_TOS` or `Hlt2Topo4BodyBBDTDecision_TOS` trigger decision. These trigger selections use a multivariate classifier, trained on topological variables, to select multibody decays. Further details about BBDT trigger decisions can be found in Ref. [77]. The overall efficiencies of these trigger requirements are $(43.29 \pm 0.21)\%$, $(38.65 \pm 0.13)\%$ and $(28.76 \pm 0.15)\%$ for the $B^+ \to K^+\eta'$, $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$ and $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ channels respectively. A full breakdown of the trigger efficiencies is given in Section 4.3.2.

4.2.2 Pre-Selection

The pre-selection consists of several loose cuts applied in order to reject background and select working data samples of a sensible size. These cuts, known as "Stripping" cuts, are largely all applied during the centralised data processing campaigns. As different "Stripping Lines" are used in the rare and control channels, the selection requirements differ for some variables. In particular, there are no requirements on the di-pion intermediate resonance in the rare channels. This is because the decay $\eta' \to \pi^+\pi^-\gamma$ (which is the only one used to reconstruct the control channel) predominantly proceeds through a ρ^0 resonance, whereas the decay $\eta' \to \pi^+\pi^-\eta$ does not.

Table 4.2(4.3) shows the pre-selection requirements for the rare (control) channel. The definitions of the variables used are:

- $\mathbf{p_T}$: Signal candidates have a harder transverse momentum (p_{T}) distribution than background candidates meaning a cut on p_{T} removes a lot of background.

- **ProbNN**: The information from the RICH PID systems and all other sub detectors is combined, with the use of a neural network, to produce a single PID variable for charged particles(see Section 2.2.4). The requirements used here are to make a positive identification of a track as being of the required particle species, which reduces background from mis-identified(mis-ID) tracks.

- **Ghost Probability**: "Ghost" tracks occur when random hits in the detector are reconstructed to form a track which is not associated with a physical particle. These are suppressed by applying the cut `ProbNN_Ghost` <0.5; the `ProbNN_Ghost` variable is the output of a neural network trained using information from several sub-detector systems.

- **IP χ^2**: The Impact Parameter (IP) is the transverse distance of closest approach between a track's trajectory and the PV. Units of χ^2 are equivalent to σ^2, therefore a requirement of IP χ^2 $>n$ is requiring the IP χ^2 to be greater than $\sqrt{n}$ standard deviations. The Λ_b^0/B^+ particles will travel a measurable distance in the detector, meaning all tracks should appear to originate from a displaced secondary vertex. Cuts on IP χ^2 are useful for rejecting tracks which were produced in the pp collision and requiring the Λ_b^0/B^+ candidate is consistent with being produced in the pp collision. The Best Primary Vertex (BPV) means the most likely PV; when BPV is specified it is the distance to the BPV, but otherwise it is the distance to any PV.

- **χ^2/ndf**: The χ^2/ndf is of the track fit. Cuts on χ^2/ndf are used to reject poorly reconstructed tracks and ghost tracks.

- **$\chi^2_{\mathbf{vtx}}$/DOF**: The χ^2/ndf of the vertex fit performed when tracks are combined.

- **DOCA χ^2**: The Distance of Closest Approach in units of χ^2, which is equivalent to standard deviations. This is calculated as the distance of closest approach between all possible pairs of particles. It is possible for background

Table 4.2: Pre-selection requirements used in the rare channels.

Particle	Variable	Requirement
Track	p_{T}	p/K^- >500 MeV, π^+/π^- >300 MeV
	Ghost Prob	< 0.5
	ProbNN	> 0.1
	BPV IP χ^2	> 20.0
	($p/$ K^- Only)	
	χ^2/ndf	< 3.0
γ	CL	>0.1
η	$m(\eta)$	$\pm$50 MeV
η'	p_{T}	>2000 MeV
	χ^2_{vtx}/DOF	<10.0
	DOCA χ^2	< 10
	$m(\eta')$	$\pm$100 MeV
Λ_b^0	χ^2_{vtx}/DOF	< 15
	DOCA χ^2	< 15
	p_{T}	> 1000 MeV
	DIRA	>0.9995
	BPV IP χ^2	<20
	$m(\Lambda_b^0)$	$\pm$750 MeV

candidates to be created from tracks which did not originate from the same vertex; requiring a minimum distance between them reduces this background.

- **DIRA**: The Direction Angle is the cosine of the angle between the Λ_b^0/B^+ momentum vector and the vector between the Λ_b^0/B^+ decay vertex and most likely PV. This is particularly effective at suppressing combinatorial background.

- $\mathbf{m_{corr}}$: The corrected B^+ mass which is defined as: $m_{\mathrm{corr}} = \sqrt{m^2 + |P_T^{miss}|^2} + |P_T^{miss}|$ where m is the invariant mass of the B^+ candidate and P_T^{miss} is the missing momentum transverse to the line of flight of the B^+ decay.

- γ **CL**: The confidence level of a reconstructed photon being a true photon. This variable is the output of a neural network trained to discriminate between real photons and background from both electrons and non-electromagnetic deposits. A full description can be found in Ref. [112].

Table 4.3: Pre-Selection requirements imposed on the control channel.

Particle	Variable	Requirement
Tracks	p_{T}	π^+/π^- >400 MeV
		K^- >1000 MeV
	Ghost Prob	< 0.5
	ProbNN	> 0.1
	(BPV) IP χ^2	> 16.0 (20.0)
γ	CL	>0.1
Di-Pion	χ^2_{vtx}/ndf	< 6.0
Resonance	p_{T}	>600 MeV
(ρ^0)	$m(\rho^0)$	<1100 MeV
η'	$m(\eta')$	$\pm$100 MeV
	p_{T}	>2000 MeV
B^+	χ^2_{vtx}/ndf	<6
	p_{T}	> 2500 MeV
	DOCA χ^2	< 15
	BPV IP χ^2	< 20
	DIRA	> 0.9995
	m_{corr}	<7000 MeV
	$m(B^+)$	$\pm$750 MeV

4.2.2.1 Cross Checks with pre-selected $B^+ \to K^+\eta'$ events

This section describes the use of the control channel to investigate the modelling of pre-selection variables and the reliability of efficiencies in MC. The modelling of variables is investigated by performing a fit to the B^+ mass distribution of pre-selected candidates, extracting *sWeights*[4] and comparing weighted distributions in data to MC [111]. The fits performed are a simple 1D fit; for the purposes of this fit a $\pm 2\sigma$ cut is made around the η' mass, where $\sigma = 12.85$ MeV which is taken from a fit to MC. The signal shape is modelled with a sum of two Crystal Ball (CB) functions with opposite side tails, from here on referred to as a Double Crystal Ball (DCB) function. A CB function is a Gaussian function in its central region with a power law tail, which is incorporated to model bremsstrahlung radiation effects.

[4] *sWeights* are described in Section 4.1

Explicitly, a CB function is given by

$$\mathrm{CB}(t;n,\alpha,\sigma) = \mathcal{N}\cdot\begin{cases} \exp(-t^2/2\sigma^2) & \text{if } t/\sigma > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \left(\frac{n-\alpha^2}{|\alpha|} - \frac{t}{\sigma}\right)^{-n} \exp(-\alpha^2/2) & \text{if } t/\sigma \leq -\alpha, \end{cases} \tag{4.6}$$

where $t = m-\mu$, with m being invariant mass; σ is the width of the central Gaussian; α defines how far from the mean (μ) the power law tail starts and n is the index of the tail. $\mathcal{N}$ is the normalisation factor but, as the sum of two CB functions is used, a fit fraction $f = \frac{\mathcal{N}_1}{\mathcal{N}_1+\mathcal{N}_2}$ is defined. This controls the relative contribution of each CB. The polarity of α is defined to be different in each of the CB functions, enforcing opposite side tails, but the mean μ and σ are shared between both CB functions.

The background is modelled with a 2nd order Chebychev polynomial, which is given by,

$$\mathrm{P}(x;a,b) = \frac{3}{2\,(3-b)}\left(1 + ax + b\left(2x^2-1\right)\right), \tag{4.7}$$

where a and b are parameters of the fit which are free to vary. x is given by,

$$x = 2\frac{m_{K^-\pi^+\pi^-\gamma} - m_{\mathrm{min}}}{m_{\mathrm{max}} - m_{\mathrm{min}}} - 1, \tag{4.8}$$

where $m_{K^-\pi^+\pi^-\gamma}$ is the invariant mass of the B^+ candidate, m_{min} is the minimum of the fit range and m_{max} is the maximum of the fit range. The projections of this fit can be seen in Figure 4.1.

The variable comparisons made between background subtracted data and MC are shown in Figure 4.2. With the exception of the $M(\pi^+\pi^-)$ distribution, these all show reasonable agreement. The disagreement in the $M(\pi^+\pi^-)$ distribution has no effect on the efficiency calculations because the cut applied is $M(\pi^+\pi^-) < 1100\,\mathrm{MeV}$, which is above the endpoint of this distribution in both data and MC.

Another cross check made is a comparison of the expected and observed number of

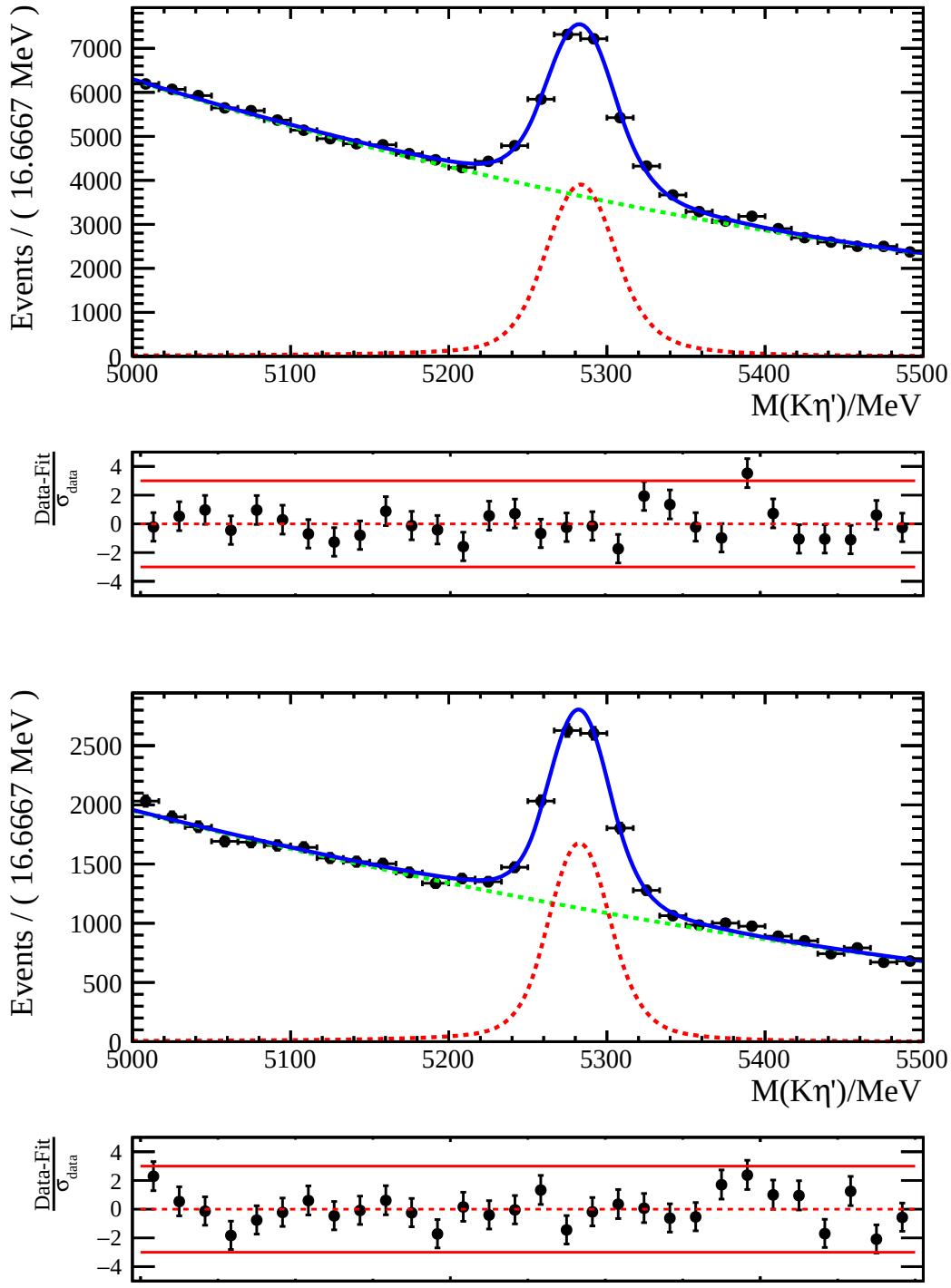


Figure 4.1: Mass distributions of control channel events passing the pre-selection. The Top(Bottom) plot is 2012(2011) data and the lower panel of each plot is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{fit}}}$, with the solid red horizontal lines representing $\pm 3\sigma$. The green dashed lined shows the background component, the red dashed line shows the signal component and the blue solid line shows the total fit function. The fit shown is used to extract *sWeights* and perform the cross checks described in Section 4.2.2.1.

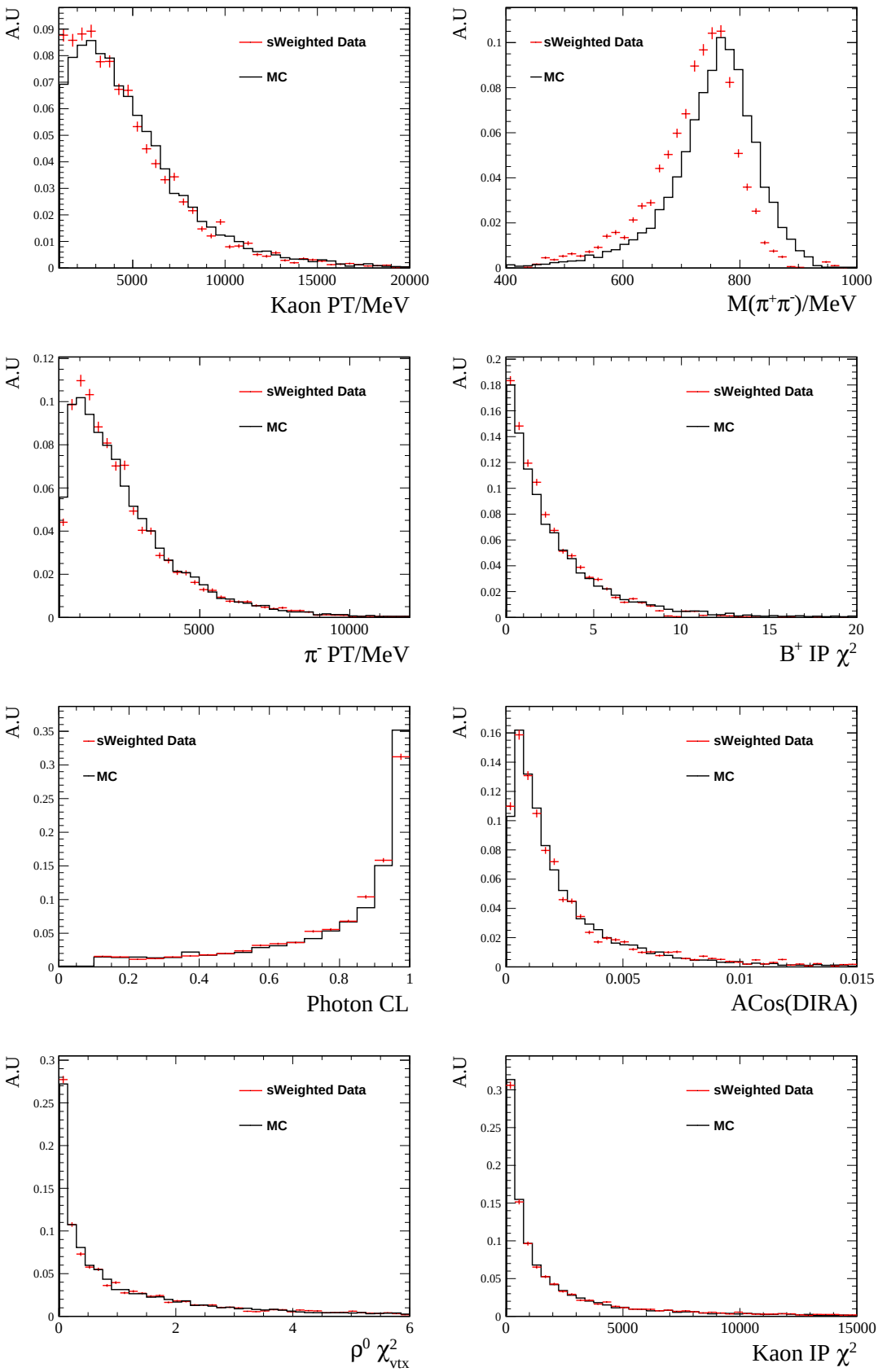


Figure 4.2: A comparison between 2012 MC and background subtracted data for pre-selected $B^+ \to K^+\eta'$ candidates. The same plots for 2011 data are shown in Appendix A.

$B^+ \to K^+\eta'$ events after the pre-selection. The observed yield in each year of data taking is extracted from the extended maximum likelihood fits shown in Figure 4.1. The expected yield can be calculated using Equation 4.9,

$$N^{\text{exp}}_{\text{presel}} = 2 \times \mathcal{L} \times \sigma(pp \to b\bar{b}) \times f_u \times \mathcal{B}(B^+ \to K^+\eta') \times \mathcal{B}(\eta' \to \pi^+\pi^-\gamma) \times \varepsilon_{\text{presel}} \quad (4.9)$$

where $\mathcal{L}$, $\sigma(pp \to b\bar{b})$ and $\varepsilon_{\text{presel}}$ are the integrated luminosity, $b\bar{b}$ production cross section and pre-selection efficiency respectively for a given year of data taking. f_u is the B^+ fragmentation fraction, which is the fraction of b quarks that hadronise with an up quark to create a B^+ meson.

The pre-selection efficiency, $\varepsilon_{\text{presel}}$, is factorised as follows,

$$\varepsilon_{\text{presel}} = \varepsilon_{geom} \times \varepsilon_{offline} \times \varepsilon_{trigger} \times \varepsilon_{PID} \quad (4.10)$$

where:

- ε_{geom} is the efficiency of requiring daughter particles to be within the LHCb acceptance at "generator level";
- $\varepsilon_{offline}$ is the combined efficiency of the candidate reconstruction and of all the cuts shown in Table 4.3 except the PID cuts;
- $\varepsilon_{trigger}$ is the efficiency of the trigger requirements described in Section 4.2.1;
- ε_{PID} [5] is the efficiency of the PID requirements applied to the charged hadrons.

The values of these efficiencies are shown in Table 4.4 along with the value of $\varepsilon_{\text{presel}}$.

All of these efficiencies are calculated using truth matched MC[6], except ε_{PID}; as outlined in Section 2.2.4 the PID variables are not well modelled in MC. Therefore,

[5] It should be noted that this PID efficiency is not used in the final branching fraction calculation as the efficiency of these loose cuts is assessed along with the tighter cuts applied later in the selection.

[6] Truth matching is performed by requiring: the true ID of all daughter particles and intermedi-

Table 4.4: Control channel pre-selection efficiencies used in the calculation of the expected number of events. Each efficiency is defined relative to the previous efficiency.

Efficiency	2012	2011
ε_{geom}	$(17.89 \pm 0.03)\%$	$(17.58 \pm 0.03)\%$
$\varepsilon_{offline}$	$(3.45 \pm 0.02)\%$	$(3.72 \pm 0.03)\%$
$\varepsilon_{L0Trigger}$	$(55.99 \pm 0.26)\%$	$(58.66 \pm 0.35)\%$
ε_{HLT1}	$(83.72 \pm 0.25)\%$	$(84.69 \pm 0.34)\%$
ε_{HLT2}	$(92.68 \pm 0.20)\%$	$(86.28 \pm 0.35)\%$
ε_{PID}	$(94.17 \pm 0.04)\%$	$(93.71 \pm 0.05)\%$
ε_{total}	$(0.2526 \pm 0.0021)\%$	$(0.2626 \pm 0.0030)\%$

Table 4.5: Comparison of expected and observed yields in the $B^+ \to K^+\eta'$ channel.

Value	2012	2011
$\mathcal{L}$	$2.057 \pm 0.072\,\text{fb}^{-1}$	$1.017 \pm 0.036\,\text{fb}^{-1}$
$\sigma(pp \to b\bar{b})$	$298 \pm 36\,\mu\text{b}$	$284 \pm 53\,\mu\text{b}$
f_u	$40.5 \pm 0.6\%$ [113]	
$\mathcal{B}\ (B^+ \to K^+\eta'\ )$	$(70.6 \pm 2.5) \times 10^{-6}$	
$\mathcal{B}\ (\eta' \to \pi^+\pi^-\gamma\ )$	$(29.1 \pm 0.5) \times 10^{-2}$	
ε_{presel}	$(0.2526 \pm 0.0021)\%$	$(0.2626 \pm 0.0030)\%$
$\varepsilon^{\eta' window}$	$(73.69 \pm 0.35)\%$	$(74.17 \pm 0.49)\%$
Expected Events	18990 ± 2436	9359 ± 1795
Observed Events	14936 ± 339	5637 ± 169
Ratio Expected/Observed	1.27 ± 0.17	1.66 ± 0.32

background-subtracted data are used as calibration samples to determine ε_{PID}; a full description of the PID efficiency calculation procedure is given in Section 4.3.3.

Table 4.5 shows the values of the individual components used in the calculation of the expected yields, the resulting overall expected yields and the observed yields. In 2012(2011) data there is a $1.6\sigma(2.1\sigma)$ excess of expected events.

In theory any of the terms in Equation (4.9) could be the cause of this discrepancy, however only discrepancies that will not cancel in the efficiency ratio (with the rare channels) are problematic. The luminosity and cross section will fully cancel in the efficiency ratio and the branching fractions are well measured so these are not

ate resonances to match the decay generated; the true ID of each particle's mother to be consistent with the decay generated; and all the daughters of a given mother particle are required to originate from the same mother candidate.

Table 4.6: Breakdown of L0 TIS and TOS efficiencies for the control and signal channels.

Efficiency	$B^+ \to K^+\eta'$	$\eta' \to \pi^+\pi^-\gamma$ Signal	$\eta' \to \pi^+\pi^-\eta$ Signal
2012 Conditions:			
$\varepsilon_{L0Hadron\_TOS}$	$(68.24 \pm 0.32)\%$	$(61.00 \pm 0.23)\%$	$(55.17 \pm 0.30)\%$
$\varepsilon_{L0Global\_TIS}$	$(55.86 \pm 0.34)\%$	$(62.76 \pm 0.23)\%$	$(66.76 \pm 0.29)\%$
$\varepsilon_{L0Hadron\_TOS(only)}$	$(44.14 \pm 0.34)\%$	$(37.24 \pm 0.23)\%$	$(33.24 \pm 0.29)\%$
$\varepsilon_{L0Global\_TIS(only)}$	$(31.76 \pm 0.32)\%$	$(38.99 \pm 0.23)\%$	$(44.83 \pm 0.30)\%$
$\varepsilon_{L0Hadron\_TOS\&\&L0Global\_TIS}$	$(24.10 \pm 0.29)\%$	$(23.77 \pm 0.20)\%$	$(21.92 \pm 0.25)\%$
2011 Conditions:			
$\varepsilon_{L0Hadron\_TOS}$	$(71.08 \pm 0.42)\%$	$(61.97 \pm 0.30)\%$	$(56.45 \pm 0.39)\%$
$\varepsilon_{L0Global\_TIS}$	$(55.56 \pm 0.47)\%$	$(63.61 \pm 0.30)\%$	$(66.71 \pm 0.37)\%$
$\varepsilon_{L0Hadron\_TOS(only)}$	$(44.43 \pm 0.47)\%$	$(36.39 \pm 0.30)\%$	$(33.28 \pm 0.37)\%$
$\varepsilon_{L0Global\_TIS(only)}$	$(28.92 \pm 0.42)\%$	$(38.03 \pm 0.30)\%$	$(43.55 \pm 0.39)\%$
$\varepsilon_{L0Hadron\_TOS\&\&L0Global\_TIS}$	$(26.65 \pm 0.41)\%$	$(25.58 \pm 0.27)\%$	$(23.17 \pm 0.33)\%$

expected to contribute significantly to this small excess. The fragmentation fraction f_u is a world average[7] of measurements from LEP, CDF and LHCb [113]; this is not expected to be the cause of this discrepancy. This just leaves the pre-selection efficiencies $\varepsilon_{\text{presel}}$.

One possible source of the discrepancy between observed and expected events seen in Table 4.5 is the L0 trigger efficiencies. It is known that these are not always well reproduced in MC. Nominally, as described in section 4.2.1 events are required to pass either the L0Global_TIS or L0Hadron_TOS requirements in order to pass the L0 trigger selection. To understand whether any data/MC discrepancies are likely to fully cancel in the ratio of control to rare channel efficiencies, the L0 TIS/TOS trigger requirements are studied. Table 4.6 shows the efficiencies of the individual L0 TIS/TOS trigger requirements and the efficiency of requiring events to pass both the TIS and TOS requirement for each channel and year of data taking. These are all measured relative to $\varepsilon_{offline}$, and are assessed using MC. These values show that the efficiencies of the trigger decisions differ between the channels, therefore a full cancellation of data/MC discrepancies is not guaranteed.

[7]It should be noted that the value used for this cross check is not used in the final branching fraction measurement because the Λ_b^0 fragmentation fraction has a pseudorapidity dependence, therefore the LHCb measurement of $\frac{f_{\Lambda_b^0}}{f_d}$ is used [109]

Table 4.7: Comparison of the fraction of events passing the TIS/TOS L0 trigger cut between MC and *sWeighted* data. All numbers are $N_{\text{trigger}}/N_{\text{presel}}$ where N_{presel} is the number of events passing the pre-selection and N_{trigger} is the number passing the pre-selection and the given trigger requirement.

Efficiency	2012		2011	
	Data	MC	Data	MC
$\varepsilon_{L0Hadron\_TOS}$	0.655 ± 0.007	0.698 ± 0.004	0.667 ± 0.10	0.735 ± 0.005
$\varepsilon_{L0Global\_TIS}$	0.556 ± 0.007	0.550 ± 0.004	0.538 ± 0.011	0.537 ± 0.006
$\varepsilon_{L0Hadron\_TOS(only)}$	0.444 ± 0.007	0.450 ± 0.003	0.462 ± 0.011	0.463 ± 0.006
$\varepsilon_{L0Global\_TIS(only)}$	0.345 ± 0.007	0.302 ± 0.004	0.333 ± 0.010	0.265 ± 0.005
$\varepsilon_{L0Hadron\_TOS\&\&L0Global\_TIS}$	0.211 ± 0.006	0.249 ± 0.003	0.205 ± 0.009	0.272 ± 0.005

To check whether the efficiencies of the individual TIS/TOS trigger decisions are compatible between data and MC, the *sWeights* extracted from the fits shown in Figure 4.1 are used. These fits are performed with the full pre-selection applied, therefore the values in Table 4.7 are the efficiencies of the trigger decisions relative to $\varepsilon_{\text{presel}}$ (rather than $\varepsilon_{offline}$). As the pre-selection includes the nominal trigger selection, the events which pass any of the studied trigger decisions are a subset of those passing the nominal trigger decision (TIS||TOS). In order to make a comparison to MC, the full pre-selection is applied to MC and the same efficiencies are assessed in MC. Table 4.7 shows a comparison of these efficiencies between data and MC. These numbers show that the largest discrepancy between data and MC is the efficiency of the requirement L0Hadron_TOS&&L0Global_TIS. This discrepancy will have minimal impact on the nominal trigger efficiencies because it is the overlap between the TIS and TOS trigger decisions, and the nominal decision used is TIS||TOS. However, a systematic uncertainty is assigned to account for any residual data/MC discrepancy in the efficiency ratios. This is assessed using data calibration samples; a full description is given in Section 5.6.1.1.

Another possible source of non-cancellation in ϵ_{presel} is the HLT 2 trigger efficiencies; the extra track in the signal channels could mean a significant fraction of events are only triggered by the HLT2Topo4BodyBBDTDecision line which would not be the case in the control channel. Therefore, any discrepancy in the efficiency of this

Table 4.8: Breakdown of relative HLT2 efficiencies for the control and signal channels.

Efficiency	$B^+ \to K^+\eta'$	$\eta' \to \pi^+\pi^-\gamma$ Signal	$\eta' \to \pi^+\pi^-\eta$ Signal
2012 Conditions:			
$\varepsilon_{HLT2Topo2Body}$	$(96.18 \pm 0.15)\%$	$(87.91 \pm 0.18)\%$	$(85.46 \pm 0.28)\%$
$\varepsilon_{HLT2Topo3Body}$	$(80.97 \pm 0.31)\%$	$(94.44 \pm 0.13)\%$	$(87.80 \pm 0.26)\%$
$\varepsilon_{HLT2Topo4Body}$	$(0.031 \pm 0.014)\%$	$(63.9 \pm 0.27)\%$	$(50.75 \pm 0.39)\%$
$\varepsilon_{HLT2Topo2Body\|\|HLT2Topo3Body}$	100.0%	$(99.04 \pm 0.05)\%$	$(98.86 \pm 0.08)\%$
$\varepsilon_{HLT2Topo4Body(only)}$	0.0%	$(0.96 \pm 0.05)\%$	$(1.14 \pm 0.08)\%$
2011 Conditions:			
$\varepsilon_{HLT2Topo2Body}$	$(96.28 \pm 0.21)\%$	$(88.56 \pm 0.23)\%$	$(86.16 \pm 0.37)\%$
$\varepsilon_{HLT2Topo3Body}$	$(75.05 \pm 0.47)\%$	$(90.38 \pm 0.22)\%$	$(83.67 \pm 0.40)\%$
$\varepsilon_{HLT2Topo4Body}$	$(0.76 \pm 0.09)\%$	$(55.41 \pm 0.36)\%$	$(45.24 \pm 0.53)\%$
$\varepsilon_{HLT2Topo2Body\|\|HLT2Topo3Body}$	100.0%	$(99.26 \pm 0.06)\%$	$(98.97 \pm 0.11)\%$
$\varepsilon_{HLT2Topo4Body(only)}$	0.0%	$(0.74 \pm 0.06)\%$	$(1.03 \pm 0.11)\%$

trigger line would not cancel in the efficiency ratio. Table 4.8 shows the relative HLT 2 efficiencies for each individual trigger line for all channels. This shows that the fraction of events passing the HLT 2 cut in the signal channels that only pass the `HLT2Topo4BodyBBDTDecision` line is $< 1.2\%$ for all channels. Therefore, any discrepancy in the `HLT2Topo4BodyBBDTDecision` efficiency will have a negligible effect on the overall efficiency ratio.

The PID efficiencies are accurate as they are determined using calibration samples from data. Furthermore, the variables involved in other pre-selection cuts, as shown in Figures 4.2, are well modelled. Therefore, the selection and PID efficiencies are not considered to be the cause of the observed discrepancy.

Unfortunately the exact cause of this discrepancy is not known despite extensive cross checks. Similar levels of discrepancy have been seen by two previous LHCb analyses (Ref. [107] and Ref. [108]) using the $B^+ \to K^+\eta'$ decay channel. In the former, the assumption that this discrepancy will cancel in the ratio of branching fractions was tested by performing the same comparison for $B^+ \to \phi K^+$ decays. The same level of discrepancy was seen. The assumption that this discrepancy cancels in the ratio of efficiencies was validated by making a measurement of the ratio of

branching fractions $\mathcal{B}(B^+ \to K^+\eta')/\mathcal{B}(B^+ \to \phi K^+)$ and multiplying by the world average value for $\mathcal{B}$ $(B^+ \to \phi K^+)$. A value compatible with the world average value for $\mathcal{B}$ $(B^+ \to K^+\eta'$) was recovered, thus showing that the ratio of efficiencies is reliable. Therefore, adopting the same approach as in the previous publications that see this discrepancy in the control channel, it is assumed that these factors are the same for the control channel and rare channels, meaning the efficiency ratios are reliable. Therefore, no additional systematic uncertainty is assigned specifically for this discrepancy.

4.2.3 Multivariate Selection

After the pre-selection, large levels of combinatorial background remain in all channels. In order to make the best possible use of the information available and maximise the sensitivity of the selection, a Boosted Decision Tree (BDT) based multivariate selection is developed to reject combinatorial background. The BDTs trained make use of the AdaBoost algorithm within the TMVA package [114, 115]. The use of a gradient boosted decision tree and a multi layer perceptron neural network were also considered, but the AdaBoost algorithm gave slightly better performance and is less susceptible to over training[8]. Separate BDTs are trained for each year of data taking and decay channel, with MC used for the signal training sample and upper mass sideband data used for the background training sample. In the case of the rare channels the upper sideband (USB) is defined as $5819\,\text{MeV} < m_{\Lambda_b^0} < 6200\,\text{MeV}$ and in the control channel it is defined as $5500\,\text{MeV} < m_{B^+} < 5900\,\text{MeV}$. The lower sideband (LSB) ($\text{M}(pK^-\,\eta') < m_{\Lambda_b^0}$) is not used because there could be partially reconstructed backgrounds present and the primary purpose of this BDT is to separate signal and combinatorial background.

Each training sample is split into two separate samples, `Sample A` and `Sample B`

[8] Over training occurs when the BDT learns statistical fluctuations in the training sample, causing it to have decreased performance when applied to a separate data sample

(based on odd or even `eventNumber`). For each year of data taking and each channel two BDTs are used; one is trained on `Sample A` and applied to `Sample B` whilst the opposite process is used for the second BDT. This is done to avoid introducing bias arising from classifying events with a BDT that was trained with the same set of events as those that are being classified.

Many variables are considered for inclusion in the BDT, but only those that provide an appreciable level of discrimination between signal and background are retained; variables that have a TMVA variable importance <0.01 are removed. The Receiver Operator Curve (ROC) integral of the trained BDT is then studied for various sets of the remaining variables until an optimal set is found. The ROC curve is a plot of signal efficiency vs. background rejection; a perfect ROC curve would show 100% signal efficiency for 100% background rejection, which would have an integral of 1. The variables chosen are summarised in Table 4.9 and the distributions for signal and background training samples are shown in Figure 4.3 and Figure 4.4. A description of the variables used is as follows:

- $\frac{P\sin(\theta_{\rm DIRA})}{P\sin(\theta_{\rm DIRA})+\sum p_{\rm T}}$, "Pointing Angle" : P is the momentum of the Λ_b^0; $\sin(\theta_{\rm DIRA})$ is the sine of the DIRA angle which is defined in Section 4.2.2; and $\sum p_{\rm T}$ is a sum of the transverse momenta of all stable Λ_b^0 daughter particles. Explicitly, these are the p,K^-,π^+, π^- and photon(s).

- $\ln(\chi^2_{\rm vtx})$: The natural logarithm of the χ^2 of the fit performed to the Λ_b^0 decay vertex.

- $\ln(\tau\chi^2)$: The natural logarithm of the χ^2 of the fit used to extract the Λ_b^0 lifetime.

- $\ln(1-\cos(\theta_{\rm DIRA}))$: The natural logarithm of one minus the cosine of $\theta_{\rm DIRA}$.

- $\ln(p_{\rm T})$: The natural logarithm of the transverse momentum of the candidate particle, in units of MeV. This is used for the Λ_b^0, proton, kaon and photon(s).

Table 4.9: Variables used in the BDTs trained to discriminate between signal and combinatorial background.

	Variables used
Λ_b^0	$\frac{P\sin(\theta_{\rm DIRA})}{P\sin(\theta_{\rm DIRA})+\sum p_{\rm T}}$
	$\ln(\chi^2_{\rm vtx})$
	$\ln(\tau\chi^2)$
	$\ln(1-\cos(\rm DIRA))$
	$\ln(p_{\rm T})$
	η
	$\ln(DTF\chi^2)$
p, K^-	$\ln(p_{\rm T})$
γ	$\ln(p_{\rm T})$
	CL

- $\ln(\Lambda_b^0\eta)$: The natural logarithm of the pseudorapidity of the Λ_b^0 particle.
- γ CL: As defined in Section 4.2.2.
- $\ln(DTF\chi^2)$: The natural logarithm of the χ^2 from the fit to the entire decay chain with Decay Tree Fitter.

The presence of two photons in the $\eta' \to \pi^+\pi^-\eta$ channel has been dealt with by including the photon specific variables for both photons, but ordering the photons by $p_{\rm T}$ (`gammaA` refers to the photon with the higher $p_{\rm T}$). In order to keep the selection as similar as possible between the rare and control channels, the BDTs trained and applied to the control channel use the same variables as those used in the rare channels (without the proton $p_{\rm T}$).

Checks have been carried out to ensure the variables used in the BDT are reasonably well modelled in MC. Comparisons between 2012 control channel MC and background subtracted data for the variables entering the BDT are shown in Figure 4.5, and those also used in the pre-selection are shown in Figure 4.2. The background subtraction is performed using the mass fit described in Section 4.2.2.1. It is clear that $\ln(DTF\chi^2)$ is not as well modelled as one would hope. However, removing this variable significantly reduces the performance of the BDT. When this variable

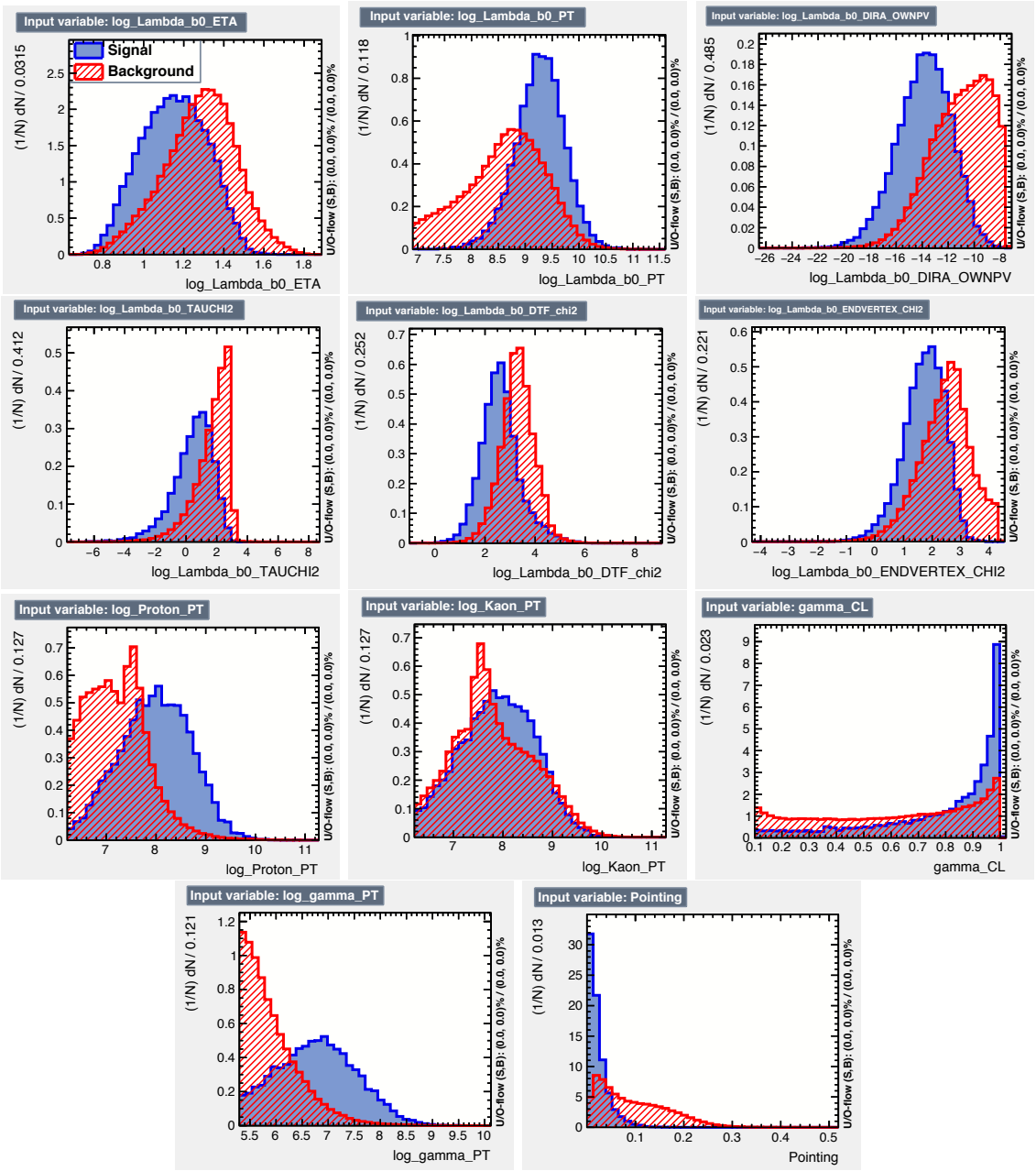


Figure 4.3: Comparisons between signal and background BDT training samples, in the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel. From top left to bottom right: Λ_b^0 η; Λ_b^0 p_{T}; Λ_b^0 DIRA; Λ_b^0 τ χ^2; DTF χ^2; Λ_b^0 χ^2_{vtx}; p p_{T}; K^- p_{T}; γ CL; γ p_{T} and "PointingAngle". Descriptions of these variables are given in the text.

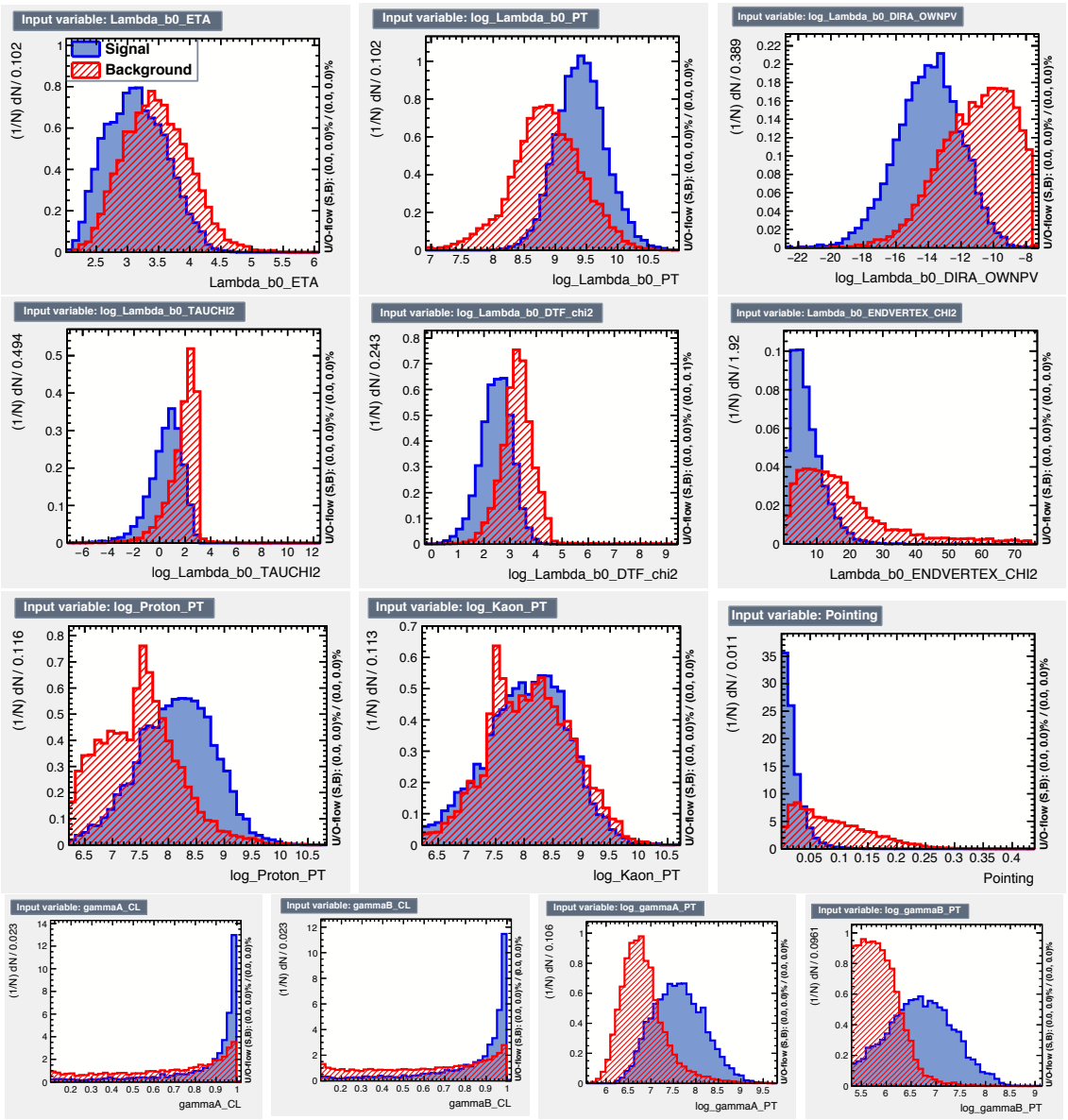


Figure 4.4: Comparisons between signal and background BDT training samples, in the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) channel. From top left to bottom right: Λ_b^0 η; Λ_b^0 p_{T}; Λ_b^0 DIRA; Λ_b^0 τ χ^2; DTF χ^2; Λ_b^0 χ^2_{vtx}; p p_{T}; K^- p_{T}; "Pointing Angle"; γ A CL; γ B CL; γ A p_{T}; γ B p_{T}. Descriptions of these variables are given in the text.

is removed and the optimisation procedure is repeated, the number of events in the signal window for the $\eta' \to \pi^+\pi^-\gamma$ channel approximately doubles whilst the signal efficiency stays approximately the same. Therefore this variable is still used; a systematic uncertainty on the BDT efficiency ratio will be assigned (as described in Section 5.6.1.4) to account for any non-cancellation of this discrepancy with the control channel.

Figure 4.6 shows the BDT classifier distributions for the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$), $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) and control channels respectively. These show no signs of over training and a good level of separation between signal and background events.

4.2.3.1 Optimisation of BDT Cuts

In the rare channels the BDT cut is optimised for the so-called "Punzi" figure of merit (FoM) [116], which is given by

$$FoM = \frac{\varepsilon_{BDT}}{\frac{a}{2} + \sqrt{B}}, \tag{4.11}$$

where ε_{BDT} is the signal efficiency of a given BDT cut, B is the estimated number of background events within 3σ of the Λ_b^0 mass after applying a given BDT cut and a is the desired signal significance in units of Gaussian standard deviations (σ). $a = 5$ is used for both rare channels, but the position of the optimal cuts was found to have no dependence on a for reasonable values of a. As shown in Equation 4.11, the Punzi FoM has no dependence on the number of signal events in the sample which makes it ideal for the blind rare channels. The signal efficiency (ε_{BDT}) is determined by applying the BDT to signal MC samples that pass the pre-selection and the number of background events is estimated by fitting the data sidebands and interpolating into the signal region; a second order Chebychev polynomial is used in the $\eta' \to \pi^+\pi^-\gamma$ channel and an exponential is used in the $\eta' \to \pi^+\pi^-\eta$ channel. This is applied nominally in both the upper and lower mass sidebands but has

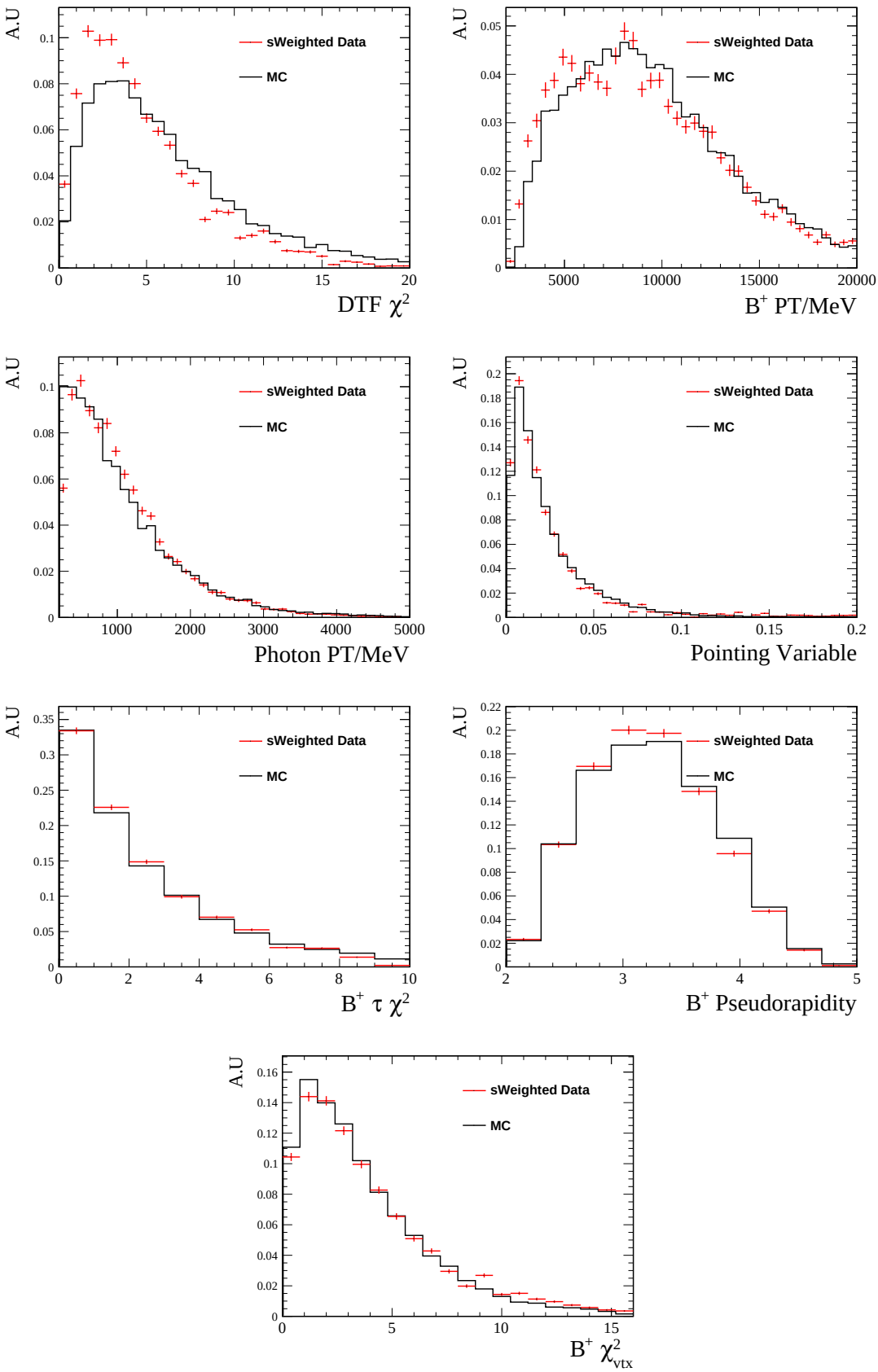


Figure 4.5: Further comparisons of 2012 MC and background subtracted data in the control channel for variables used in the BDT. The same comparisons for 2011 data can be found in Appendix A

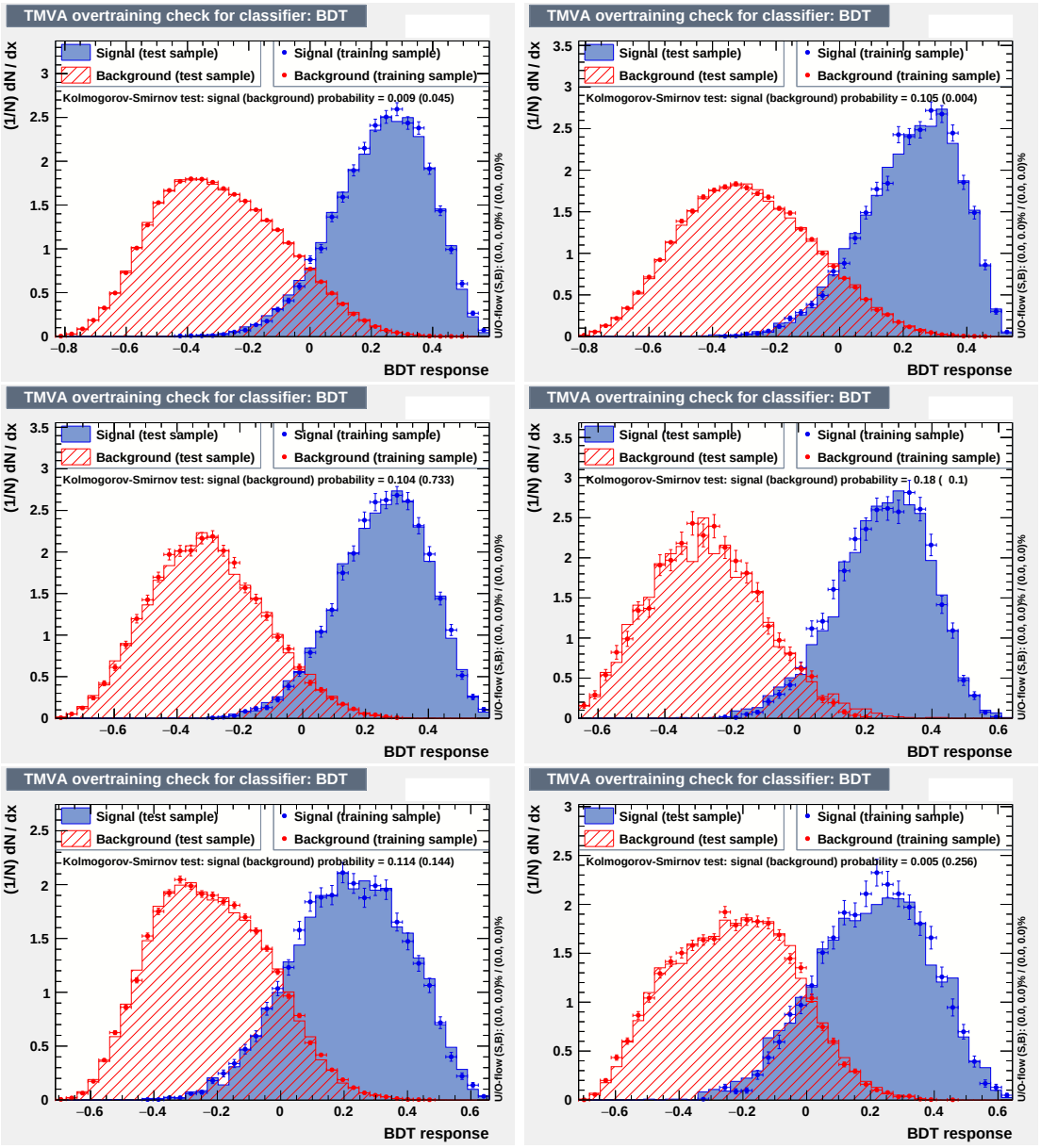


Figure 4.6: BDT classifier distributions for the signal and background training and test samples in the $\eta' \to \pi^+\pi^-\gamma$ (top), $\eta' \to \pi^+\pi^-\eta$ (middle) and control channels (bottom) for 2012(left) and 2011(right) data.

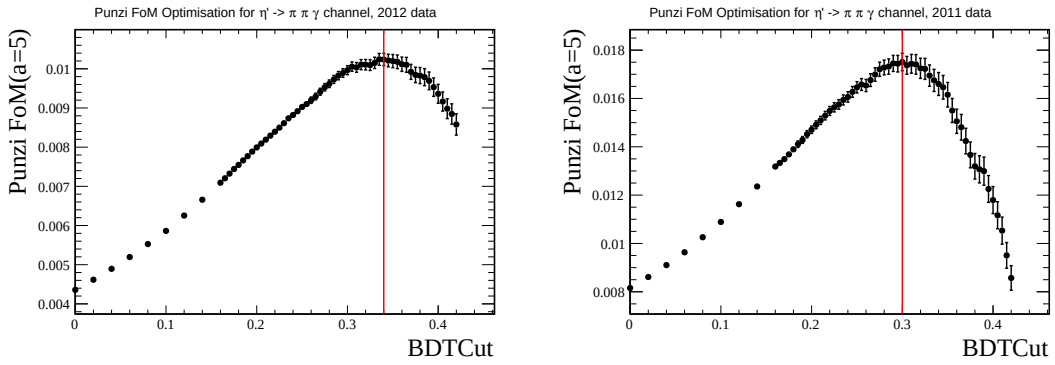


Figure 4.7: Plots of Punzi FoM as a function of BDT cut in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma$) channel for 2012(left) and 2011(right) data. The red vertical line indicates the positon of the chosen cut. The calculation of Punzi FoM is discussed in the text.

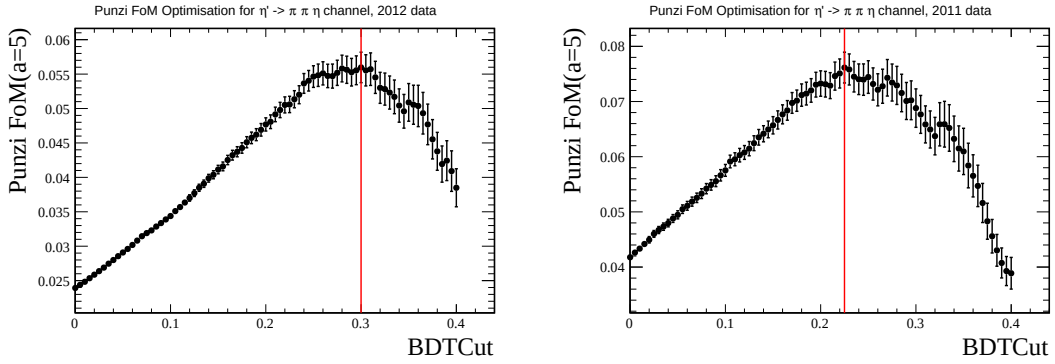


Figure 4.8: Plots of Punzi FoM as a function of BDT cut in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta$) channel for 2012(left) and 2011(right) data. The red vertical line indicates the position of the chosen cut. The calculation of Punzi FoM is discussed in the text.

also been performed for each sideband individually and the results are consistent. Figures 4.7 and 4.8 show plots of Punzi FoM against BDT cut for the rare channels with the position of the chosen cut indicated.

The control channel BDT cut is optimised for signal significance because it is a well studied channel and optimisation for signal significance retains a larger signal yield than an optimisation for Punzi FoM; this is important to minimise systematic uncertainties arising from a limited control channel yield. Signal significance is estimated by

$$\Sigma = \frac{S}{\sqrt{S+B}} = \frac{\varepsilon_{BDT} S_0}{\sqrt{\varepsilon_{BDT} S_0 + B}}, \tag{4.12}$$

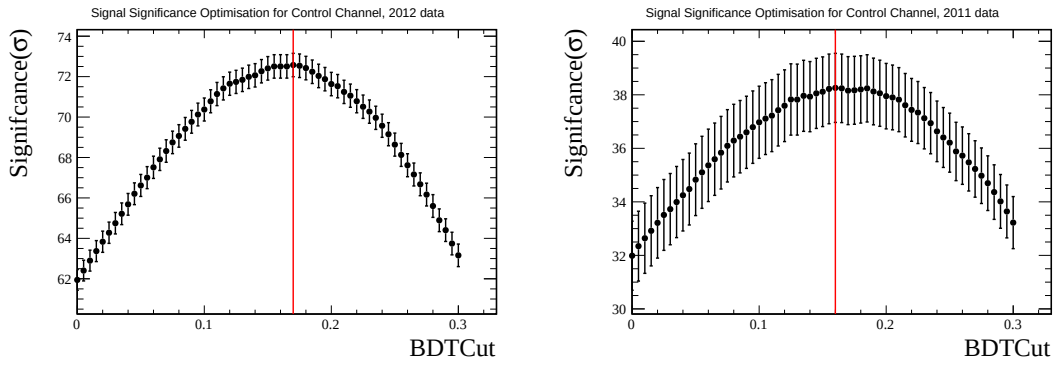


Figure 4.9: Plots of Signal Significance as a function of BDT cut in the $B^+ \to K^+\eta'$ control channel for 2012(left) and 2011(right) data. The red vertical line indicates the position of the chosen cut. The estimation of significance is discussed in the text.

Table 4.10: A Summary of BDT cuts chosen for each channel and year of data taking.

Channel	2012	2011
$\Lambda_b^0 \to pK\eta' \quad (\eta' \to \pi^+\pi^-\gamma\ )$	0.34	0.30
$\Lambda_b^0 \to pK\eta' \quad (\eta' \to \pi^+\pi^-\eta\ )$	0.30	0.225
$B^+ \to K^+\eta' \quad (\eta' \to \pi^+\pi^-\gamma\ )$	0.17	0.16

where $S(B)$ is the number of signal(background) events within 3σ of the B^+ mass, ε_{BDT} is the signal efficiency of the BDT and S_0 is the number of signal events in the sample with no BDT cut applied. ε_{BDT} is determined by applying the BDT to signal MC, S_0 is taken from the fits shown in Figure 4.1 and B is estimated by fitting the sidebands with a first order Chebychev polynomial and interpolating into the signal region. S is determined using $\varepsilon_{BDT} \times S_0$ rather than fitting the signal peak at each BDT cut as this could introduce bias from forcing signal to be in that region. As discussed in Section 4.2.2.1, the efficiencies determined from MC are overestimated, therefore a Data/MC correction factor is applied to ε_{BDT} for the significance optimisation procedure. The Data/MC correction factors are taken as the ratio of Observed/Expected events in the $B^+ \to K^+\eta'$ channel after pre-selection; these values are shown in Table 4.5. Figure 4.9 shows plots of signal significance as a function of BDT cut for each year of data taking with the position of the chosen cut indicated. All of the chosen cuts are summarised in Table 4.10.

4.2.4 PID Selection

Following the application of the BDT, the PID ability of the LHCb detector is used to further reduce backgrounds and ensure backgrounds involving mis-identified hadrons are suppressed. The `ProbNN` set of variables is used, specifically the variables `Kaon_ProbNNk`, `Proton_ProbNNp` and `Pion_ProbNNpi`. These are the probabilities that each particle species has been correctly identified[9].

Various sources of mis-ID backgrounds have been considered. In the $\eta' \to \pi^+\pi^-\gamma$ channel, backgrounds from decays such as $B^0/B_s^0 \to hK^-\pi^+\pi^-$, where h is a Kaon or Pion mis-identified as a proton, are considered. This decay could then be combined with a random photon to create a fake $\Lambda_b^0 \to pK\eta'$ candidate. The addition of the random photon causes these decays to peak higher than the B^0 mass, meaning they could potentially be in the Λ_b^0 signal window. To investigate possible contamination from these backgrounds the invariant mass of the four hadrons in the final state, $M(pK^-\pi^+\pi^-)$, is reconstructed with the mass hypothesis of the p changed for a K/π. No peaking structure is apparent, suggesting negligible contamination from these backgrounds. Furthermore, the pre-selection and multivariate selections have been applied to MC samples of $B^0 \to K^{*0}\rho^0$ ($K^{*0} \to K^+\pi^-$, $\rho^0 \to \pi^+\pi^-$), $B^0 \to K^+\pi^-\pi^+\pi^-$ and $B^0 \to \phi K^{*0}$ ($\phi \to K^+K^-$, $K^{*0} \to K^-\pi^+$) events; fewer than 10 events remain in the MC samples in all cases, which leads to a predicted yield at this stage of the selection (normalised to the control channel, see Equation (4.25)) of <0.1 events in all of the mentioned channels. In the $\eta' \to \pi^+\pi^-\eta$ channel, backgrounds from $B^0 \to hK^-\pi^+\pi^-$ combined with a random photon are significantly less likely because there are two photons that are constrained to have an invariant mass within $\pm 50\,\text{MeV}$ of the η mass ($548\,\text{MeV}$). As with the $\eta' \to \pi^+\pi^-\gamma$ channel, the invariant mass of the four hadrons in the final state is reconstructed with the mass hypothesis of the p changed for that of a K/π. No peaking structure is observed which again strongly suggests no background contamination from $B^0/B_s^0 \to hK^-\pi^+\pi^-$ decays.

[9] See Section 2.2.4 for a description of how these variables are determined.

The decay $B^0 \to \eta' K^{*0}$ is also considered, due to the fact a pion from the K^{*0} could be mis-indentified as a proton to create a fake Λ_b^0 candidate. Again the pre-selection and multivariate selection have been applied to a sample of $B^0 \to \eta' K^{*0}$ MC events in the $\eta' \to \pi^+\pi^-\gamma$ channel. It is consequently predicted that there are <10 events surviving the previous stages of the selection in the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) data sample. As this is an almost negligible number of events and even a loose PID cut on the proton of `Proton_ProbNNp`>0.2 reduces this number to ≈ 1, a dedicated PID optimisation is not pursued for this background.

The decay $B_s^0 \to (\phi \to K^+K^-)(\phi \to \pi^+\pi^-\pi^0)$ could appear as both a mis-ID and partially reconstructed background because missing a photon from the π^0 causes the ϕ, which decays to $\pi^+\pi^-\pi^0$, to peak very close to the η' mass. As with the previously discussed decays, the pre-selection and multivariate selection are applied to a sample of fully simulated MC and it is predicted that ≈ 12 events would remain in the data samples. As with the $B^0 \to \eta' K^{*0}$ decay, even loose PID cuts remove this background completely therefore a dedicated PID optimisation is not pursued for this background.

In theory there could also be mis-ID background from $\Lambda_b^0 \to p\pi^-\eta'$ decays, but there are no theory predictions for the BF of this decay which makes estimating its contribution difficult. However this decay is expected to be heavily suppressed by the presence of a $b \to u$ transition; approximately by a factor $\frac{|V_{td}|^2}{|V_{ts}|^2} = 0.046 \pm 0.004$, with which the ratio $\mathcal{B}(B^+ \to \pi^+\eta')/\mathcal{B}(B^+ \to K^+\eta')$=0.038 $\pm$ 0.012 [1] is consistent. Consequently, if a very high BF of 5×10^{-5} is assumed for the rare channel $\Lambda_b^0 \to pK\eta'$, and a suppression factor of 0.042 is assumed for the $\Lambda_b^0 \to p\pi^-\eta'$ channel, it is predicted that approximately only 6 $\Lambda_b^0 \to p\pi^-\eta'$ events would be present in the $\eta' \to \pi^+\pi^-\gamma$ channel data sample after the BDT selection. Therefore, a dedicated PID optimisation for this background is also not pursued. The efficiencies for this decay mode are calculated by applying the pre-selection and BDT selection to a sample of $\Lambda_b^0 \to p\pi^-\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) MC events.

As no specific mis-ID backgrounds with an appreciable contribution peaking in or near the signal window have been identified, the PID selection is optimised in the same way as the BDT optimisation. The Punzi FoM is used for the PID optimisation with $a = 5$ and the number of background events in the signal window is estimated by fitting the data sidebands with a 2nd order Chebychev polynomial in the $\eta' \to \pi^+\pi^-\gamma$ channel and an exponential in the $\eta' \to \pi^+\pi^-\eta$ channel.

Unlike the BDT optimisation, the efficiencies for a given set of cuts cannot be evaluated directly with signal MC as the PID variables are not correctly reproduced in MC. Therefore, in a similar manner to that adopted to determine the pre-selection efficiency in Section 4.2.2.1, the PID efficiencies are determined using data calibration samples. A full description of this procedure is given in Section 4.3.3. To ensure that this calibration and efficiency calculation is accurate it is firstly required that all hadrons have interacted with the RICH system. There is also a cut on K^- momentum applied, $p < 450.0\,\text{GeV}$, because the calibration samples provide no coverage at such high momentum. Furthermore, in the case of protons some kinematic cuts also have to be applied before the optimisation to ensure the efficiency calibration is accurate. As there is particularly low coverage in the low η, high p region, the cut on momentum is separated into three regions of η. The cuts applied are:

- Proton η>1.8
- Proton 1.8<η<2.35 : Proton p <66.05 GeV
- Proton 2.35<η<2.8 : Proton p <87.95 GeV
- Proton η>2.8 : Proton p <150 GeV

The PID optimisation is performed using the combined 2011 and 2012 datasets, simultaneously as a function of the ID variables of the proton, kaon and pions in the rare channels. The PID selection chosen for the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel is also applied to the control channel. All possible combinations of PID cuts[10] are

[10]The `MC12TuneV3` variables are used. This is the third tuning of the neural network based PID variables.

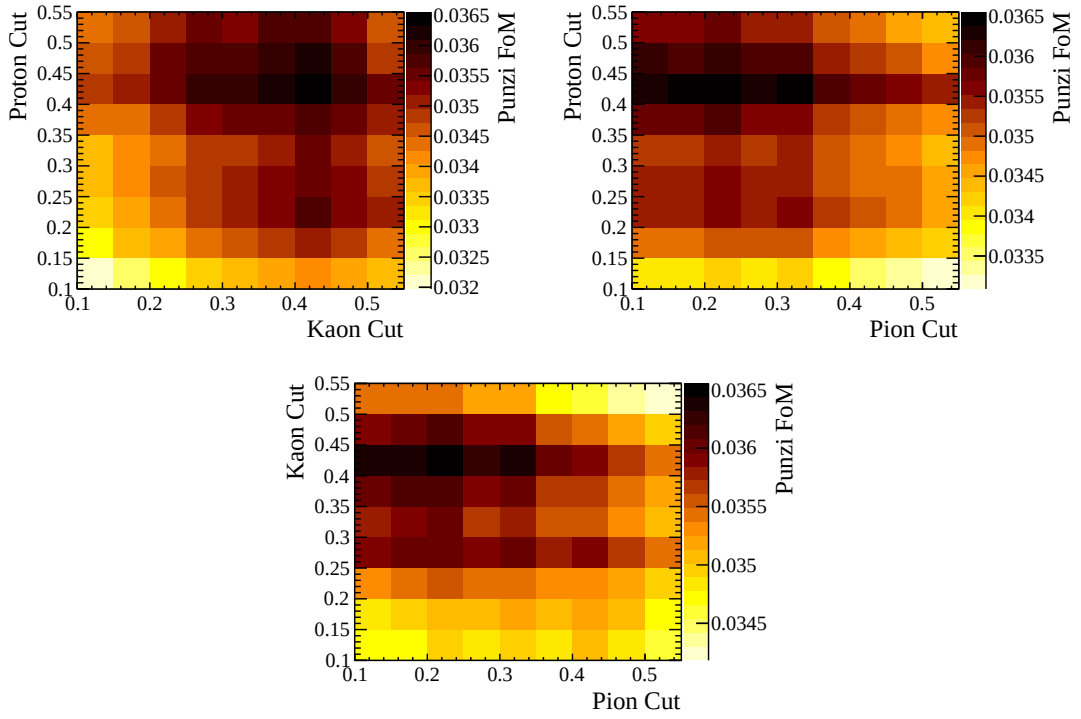


Figure 4.10: 2D Slices of the PID Optimisation in the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel where the PID variable not present in the plot is fixed to its optimum value.

considered in the set {0.1,0.15.....0.5}. In the $\eta' \to \pi^+\pi^-\eta$ channel, not applying any pion ID cut (in addition to the `ProbNN`>0.1 applied in the pre-selection) is also considered. Figures 4.10 and 4.11 show 2D slices of the optimisation in each channel at the optimum value of the third PID variable (the variable not present in the plot). The values of the cuts chosen are shown in Table 4.11. These cuts reduce the specific mis-ID backgrounds discussed previously in this section to negligible levels.

It is possible, even after these cuts are applied, that there are a small number of events in the $\eta' \to \pi^+\pi^-\gamma$ channel where one of the hadrons is a mis-identified muon. This is dealt with by applying further PID requirements on the hadrons which require them to be inconsistent with being a muon[11].

[11]The explicit requirement is `X_isMuon==0`, where X is $\in (p, K^-, \pi^+, \pi^-)$

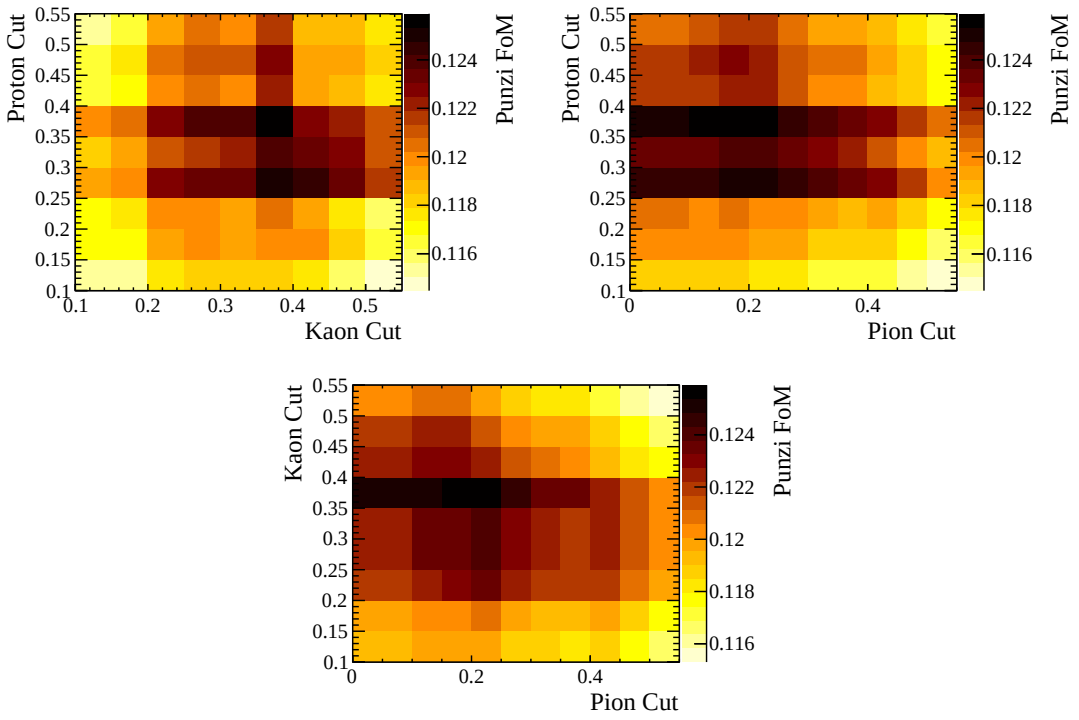


Figure 4.11: 2D Slices of the PID Optimisation in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ channel where the PID variable not present in the plot is fixed to its optimum value. In the case of the pion ID variables, the bin between 0.0 and 0.1 represents no cut in addition to the `ProbNNpi` requirement applied in the stripping.

Table 4.11: A Summary of the PID cuts chosen.

Variable	$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$	$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$
`Proton_ProbNNp`	0.40	0.35	
`Kaon_ProbNNk`	0.40	0.35	0.40
`Pion_ProbNNpi`	0.20	0.15	0.20

Table 4.12: A summary of the mass vetoes applied in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ channel. The phase space integrated efficiencies of these vetoes, relative to the previous stages of the selection, are also shown for both years of data taking.

Resonance	Mass Window	2012 Efficiency	2011 Efficiency
$D^0 \to K^-K^+$	$\left\lvert M(K^-(p\to K^+)) - M(D^0)\right\rvert > 30.0\,\text{MeV}$	$(98.1 \pm 0.2)\%$	$(98.3 \pm 0.2)\%$
$D^0 \to K^-\pi^+$	$\left\lvert M(K^-(p\to \pi^+)) - M(D^0)\right\rvert > 30.0\,\text{MeV}$	$(98.2 \pm 0.2)\%$	$(98.3 \pm 0.2)\%$
$D^0 \to K^-\pi^+$	$\left\lvert M(K^-\pi^+) - M(D^0)\right\rvert > 30.0\,\text{MeV}$	$(97.0 \pm 0.2)\%$	$(97.3 \pm 0.2)\%$
$D^0 \to \pi^-K^+$	$\left\lvert M(\pi^-(p\to K^+)) - M(D^0)\right\rvert > 30.0\,\text{MeV}$	$(97.2 \pm 0.2)\%$	$(97.2 \pm 0.3)\%$
$\Lambda_c^+ \to p\pi^+K^-$	$\lvert M(p\pi^+K^-) - M(\Lambda_c^+)\rvert > 30.0\,\text{MeV}$	$(98.0 \pm 0.2)\%$	$(98.0 \pm 0.2)\%$
$\Lambda_c^+ \to p\pi^+\pi^-$	$\lvert M(p\pi^+\pi^-) - M(\Lambda_c^+)\rvert > 30.0\,\text{MeV}$	$(99.8 \pm 0.1)\%$	$(99.9 \pm 0.1)\%$
$\Lambda_b^0 \to pK^-\pi^+\pi^-$	$\left\lvert M(pK^-\pi^+\pi^-) - M(\Lambda_b^0)\right\rvert > 60.0\,\text{MeV}$	$(99.98 \pm 0.02)\%$	$(99.97 \pm 0.03)\%$
Overall		$(88.7 \pm 0.4)\%$	$(89.5 \pm 0.5)\%$

4.2.5 Mass Vetoes

With the pre-selection, BDT selection and PID selection applied particularly large levels of background still remain in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ channel. Furthermore, inspection of the $M(\Lambda_b^0)$ sidebands suggests it is not just combinatorial background present; there are other sources of background present in the data sample. After further investigation it is found that there are several backgrounds present that involve intermediate (largely charm) resonances; these are removed by applying specific mass vetoes, which are summarised in Table 4.12.

In the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$ channel, the background levels are considerably lower and there is only combinatorial background present. The first reason for this is the presence of two photons in the final state which are constrained to have an invariant mass within 50 MeV of the η mass. Secondly the η having a mass of 548 MeV [1], compared to the massless photon, means the two pions in the final state have considerably softer kinematics compared to the two pions in the $\eta' \to \pi^+\pi^-\gamma$ channel. This means the pion kinematics are significantly less similar to the daughters of other heavy hadron decays, which are the sources of non-combinatorial background in the $\eta' \to \pi^+\pi^-\gamma$ channel for which mass vetoes are applied.

Figure 4.12 shows intermediate resonant peaks from $D^0 \to K^-K^+$ decays where the

K^+ is mis-identified as a proton, $D^0 \to K^-\pi^+$ decays where the π^+ is mis-identified as a proton and $D^0 \to \pi^-K^+$ decays where the K^+ is mis-identified as a proton. In each case the invariant mass of the D^0 is reconstructed with the mass hypothesis of the proton swapped for that of the mis-identified particle. Figure 4.13 shows resonant peaks from $D^0 \to K^-\pi^+$, $\Lambda_c^+ \to p\pi^+K^-$ and $\Lambda_c^+ \to p\pi^+\pi^-$ decays where there is no mis-identification. These resonant backgrounds are vetoed by removing events within $\pm 30\,\text{MeV}$ of the known D^0/ Λ_c^+ mass [1].

As shown in Figure 4.14, there are also background events present involving $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decays. These events can be associated with random photons to create fake $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ candidates, which are present entirely in the upper sideband of the M($pK\ \eta'$) spectrum. These background events are vetoed by removing events where $M(pK\pi^+\pi^-)$ falls within $\pm 60\,\text{MeV}$ of the known Λ_b^0 mass [1]; this width was chosen by studying large, fast, MC samples of $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decays (generated with the `RapidSim` package [94]) that have the kinematic stripping cuts applied.

4.2.5.1 Effect of Mass Vetoes on the Dalitz Distribution

The application of these mass vetoes unavoidably causes events to be removed non-uniformly from the phase space of the $\Lambda_b^0 \to pK\eta'$ decay; they introduce extra dependence on phase space position to the signal efficiency. As discussed in section 4.1, it is expected that the $\Lambda_b^0 \to pK\eta'$ decay will proceed through Λ^* resonances which are not *a priori* known. Therefore, the signal efficiency will be corrected for phase space variations in the event of an observation. If the mass vetoes discussed in this section remove a disproportionately large number of events from the area of the phase space where Λ^* resonances are present, the corrected efficiencies could be significantly lower than the phase space integrated efficiencies shown in Table 4.12. This would be detrimental to the sensitivity of this search. In order to check that

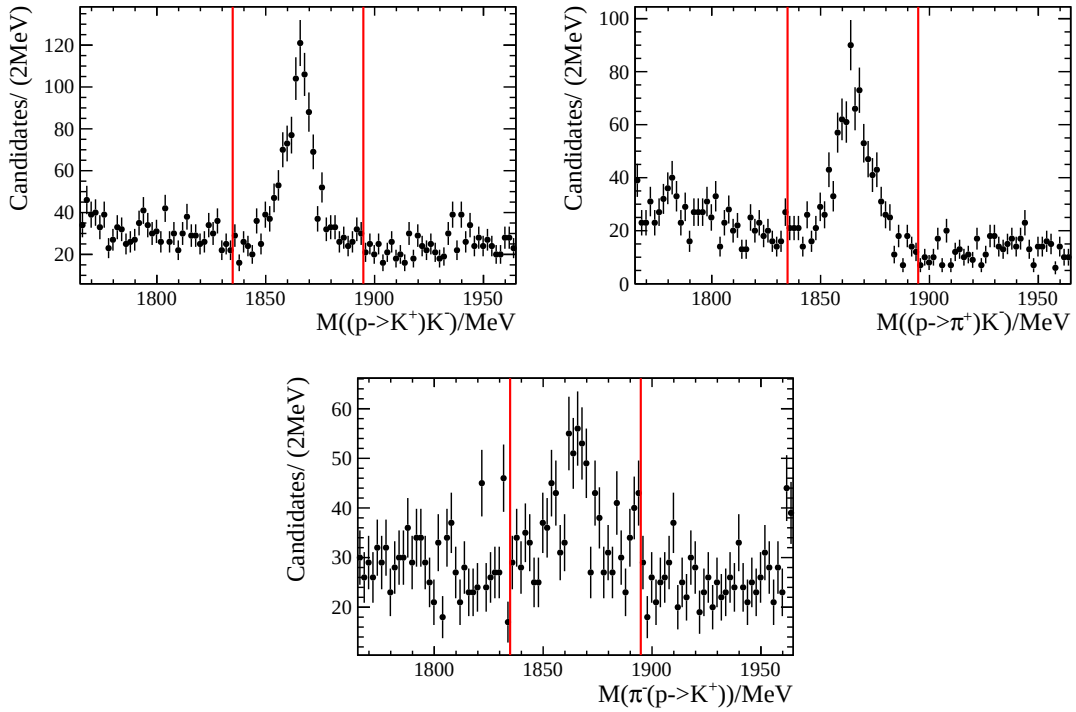


Figure 4.12: Top Left: Invariant mass of the K^- and proton system in the $\eta' \to \pi^+\pi^-\gamma$ channel (with the proton reconstructed under the K^+ mass hypothesis) showing an excess of $D^0 \to K^+K^-$ events. Top Right: Invariant mass of the K^- and proton in the $\eta' \to \pi^+\pi^-\gamma$ channel (with the proton reconstructed under the π^+ mass hypothesis) showing an excess of $D^0 \to K^-\pi^+$ events. Bottom: Invariant mass of the π^- and proton in the $\eta' \to \pi^+\pi^-\gamma$ channel (with the proton reconstructed under the K^+ mass hypothesis) showing a slight excess of $D^0 \to \pi^-K^+$ events. These plots are created with a looser BDT cut but nominal PID selection. The red vertical lines indicate the vetoed mass ranges.

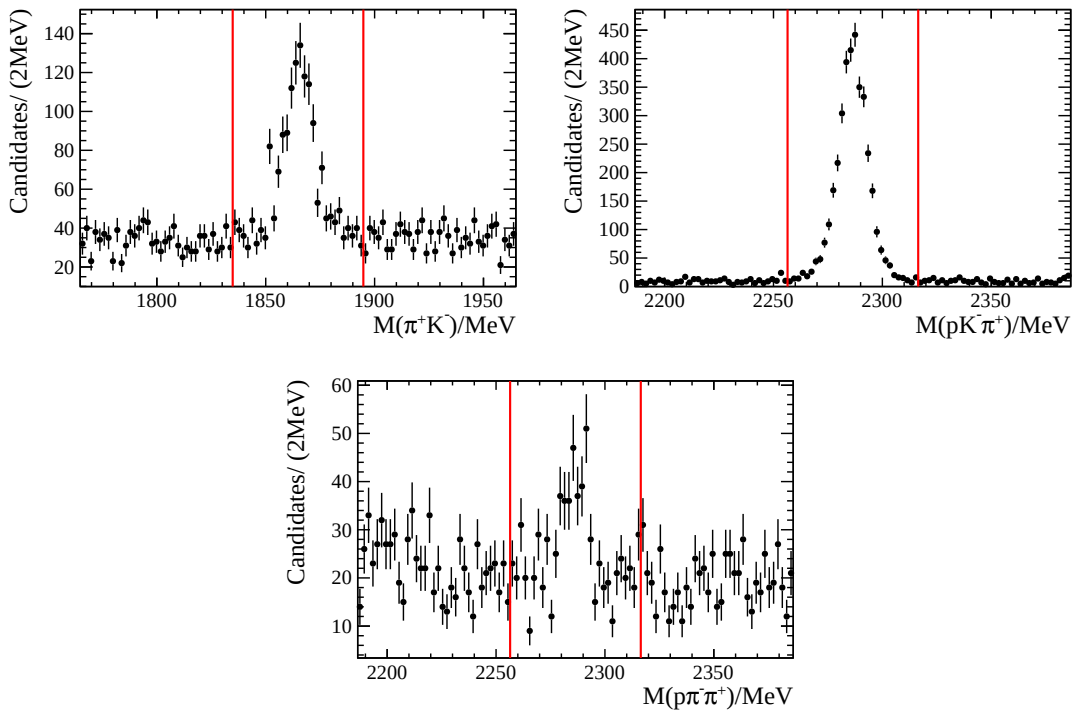


Figure 4.13: Top Left: Invariant mass of the K^- and π^+ system in the $\eta' \to \pi^+\pi^-\gamma$ channel showing an excess of $D^0 \to K^-\pi^+$ events. Top Right: Invariant mass of the p, K^- and π^+ system in the $\eta' \to \pi^+\pi^-\gamma$ channel showing an excess of $\Lambda_c^+ \to pK^-\pi^+$ events. Bottom: Invariant mass of the p, π^+ and π^- system in the $\eta' \to \pi^+\pi^-\gamma$ channel showing an excess of $\Lambda_c^+ \to p\pi^-\pi^+$ events. These plots are created with a looser BDT cut but nominal PID selection. The red vertical lines indicate the vetoed mass ranges.

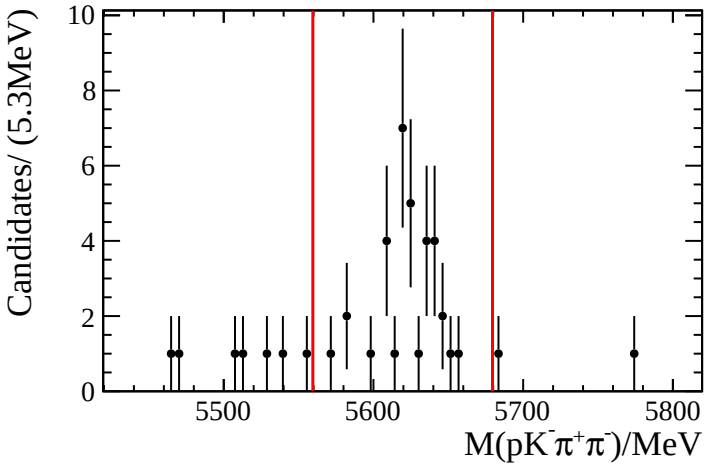


Figure 4.14: The M($pK^-\pi^+\pi^-$) spectrum with the nominal selection applied. The peak is from $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decays that have been associated with a random photon to create a fake $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) candidate. The red vertical lines indicate the veto applied to remove these events.

this is not the case, the Dalitz distribution of signal MC events removed by the mass vetoes is shown in Figure 4.15.

The Dalitz plot shows that the region with $4\,\mathrm{GeV}^2 < M^2_{pK} < 6\,\mathrm{GeV}^2$ is most heavily depleted by these vetoes. The majority of the predicted Λ^* resonances have a mass below 2 GeV [1], and will therefore populate the region with $M^2(pK) < 4\,\mathrm{GeV}$. Therefore, the area of the phase space expected to be most heavily populated by signal decays is not significantly affected by these vetoes; they should not cause the corrected signal efficiency to be significantly lower than the phase space integrated efficiency. A systematic uncertainty will however be assigned to the efficiency ratios to account for the effect of these vetoes, which is described in Section 5.6.

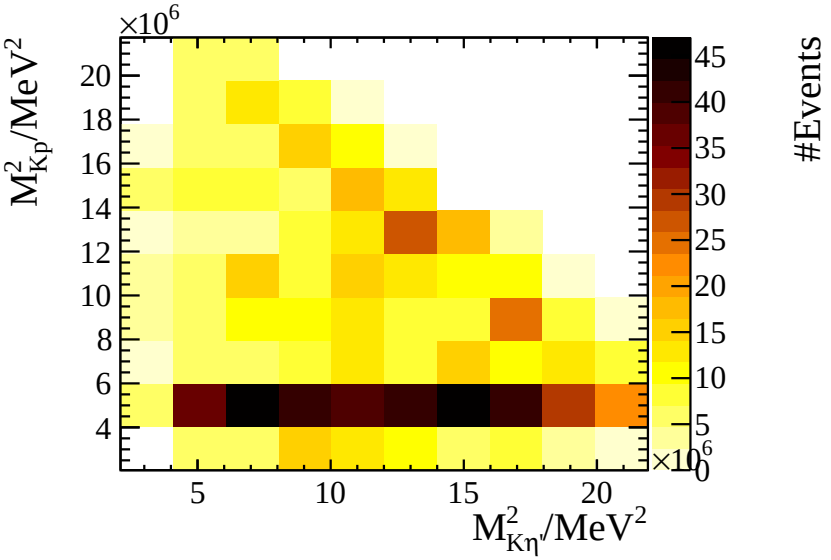


Figure 4.15: Dalitz plot distribution of $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) signal MC events removed by the mass vetoes

4.2.6 Further Requirements

In the $\eta' \to \pi^+\pi^-\gamma$ channel, it is possible for backgrounds involving $\eta' \to \pi^+\pi^-\eta$ decays or light mesons decaying to $\pi^+\pi^-$ such as $K^0_S \to \pi^+\pi^-$ to be present. In the former case, the η is partially reconstructed from one of the several decay channels involving photons in the final state. These are removed by requiring that

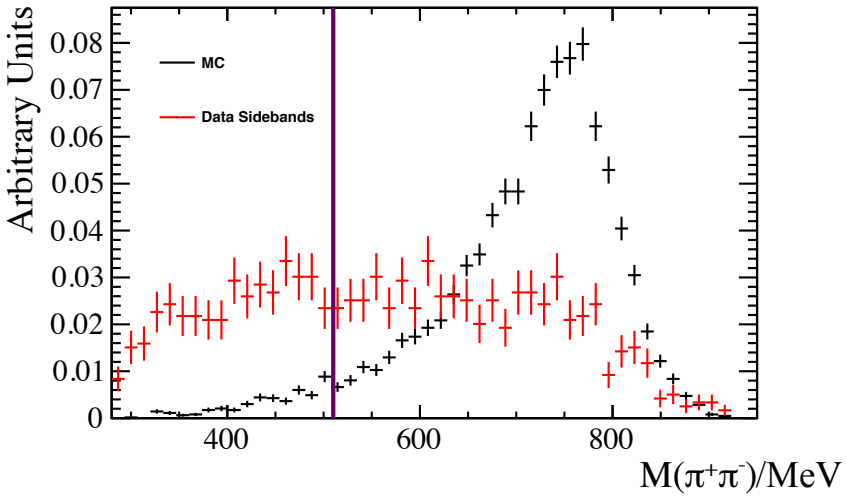


Figure 4.16: The M($\pi^+\pi^-$) normalised mass spectra for Λ_b^0 sideband data and MC in the $\eta' \to \pi^+\pi^-\gamma$ channel. The position of the cut described in the text is indicated by the magenta vertical line.

$M(\pi^+\pi^-)$>510 MeV; a comparison of the $M(\pi^+\pi^-)$ distribution between MC and *sWeighted* control channel data can be found in Figure 4.2. As shown in Figure 4.16, this requirement has a high signal efficiency ($\approx 98\%$) whilst removing a significant amount of background. The same requirement is applied to the control channel for consistency.

As Section 4.1.2 describes, DTF is used to refit the entire decay chain with the η' mass fixed but the η mass, $M_{DTF}(\eta)$, unconstrained in the $\eta' \to \pi^+\pi^-\eta$ channel. The η' mass constraint can result in a value of $M_{DTF}(\eta)$ a long way from the known value for background events that do not involve a true $\eta' \to \pi^+\pi^-\eta$ decay; this is exploited to improve the selection by applying a mass window of 480.0 MeV<$M_{DTF}(\eta)$<620.0 MeV[12]. As shown in Figure 4.17, this requirement is chosen conservatively because it is known that mass resolutions are slightly narrower in MC than data. This cut has a very high signal efficiency ($> 99\%$) whilst removing combinatorial background.

[12]It should be noted that there is already a ±50 MeV window around the kinematically reconstructed η mass in the stripping, as described in Table 4.2

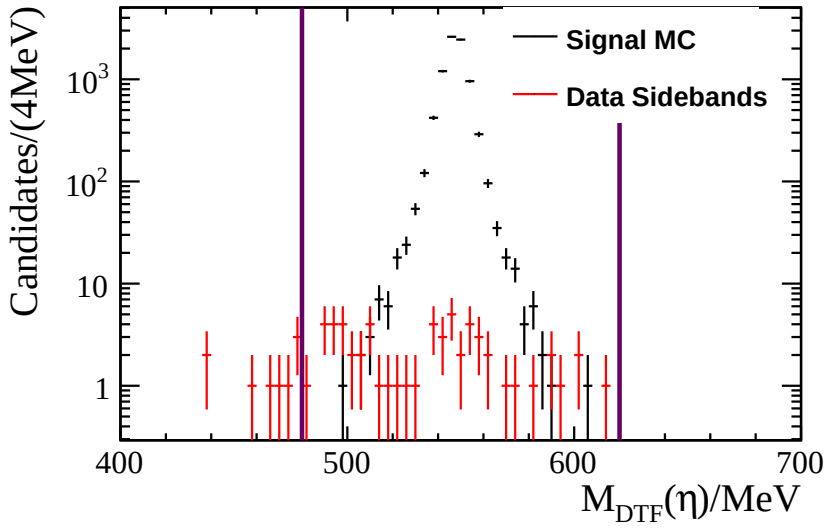


Figure 4.17: Invariant mass of the η determined by DTF in the $\eta' \to \pi^+\pi^-\eta$ channel. The vertical lines indicate the position of the mass window applied, as described in the text.

4.2.7 Multiple Candidates

With the full selection applied, there are some collision events in which more than one reconstructed candidate satisfies all of the requirements. Figure 4.18 shows the distribution of the number of candidates per event in each channel. Multiple candidates most commonly occur when either the same tracks are associated with different soft photons, or the same p, K^- and γ candidates are associated with different $\pi^+\pi^-$ tracks. Table 4.13 shows the fraction of events that pass the full selection and contain >1 candidate. Only one candidate per event is retained; to avoid bias this candidate is chosen randomly.

It is also of interest to investigate the distribution of multiple candidates across the phase space of the signal channel decays; Figure 4.19 shows the distributions of multiple candidates across the Dalitz plot. In the $\eta' \to \pi^+\pi^-\gamma$ channel there is a clear concentration of multiple candidates at low $m^2_{K^-p}$, which is where a rich spectrum of Λ^* resonances is expected. Therefore, these multiple candidates could be arising from the same Λ^* resonance being associated with more than one η'

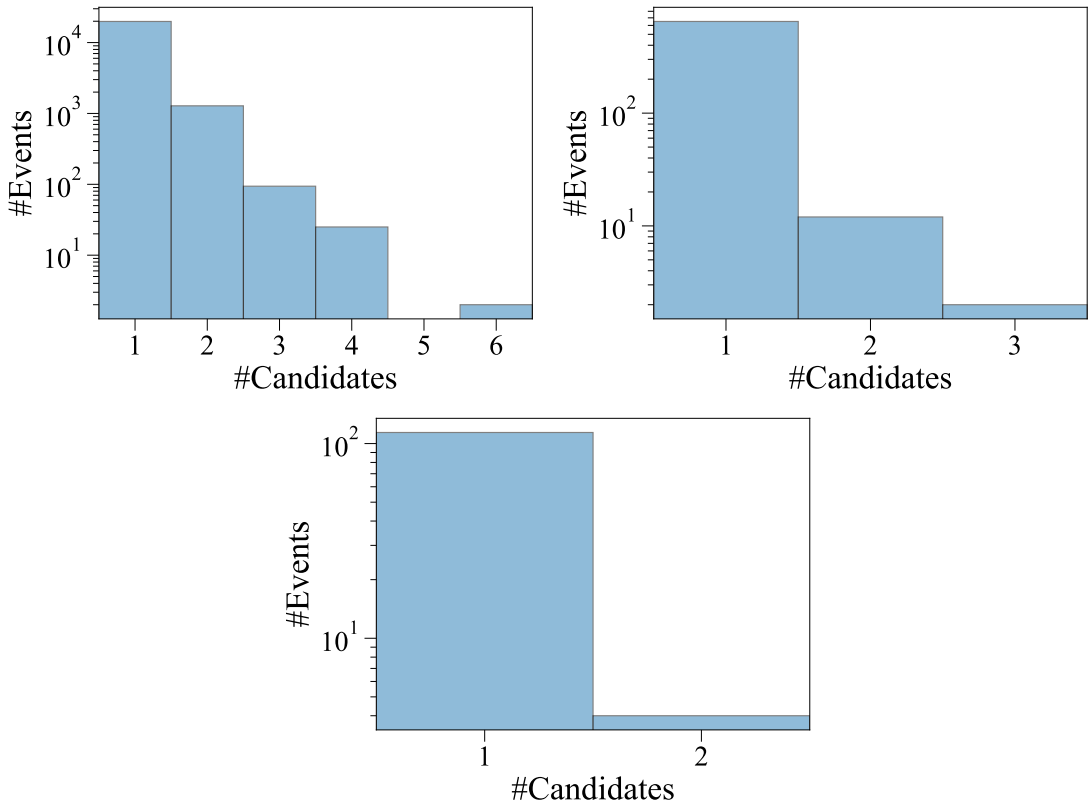


Figure 4.18: Number of candidates per event, for events passing the full selection in the Control channel (Top Left), $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$ channel (Top Right) and $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ channel (Bottom).

candidate. In the $\eta' \to \pi^+\pi^-\eta$ channel the low number of multiple candidates makes it difficult to conclude whether any particular area of the phase space is preferentially populated.

Table 4.13: The fraction of events that contain >1 candidate after all previous stages of the selection in each channel.

Channel	Data	MC
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$(6.61 \pm 0.18)\%$	$(5.33 \pm 0.21)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$(2.02 \pm 0.56)\%$	$(1.15 \pm 0.12)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$	$(4.31 \pm 1.94)\%$	$(5.40 \pm 0.22)\%$

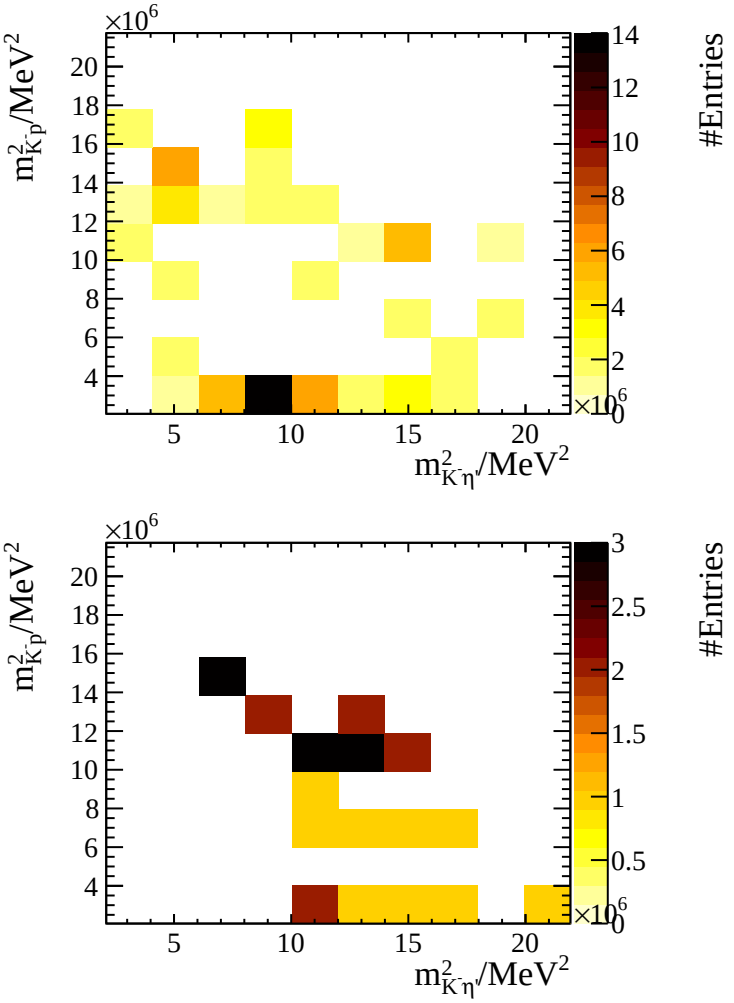


Figure 4.19: Top(Bottom): Distribution of multiple candidates across the phase space of the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) signal channel in data sidebands.

4.3 Efficiencies

As shown in Equation 4.1, the relative efficiency of each rare channel with respect to the control channel is required to calculate the branching fraction of the $\Lambda_b^0 \to pK\eta'$ decay. In order to calculate these efficiencies they are factorised into four main components, each calculated relative to the preceding one, as

$$\varepsilon_{tot} = \varepsilon_{geom} \times \varepsilon_{sel|geom} \times \varepsilon_{PID|sel\&geom} \times \varepsilon_{MultCands|sel\&geom\&PID}, \tag{4.13}$$

where:

- ε_{geom} is the efficiency for having all of the decay products of the signal or control channel within the geometrical acceptance of the LHCb detector[13]. These efficiencies are discussed in Section 4.3.1.
- $\varepsilon_{sel|geom}$ is the efficiency of the entire selection except PID requirements. This includes the reconstruction, trigger, BDT, mass vetoes, and the requirements described in section 4.2.6. These efficiencies are discussed in Section 4.3.2.
- $\varepsilon_{PID|sel\&geom}$ is the efficiency of the particle identification requirements. These efficiencies are discussed in Section 4.3.3.
- $\varepsilon_{MultCands|sel\&geom\&PID}$ is the efficiency of randomly keeping a single candidate per event. These efficiencies are discussed in Section 4.3.4.

There are also photon reconstruction efficiency correction factors which are discussed in Section 4.3.5.

It is possible for the $\Lambda_b^0 \to pK\eta'$ decay to be the result of many intermediate resonances, particularly in the M(pK^-) system. However, as these resonances are

[13]Referred to within LHCb as `DaughtersInLHCb`

not *a priori* known it is not possible to model these in the production of the MC and therefore a phase space model is used. As the signal efficiency of the $\Lambda_b^0 \to pK\eta'$ decay is not expected to be independent of the phase space position, and intermediate resonances would cause the phase space to be unequally populated, a naive calculation of the $\Lambda_b^0 \to pK\eta'$ channel efficiencies using phase space MC is unlikely to be correct.

In order to correct for the variation of efficiencies over the phase space of the $\Lambda_b^0 \to pK\eta'$ decay, all efficiencies need to be calculated as a function of phase space position. The presence of spin-$\frac{1}{2}$ particles in both the initial and final state means the dynamics of this decay are described by a five-dimensional phase space which can be described by the two Dalitz plot variables (m^2_{13} and m^2_{23}) and three angular variables [117]. However, as the polarisation of the Λ_b^0 particle is consistent with 0 the angular variables are spherically symmetric and consistent with phase space MC [118]. This means that the three angular variables do not need to be considered in the efficiency correction procedure, therefore 2D histograms are used to describe the phase space dependence of the efficiency.

When the traditional Dalitz plot variables, m^2_{13} and m^2_{23}, are used to describe efficiencies there can be strong variations near the edges of the Dalitz plot. This occurs because at least one of the final state particles will have low momentum in these regions. It is undesirable to have strong variations over small areas of the Dalitz plot because a binned approach to describing the efficiency assumes smooth variation within each bin. Furthermore, the use of square bins and curved bin boundaries can lead to additional complications. To avoid these issues the traditional Dalitz plot variables are often transformed to the Square Dalitz Plot (SDP) variables, m' and θ', which are given by:

$$m' = \frac{1}{\pi} \arccos\left(2\frac{m_{\eta' p} - m^{min}_{\eta' p}}{m^{max}_{\eta' p} - m^{min}_{\eta' p}} - 1 \right) \tag{4.14}$$

and

$$\theta' = \frac{1}{\pi}\theta_{\eta' p}, \tag{4.15}$$

where $m_{\eta' p}$ is the invariant mass of the η' and proton system, $m_{\eta' p}^{max} = m_{\Lambda_b^0} - m_{K^-}$ and $m_{\eta' p}^{min} = m_{\eta'} + m_p$. $\theta_{\eta' p}$ is the angle between the η' and the K^- in the rest frame of the $\eta' p$ system. Explicitly, these variables can also be described in terms of Lorentz invariant quantities as:

$$m' = \frac{1}{\pi}\arccos\left(2\frac{m_{\eta' p} - (m_{\eta'} + m_p)}{m_{\Lambda_b^0} - (m_{\eta'} + m_p + m_{K^-})} - 1\right) \tag{4.16}$$

and

$$\theta' = \frac{1}{\pi}\arccos\left(\frac{m_{\eta' p}^2(m_{pK^-}^2 - m_{\eta' K^-}^2) - (m_p^2 - m_{\eta'}^2)(m_{\Lambda_b^0}^2 - m_{K^-}^2)}{\sqrt{(m_{\eta' p}^2 + m_{\eta'}^2 - m_p^2)^2 - 4m_{\eta' p}^2 m_{\eta'}^2}\sqrt{(m_{\Lambda_b^0}^2 - m_{K^-}^2 - m_{\eta' p}^2)^2 - 4m_{\eta' p}^2 m_{K^-}^2}}\right). \tag{4.17}$$

The variables m' and θ' are distributed between 0 and 1, which makes the use of square bins far more convenient. The transformation also draws events from the edges of the traditional Dalitz plot towards the centre of the square Dalitz plot; this avoids large variations in efficiency over small areas of the phase space. The disadvantage of this is that when a SDP is filled with phase space MC, bins near the edges of the SDP can have low bin content. Consequently, in order to bin in these variables larger bins have to be used near the edges of the SDP. The rare channel efficiencies as a function of the SDP variables can be found in Appendix B.

However, the use of larger and non-uniform bins is undesirable, because the efficiency variation is less likely to be smooth across each bin and it will lead to increased systematic uncertainties due to the use of a binned approach. Therefore, the variables m'' and $\cos(\theta_{\eta' p})$ are instead used to parameterise the efficiency variation; these are defined as:

$$m'' = \frac{m_{\eta' p} - m_{\eta' p}^{min}}{m_{\eta' p}^{max} - m_{\eta' p}^{min}} = \frac{1}{2}\left(\cos(\pi m') + 1\right) \tag{4.18}$$

and

$$\cos(\theta_{\eta' p}) = \cos(\pi\theta'). \tag{4.19}$$

These variables remove the non-linearity from the SDP variables, which leads to a more uniform spread of phase space MC. Consequently, there is sufficient coverage near the edges of the efficiency map to use a uniform binning scheme. This also means that the efficiency maps can be smoothed with the use of 2D cubic splines, which are given by

$$p(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j, \tag{4.20}$$

where a_{ij} are coefficients to be determined and (x, y) are the coordinates of the midpoints of four adjacent bins. The coefficients a_{ij} are determined in unit cells of four bins by requiring that the function value, first derivative and the mixed second derivative agree with that of the histogram content at the four bin midpoints. A linear approximation is used for the derivatives. The use of cubic splines should further reduce systematic uncertainties because the efficiencies will be more stable against variations in the binning scheme.

The necessity to correct for the variation of the efficiency over the phase space of the $\Lambda_b^0 \to pK\eta'$ decay leads to a two-step unblinding procedure. Firstly, the nominal mass fit will be performed to extract raw signal yields and the significance of any signal in each rare channel will be calculated using Wilks' theorem. In any rare channel where $>3\sigma$ signal significance is observed, *sWeights* will be extracted from the nominal mass fit [111]. The corrected efficiency will then be calculated as

$$\bar{\varepsilon} = \frac{\sum_i w_i}{\sum_i \frac{w_i}{\varepsilon_i}}, \tag{4.21}$$

where w_i is the *sWeight* of event i and ε_i is the per event efficiency obtained for event i. To evaluate the per event efficiencies, m_i'' and $\cos(\theta_i)$ will be calculated for each event in data and $\varepsilon_i(m_i'', \cos(\theta_i))$ will be obtained by evaluating the 2D cubic spline efficiencies at m_i'' and $\cos(\theta_i)$.

In any rare channel where the signal significance is $<3\sigma$ there will not be sufficient knowledge of the signal distribution to perform a phase space correction. Therefore, the phase space integrated efficiency will be used. The bin by bin standard deviation will then be assigned as a systematic uncertainty on the efficiency due to phase space variations.

The following sections detail the efficiencies for all channels; both phase space integrated and as a function of m'' and $\cos(\theta_{\eta' p})$ for the rare channels. The uncertainties presented are purely statistical; systematic uncertainties are discussed in Section 5.6.1. The statistical uncertainties on the phase space integrated efficiencies are calculated as the standard deviation of the binomial distribution. Explicitly,

$$\sigma(\varepsilon) = \sqrt{\frac{\varepsilon(1-\varepsilon)}{N}} = \frac{\sqrt{m(1-m/N)}}{N}, \tag{4.22}$$

where m is the number of candidates passing a cut and N is the total number of candidates in the sample. However, for the case of the (phase space dependent) multiple candidate efficiencies there are bins where 100% of the candidates are retained. In this scenario the standard deviation is not a reliable estimator of the uncertainty on the efficiency. Therefore, for the phase space dependent efficiencies the Clopper Pearson method is used to calculate 68% confidence limits on the efficiency in each bin [119]. This method provides an exact confidence interval, rather than an approximation of the binomial distribution, and is the method recommended by the PDG [1].

4.3.1 Geometric Efficiencies

All of the MC samples used in this analysis are produced with the daughters required to be in the LHCb acceptance, this section assesses the efficiency of this requirement. More specifically, it is required that all charged stable daughters satisfy

Channel	2012	2011	Combined
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	$(17.89 \pm 0.026)\%$	$(17.58 \pm 0.031)\%$	$(17.79 \pm 0.020)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	$(17.21 \pm 0.024)\%$	$(16.90 \pm 0.024)\%$	$(17.11 \pm 0.018)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$	$(15.93 \pm 0.047)\%$	$(15.73 \pm 0.039)\%$	$(15.86 \pm 0.034)\%$

Table 4.14: Phase space integrated geometric efficiencies

10 mrad$<\theta<$400 mrad and the neutral stable daughters satisfy 5 mrad$<\theta<$400 mrad, where θ is the angle between the particle trajectory and beam pipe at generation. In the case of photons, these cuts are only applied if they are the daughter of a π^0 or η.

The phase space integrated geometric efficiencies are shown in Table 4.14. The efficiency maps[14] for each rare channel and the corresponding uncertainties are shown in Figure 4.20 for 2012 data and the corresponding plots for 2011 data can be found in Appendix C.

4.3.2 Selection Efficiencies

The selection efficiency includes reconstruction, trigger, pre-selection and every aspect of the offline selection efficiency except PID cuts. It can be factorised as

$$\varepsilon_{Selection} = \varepsilon_{offline} \times \varepsilon_{L0Trigger} \times \varepsilon_{HLT1} \times \varepsilon_{HLT2} \times \varepsilon_{BDT} \times \varepsilon_{CalibCuts} \times \varepsilon_X, \quad (4.23)$$

where:

- $\varepsilon_{offline}$ is the efficiency of the reconstruction and all the cuts shown in Tables 4.2 and 4.3 except the PID cuts.
- $\varepsilon_{L0Trigger}, \varepsilon_{HLT1}$ and ε_{HLT2} are the efficiencies of the trigger requirements discussed in Section 4.2.1.

[14]The phase space dependent efficiencies are assessed using privately generated samples, where GAUSS is only run up to the EVTGEN stage.

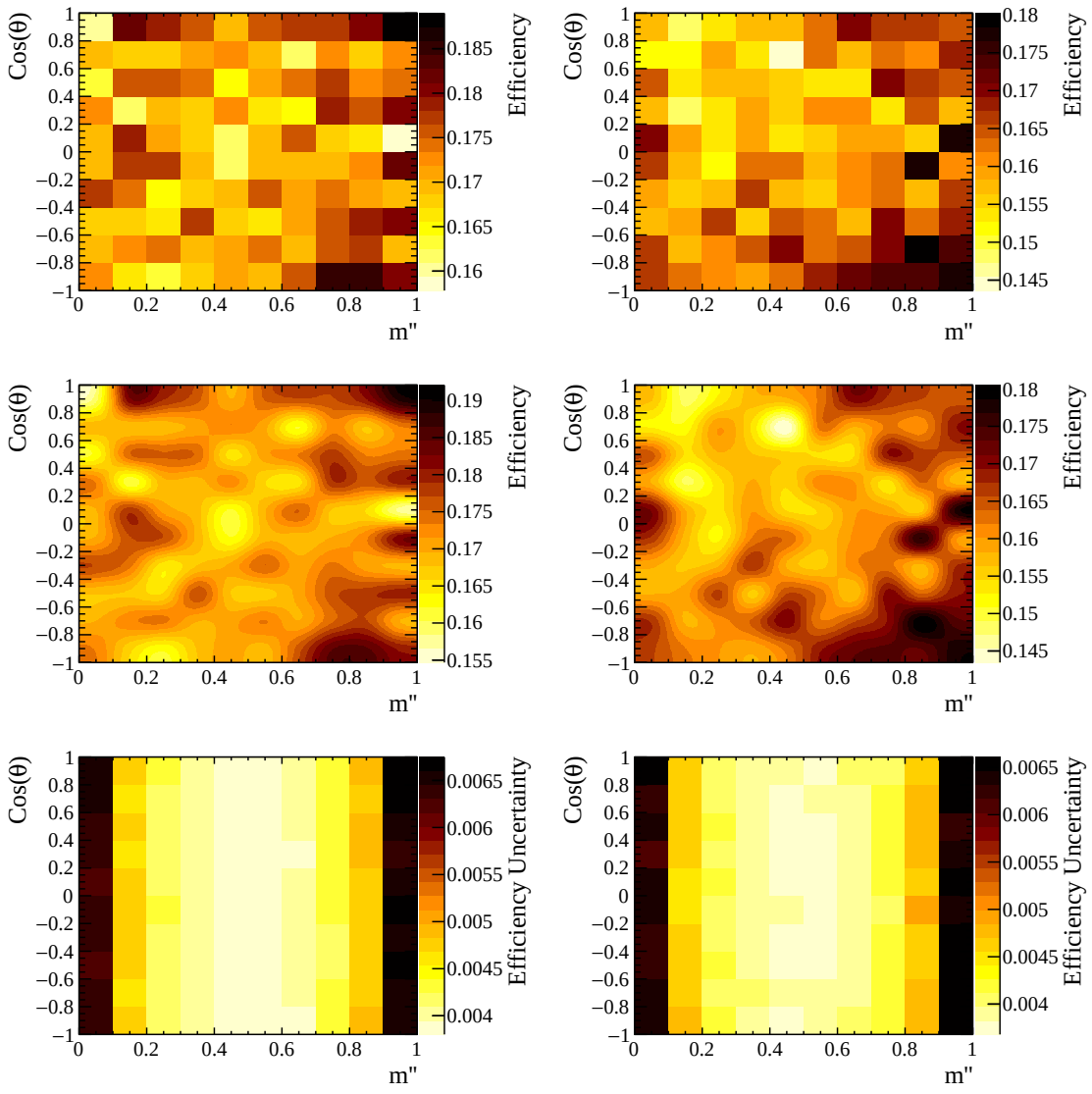


Figure 4.20: Left(Right): Geometric efficiencies (top), with the results of the cubic spline interpolation(middle) and statistical uncertainties(bottom) for the `DaughtersInLHCb` cut as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2012 data. The uncertainties are calculated using the Clopper-Pearson method [119].

Table 4.15: Phase space integrated selection efficiencies for the $B^+ \to K^+\eta'$ channel. See text for description of each efficiency.

Requirement	2012	2011	Combined
$\varepsilon_{offline}$	$(3.452 \pm 0.017)\%$	$(3.726 \pm 0.026)\%$	$(3.541 \pm 0.017)\%$
$\varepsilon_{L0Trigger}$	$(55.99 \pm 0.26)\%$	$(58.66 \pm 0.35)\%$	$(56.85 \pm 0.21)\%$
ε_{HLT1}	$(83.72 \pm 0.25)\%$	$(84.69 \pm 0.34)\%$	$(84.03 \pm 0.20)\%$
ε_{HLT2}	$(92.68 \pm 0.20)\%$	$(86.28 \pm 0.35)\%$	$(90.63 \pm 0.17)\%$
ε_{BDT}	$(55.79 \pm 0.39)\%$	$(57.31 \pm 0.54)\%$	$(56.28 \pm 0.32)\%$
$\varepsilon_{CalibCuts}$	$(99.71 \pm 0.06)\%$	$(99.83 \pm 0.06)\%$	$(99.75 \pm 0.04)\%$
$\varepsilon_{M(\pi^+\pi^-)}$	$(98.26 \pm 0.14)\%$	$(97.94 \pm 0.21)\%$	$(98.16 \pm 0.11)\%$
$\varepsilon_{Selection}$	$(0.8198 \pm 0.0086)\%$	$(0.8951 \pm 0.013)\%$	$(0.8439 \pm 0.0072)\%$

- ε_{BDT} is the efficiency of the requirement on the output of the BDT classifier.
- $\varepsilon_{CalibCuts}$ is the efficiency of the cuts applied to ensure the accuracy of the PID efficiency calibration. This includes: the requirement that all hadrons in all channels contain information from the RICH[15]; the requirement that K^- momentum <450.0 GeV in the control channel and the kinematic cuts on the proton described in Section 4.2.4 for the rare channels.
- ε_X is the efficiency of any further selection cuts applied after the BDT.
  - In the control channel this includes the cut M($\pi^+\pi^-$)>510.0 MeV.
  - In the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel this includes the efficiency of the vetoes discussed in Section 4.2.5 and the M($\pi^+\pi^-$)>510.0 MeV cut discussed in Section 4.2.6.
  - In the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) channel it is the efficiency of the η mass window described in Section 4.2.6.

Detailed breakdowns of the phase space integrated efficiencies for each channel are shown in Tables 4.15, 4.16 and 4.17. The 2D selection efficiency maps for the rare channels in 2012 data are shown in Figure 4.21 and the same maps for 2011 data are shown in Appendix C.

[15]The so called `HasRich` requirement

Table 4.16: Phase space integrated selection efficiencies for the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel. See text for a description of each efficiency.

Requirement	2012	2011	Combined
$\varepsilon_{offline}$	$(2.044 \pm 0.007)\%$	$(2.230 \pm 0.011)\%$	$(2.13 \pm 0.06)\%$
$\varepsilon_{L0Trigger}$	$(52.14 \pm 0.20)\%$	$(55.35 \pm 0.23)\%$	$(53.17 \pm 0.13)\%$
ε_{HLT1}	$(82.69 \pm 0.18)\%$	$(82.09 \pm 0.24)\%$	$(82.50 \pm 0.14)\%$
ε_{HLT2}	$(89.40 \pm 0.16)\%$	$(85.35 \pm 0.24)\%$	$(88.10 \pm 0.13)\%$
ε_{BDT}	$(23.40 \pm 0.23)\%$	$(27.64 \pm 0.33)\%$	$(24.76 \pm 0.19)\%$
$\varepsilon_{CalibCuts}$	$(94.74 \pm 0.25)\%$	$(96.20 \pm 0.27)\%$	$(95.21 \pm 0.19)\%$
ε_{Vetos}	$(88.81 \pm 0.37)\%$	$(89.58 \pm 0.43)\%$	$(89.06 \pm 0.29)\%$
$\varepsilon_{M(\pi^+\pi^-)}$	$(95.58 \pm 0.26)\%$	$(96.04 \pm 0.29)\%$	$(95.73 \pm 0.20)\%$
$\varepsilon_{Selection}$	$(0.1483 \pm 0.0019)\%$	$(0.2039 \pm 0.0031)\%$	$(0.1661 \pm 0.0016)\%$

Table 4.17: Phase space integrated selection efficiencies for the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) channel. See text for a description of each efficiency.

Requirement	2012	2011	Combined
$\varepsilon_{offline}$	$(0.929 \pm 0.004)\%$	$(1.029 \pm 0.006)\%$	$(0.961 \pm 0.003)\%$
$\varepsilon_{L0Trigger}$	$(47.94 \pm 0.21)\%$	$(50.82 \pm 0.28)\%$	$(48.86 \pm 0.17)\%$
ε_{HLT1}	$(76.43 \pm 0.26)\%$	$(74.61 \pm 0.34)\%$	$(75.84 \pm 0.21)\%$
ε_{HLT2}	$(79.34 \pm 0.28)\%$	$(73.98 \pm 0.40)\%$	$(77.62 \pm 0.23)\%$
ε_{BDT}	$(38.13 \pm 0.38)\%$	$(56.41 \pm 0.53)\%$	$(43.98 \pm 0.31)\%$
$\varepsilon_{CalibCuts}$	$(94.65 \pm 0.29)\%$	$(96.33 \pm 0.27)\%$	$(95.19 \pm 0.21)\%$
$\varepsilon_{EtaWindow}$	$(100.0 \pm 0.0)\%$	$(99.89 \pm 0.05)\%$	$(99.97 \pm 0.01)\%$
$\varepsilon_{Selection}$	$(0.0975 \pm 0.0013)\%$	$(0.1568 \pm 0.0023)\%$	$(0.1164 \pm 0.0013)\%$

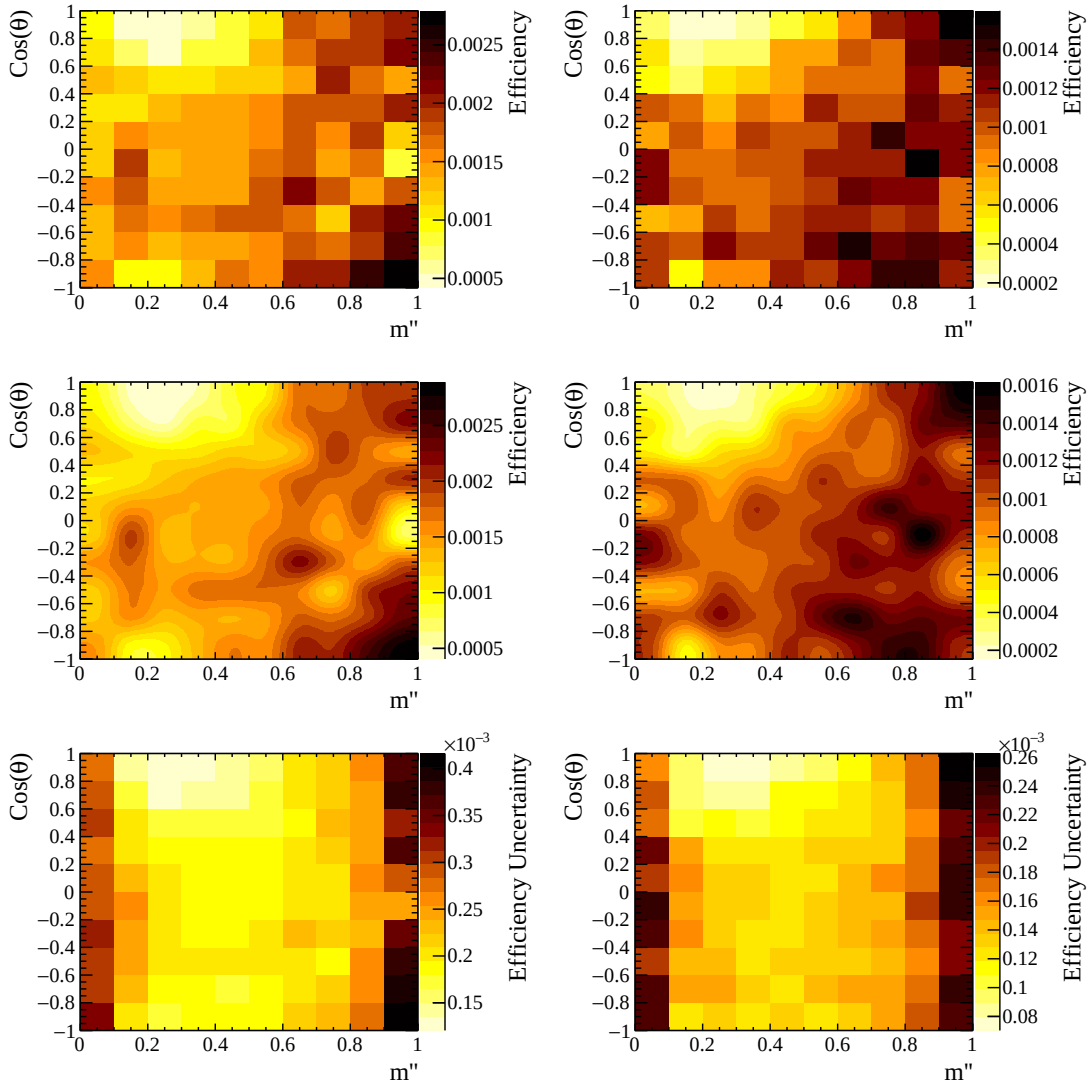


Figure 4.21: Left(Right): Selection Efficiencies (Top) along with the results of the cubic spline interpolation (middle) and uncertainties(bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2012.

4.3.3 PID Efficiencies

The PID efficiencies include all of the PID requirements placed on charged particles. Specifically, these are the loose PID requirements applied in stripping (see Table 4.2 and Table 4.3) and the requirements optimised in section 4.2.4. It is well established that the PID variables are not well modelled in MC, therefore MC cannot be used to calculate efficiencies accurately; a data driven approach is needed. The `PIDCalib` package is used to calculate all PID efficiencies, which makes use of background-subtracted data calibration samples [120]. In this analysis, samples of $D^* \to D^0(K^-\pi^+)\pi^+$ decays are used to determine the PID efficiency of kaons and pions. For protons, prompt $\Lambda_c^+ \to pK^-\pi^+$ and $\Lambda_b^0 \to (\Lambda_c^+ \to pK^-\pi^+)\mu\nu$ decays are used. The PID of the daughter tracks in these decays can be determined from kinematics alone. In the $D^* \to D^0(K^-\pi^+)\pi^+$ decay for example, the bachelor π^+ has significantly less momentum than the daughters of the D^0. Therefore, once this track is identified the track with the same sign must be the π^+, and the opposite sign track must be the K^- (the decay $D^0 \to K^+\pi^-$ is suppressed by two orders of magnitude [1]).

The PID efficiencies cannot be taken straight from these calibration samples because the kinematics are very different to those of the decays studied in this analysis. It has been shown that the performance of the PID system is a function of track momentum p, pseudorapidity η and the total number of tracks in an event (nTracks). Therefore, the data calibration samples are used to determine the efficiencies of the PID cuts applied as a function of these variables, $\varepsilon(p, \eta, \text{nTracks})$. As there is particularly low statistics in the proton calibration samples, the samples are merged for both year and magnet polarity to make best use of the statistics available. Therefore, a single performance histogram is created for each particle species. The performance histograms created from the calibration samples[16], for the cuts applied in the $\Lambda_b^0 \to$

[16]The 3D histograms used in the PID calibration are sliced into 2D histograms for visualisation purposes

Table 4.18: Phase space integrated PID efficiencies and uncertainties due to calibration sample size $\delta_{CalibStat.}$, reference sample size $\delta_{Ref.Stat.}$ and total uncertainty $\sigma(\varepsilon_{PID})$.

Channel	ε_{PID}	$\delta_{CalibStat.}$	$\delta_{ref.stat.}$	$\sigma(\varepsilon_{PID})$
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	73.49%	0.01%	0.21%	0.21%
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	59.66%	0.03%	0.21%	0.21%
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$	66.03%	0.04%	0.23%	0.23%

$pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ channel, are shown in Figures 4.22, 4.23 and 4.24. The binning schemes have been chosen to ensure smooth variation of efficiencies within each bin whilst keeping statistical uncertainties to a minimum.

The performance histograms are then used to assign efficiencies to each track in signal MC (also refered to as the reference sample) and the efficiency for an event is the product of the per track efficiencies, $\prod_t^{N_{trk}} \varepsilon_t(p_t, \eta_t, \text{nTracks})$. The overall efficiency, ε_{PID}, is then calculated as the mean of all per event efficiencies. Explicitly, the PID efficiency is given by

$$\varepsilon_{\text{PID}} = \frac{1}{N_{evt}} \sum_i^{N_{evt}} \prod_t^{N_{trk,i}} \varepsilon_t(p_t, \eta_t, \text{nTracks}_i), \tag{4.24}$$

where N_{evt} is the number of events in MC and the product is over all tracks, $N_{trk,i}$, in the event i. To determine the PID efficiency as a function of m'' and $\cos(\theta_{\eta' p})$, the averaging is performed in each bin. The PID efficiency distributions, as a function of m'' and $\cos(\theta_{\eta' p})$, are shown in Figure 4.25 and the phase space integrated efficiencies are shown in Table 4.18.

4.3.3.1 Resampling of nTracks

It is known that MC does not reproduce the nTracks distribution correctly. Figure 4.26 shows a comparison of the nTracks distribution between MC for all three channels and the background subtracted distribution observed in the control channel. As these distributions are all very similar in MC (especially when it is considered

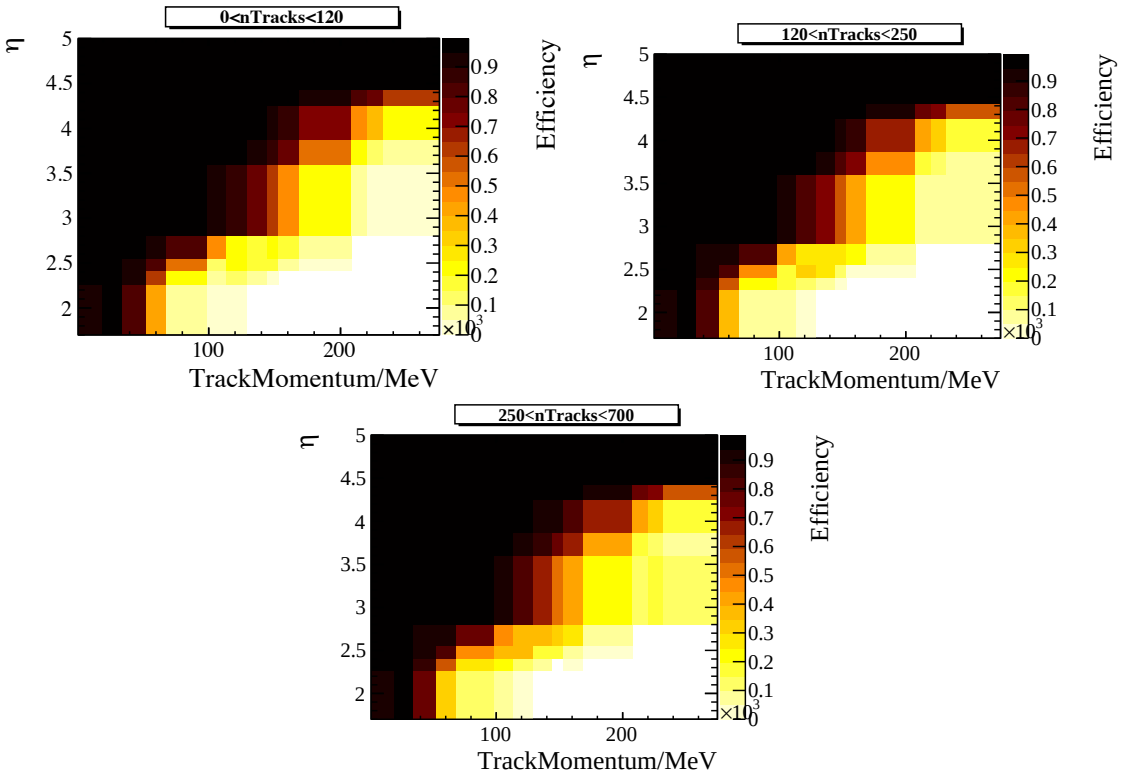


Figure 4.22: Pion PID cut efficiency (for the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ channel) as a function of p and η in each bin of nTracks.

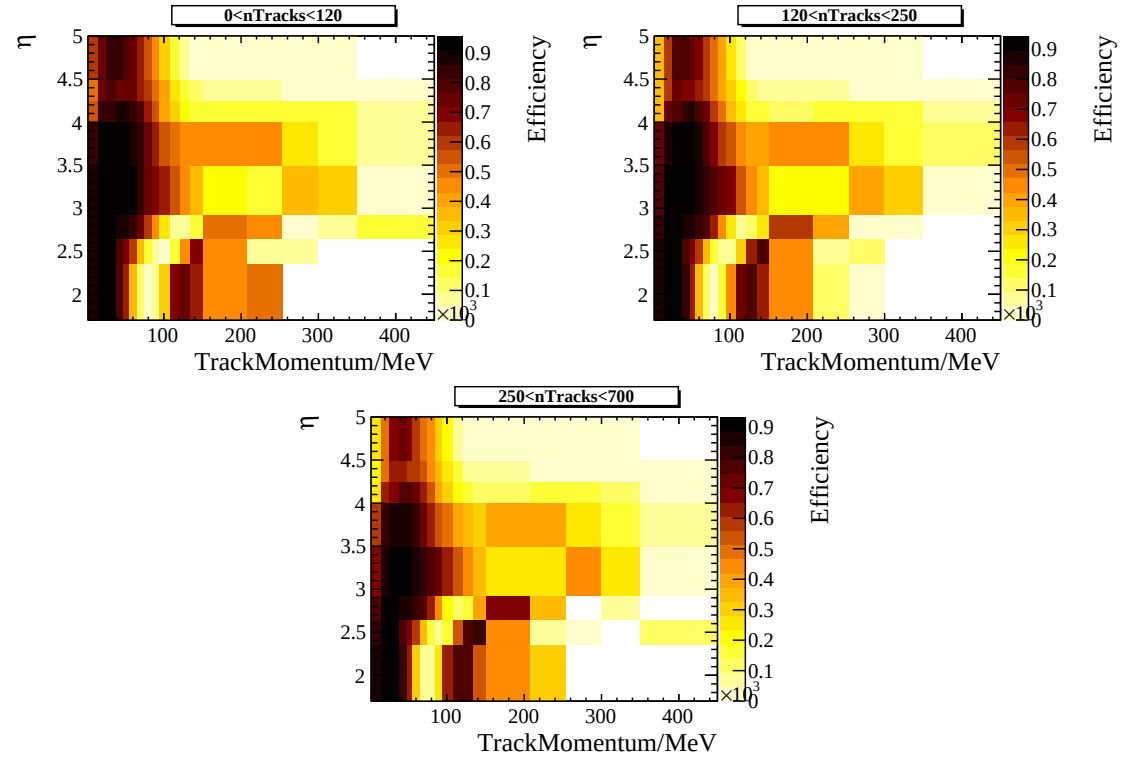


Figure 4.23: Kaon PID cut efficiency (for the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ channel) as a function of p and η in each bin of nTracks.

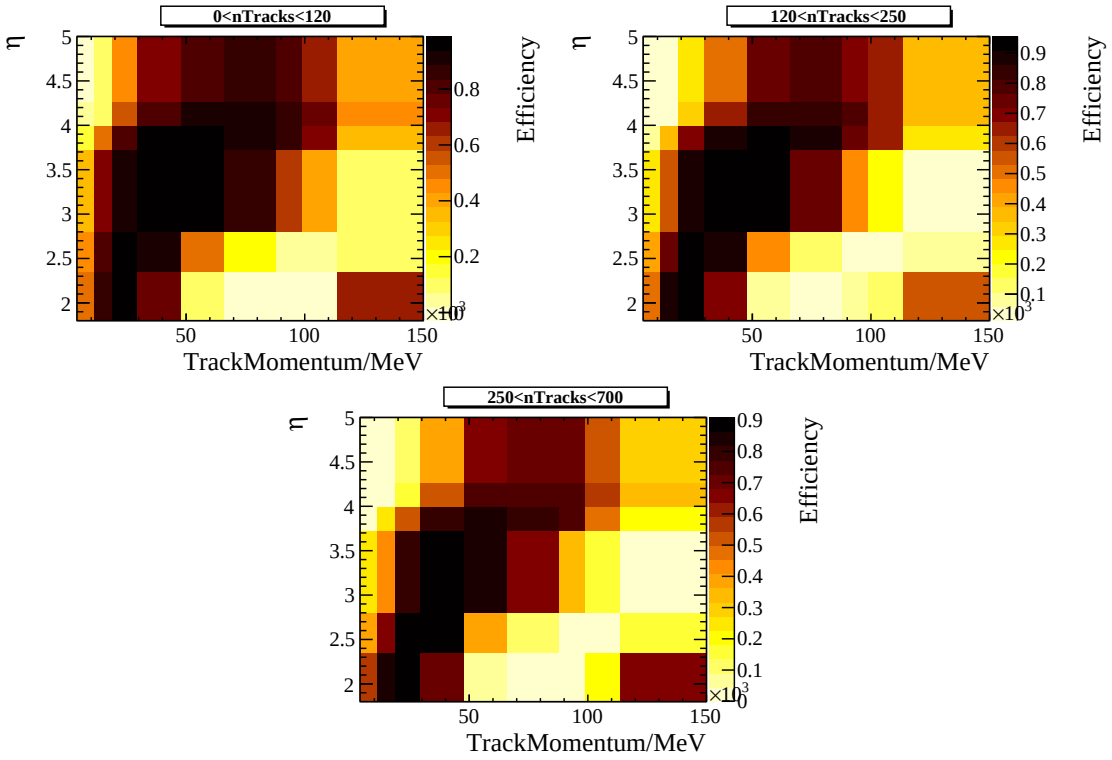


Figure 4.24: Proton PID cut efficiency (for the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel) as a function of p and η in each bin of nTracks. It should be noted that the bottom right corner of this plot is not used; see Section 4.2.4 for a description of the kinematic cuts applied

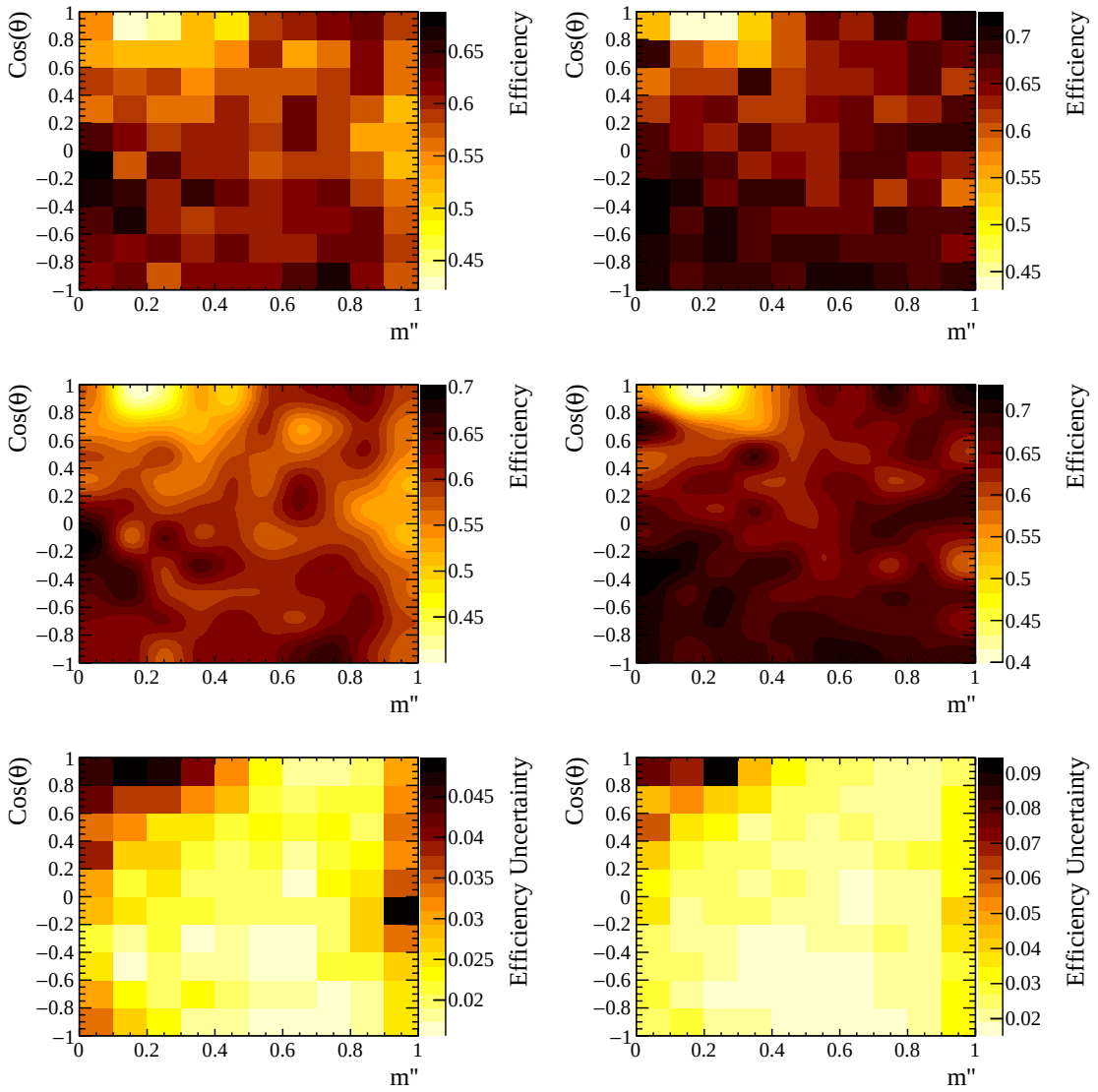


Figure 4.25: Left(Right): PID Efficiencies (Top) along with the results of the cubic spline interpolation (middle) and total uncertainties(bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

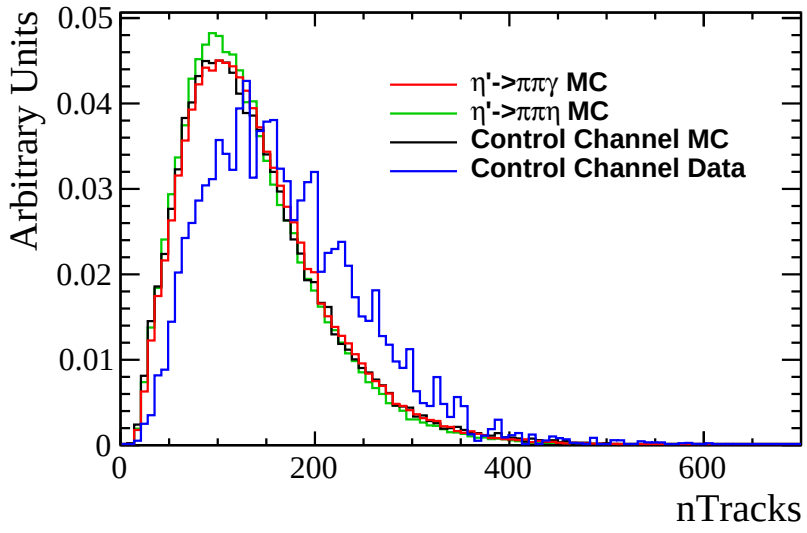


Figure 4.26: A comparison of the nTracks distributions between MC for all channels and background subtracted data in the control channel

only three bins of nTracks are used) the nTracks distributions used in the calibration, for all channels, are taken from control channel data. Per event values of nTracks are assigned to MC events by randomly sampling from the background subtracted distribution observed in the control channel.

4.3.3.2 Statistical Uncertainties

Two sources of statistical uncertainty are considered on these PID efficiencies: from the limited size of the calibration sample and from the limited size of the reference sample. The reference sample is the sample from which signal kinematics are taken, which in this case is signal MC. The uncertainty due to the limited size of the calibration sample, $\delta_{Calib.Stat.}$, is assessed by creating 1000 calibration histograms of $\varepsilon(p, \eta, nTracks)$ where the bin content is sampled from a Gaussian distribution with a mean equal to the efficiency in the corresponding bin of the nominal performance histogram and a width equal to the statistical uncertainty on the efficiency in the corresponding bin of the nominal performance histogram. The PID efficiency is then recalculated using each of the 1000 sampled histograms and the phase space inte-

Table 4.19: Phase space integrated Multiple Candidate efficiencies

Channel	2012	2011	Combined
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	$(96.98 \pm 0.20)\%$	$(97.46 \pm 0.25)\%$	$(97.14 \pm 0.16)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	$(99.40 \pm 0.12)\%$	$(99.21 \pm 0.16)\%$	$(99.34 \pm 0.09)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$	$(97.80 \pm 0.22)\%$	$(95.79 \pm 0.33)\%$	$(97.16 \pm 0.18)\%$

grated uncertainties are extracted by fitting the resulting PID efficiency distribution with a Gaussian and the width, σ, is taken as the uncertainty. The same procedure is used to extract $\delta_{Calib.Stat.}$ as a function of the phase space variables by performing the Gaussian fit in each bin of the efficiency map.

The uncertainty due to the reference sample size, $\delta_{ref.stat.}$, is assessed by creating 1000 bootstrap samples from the signal MC samples and recalculating the PID efficiency using each bootstrap sample[17]. The uncertainties on both the phase space integrated and phase space dependent PID efficiencies are again extracted from the resulting distribution using Gaussian fits. Table 4.18 shows these uncertainties along with the combined total uncertainty (added in quadrature) and the total uncertainty as a function of m'' and $\cos(\theta_{\eta' p})$ is shown in Figure 4.25.

4.3.4 Removal of Multiple Candidates Efficiency

After the full selection has been applied there are a small number of candidates that come from the same event as one or more other candidate(s) passing the selection. To avoid bias, one candidate per event is chosen randomly. The efficiency of this procedure is assessed by applying the same procedure to MC which contains the full underlying event; the truth matching is done after a single candidate per event has been randomly flagged as the candidate to keep. The phase space integrated efficiencies for this procedure are shown in Table 4.19 and the efficiencies as a function of m'' and $\cos(\theta_{\eta' p})$ are shown in Figure 4.27 for 2012 data and Appendix C for 2011 data.

[17] The bootstrap samples are created by sampling from the signal MC samples with replacement.

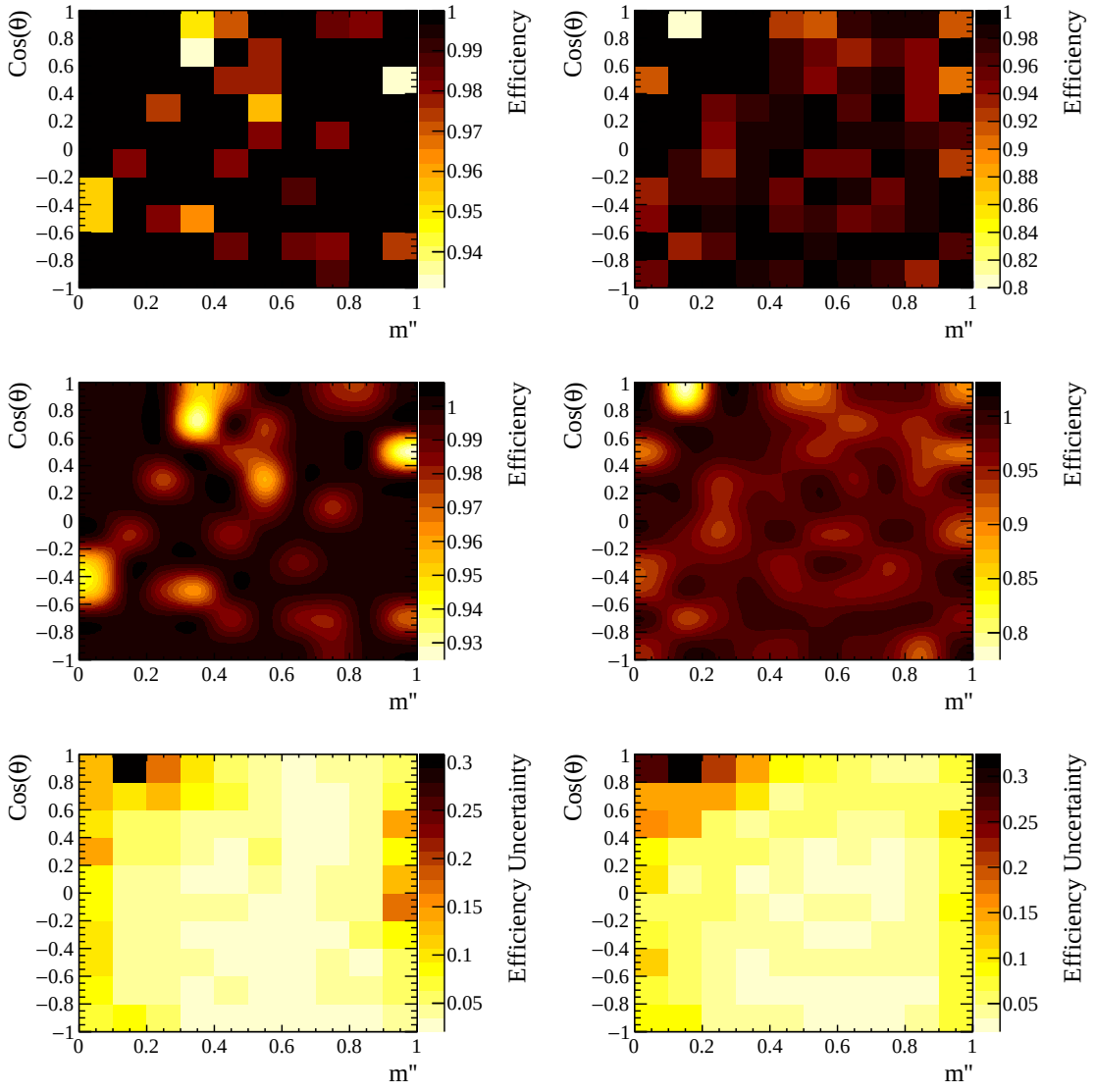


Figure 4.27: Left(Right): Multiple candidate efficiencies (Top) along with the results of the cubic spline interpolation (middle) and total uncertainties(bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2012 data.

4.3.5 Photon Efficiency Corrections

The photon reconstruction efficiencies are not perfectly modelled in MC, therefore reconstruction efficiency corrections have been determined as a function of photon E_{T}. These have been calculated by comparing the observed number of $B^+ \to J/\psi\,(K^{*+} \to K^+\pi^0)$ and $B^+ \to J/\psi\,K^+$ events [121].

The correction factor as a function of m'' and $\cos(\theta_{\eta' p})$ for both rare channels can be seen in Figure 4.28 for 2012 data and in Appendix C for 2011 data. This is calculated by assessing a weighted average correction factor, based on signal MC photon E_{T} distributions, in each bin. A separate correction factor is calculated for the control channel total efficiency, again using a weighted average based on the photon E_{T} distributions seen in MC. The value of this correction factor is 1.007 and it will be applied to the control channel efficiency when the branching fraction result is evaluated. A statistical uncertainty is not quoted here because the uncertainty due to the limited size of the signal MC sample has already been assigned to the selection efficiency. The assessment of a systematic uncertainty on the overall efficiency ratio due to the uncertainty on these correction factors is described in Section 5.6.1.6.

4.3.6 Total Efficiencies

As the analysis strategy is to fit 2012 and 2011 data together, total efficiencies are required for both years combined in each channel. The total efficiencies are obtained by first taking the product of the geometric efficiencies, selection efficiencies, multiple candidate efficiencies and photon reconstruction correction factors for each year. A weighted average of the two resulting efficiencies is then calculated, where the weights are the product of $b\bar{b}$ cross section and luminosity for each year of data taking. This weighted average efficiency is then multiplied by the PID efficiency, which is calculated for both years combined, to give a combined total efficiency for

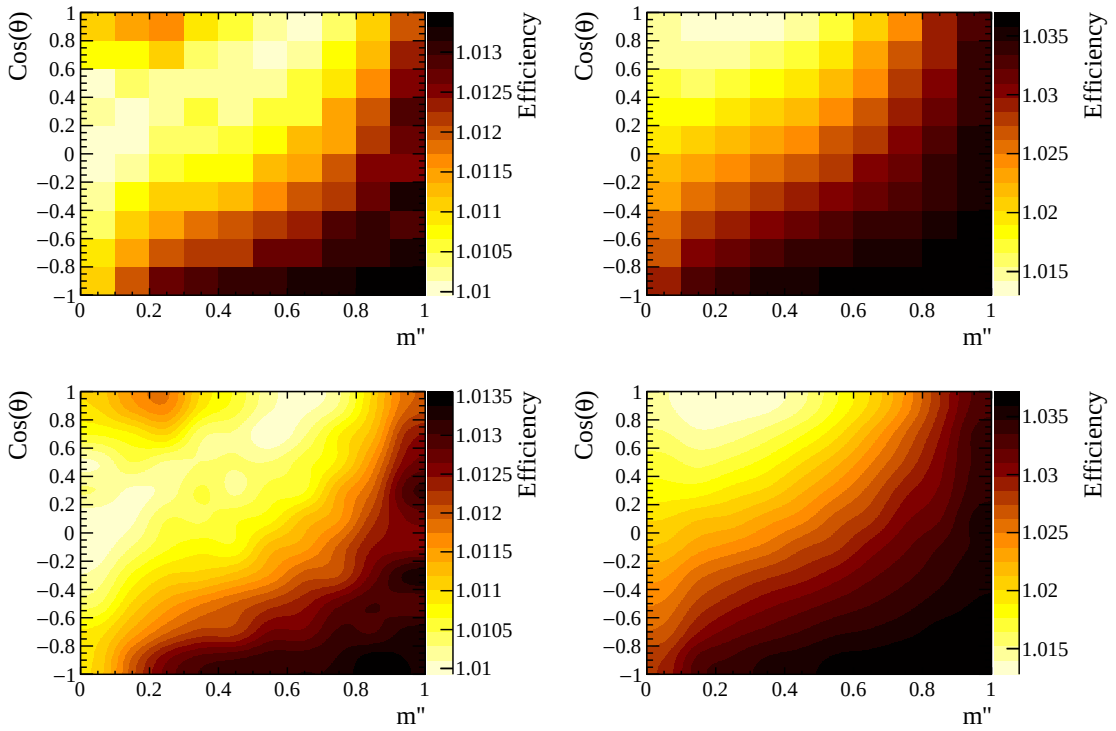


Figure 4.28: Left(Right): Photon reconstruction efficiency correction (Top) along with the results of the cubic spline interpolation (bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2012.

Channel	Total Efficiency
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	$(0.10715 \pm 0.00098)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	$(0.01682 \pm 0.00018)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$	$(0.01184 \pm 0.00013)\%$

Table 4.20: Total phase space integrated efficiencies

Channel	Efficiency	Standard Deviation	Relative Uncertainty
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$	0.0170%	0.0052%	31%
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$	0.0122%	0.0040%	33%

Table 4.21: Phase space averaged overall efficiencies and standard deviations. These are calculated by taking the mean over all efficiency map bins.

each channel.

The total efficiencies as a function of m'' and $\cos(\theta_{\eta' p})$ are shown in Figure 4.29. In the event of a significant signal yield being observed in either channel, per-event efficiencies will be obtained for each event in data by evaluating the cubic spline functions at the value of m'' and $\cos(\theta_{\eta' p})$ observed in data. These per-event efficiencies will be used in Equation.(4.21) to obtain an efficiency that is corrected for phase space variations.

The total phase space integrated efficiencies, which will be used in the event of no significant signal yield (and always used in the control channel), are shown in Table 4.20. Histograms are filled for each bin of the total efficiency maps, as shown in Figure 4.30. The standard deviation taken from these histograms will be assigned as a systematic uncertainty in the event of no significant signal yield; these values are shown in Table 4.21.

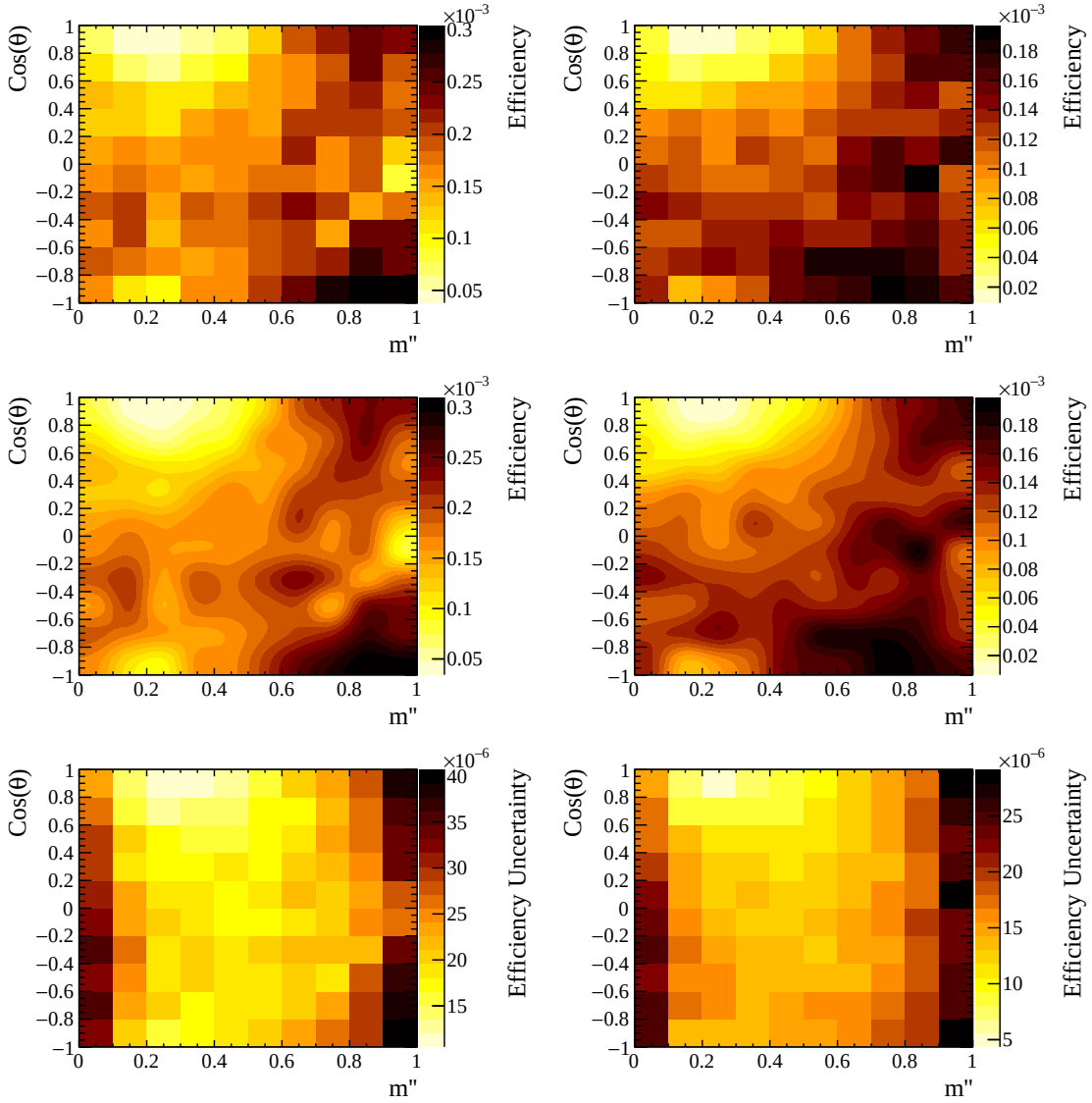


Figure 4.29: Left(Right): Total Efficiencies (Top) along with the results of the cubic spline interpolation (middle) and total uncertainties(bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

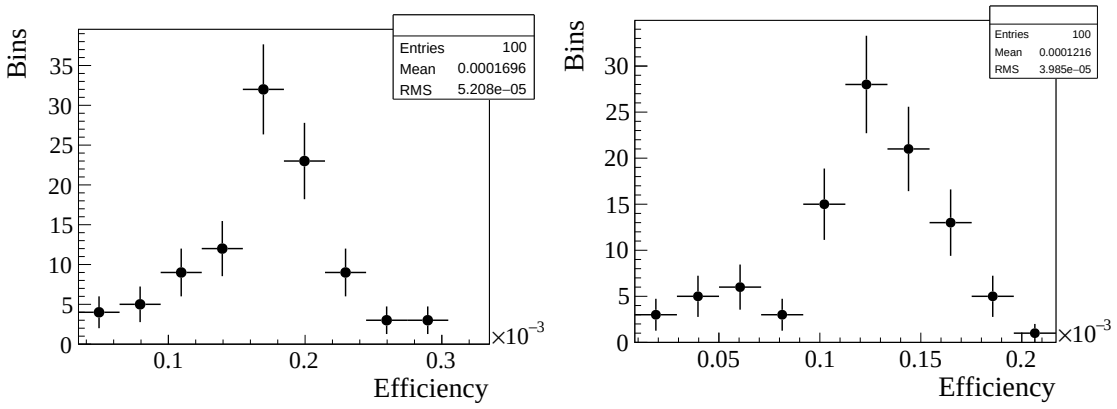


Figure 4.30: Left (Right): 1D projection of the total efficiency for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

Background	Predicted Yield in Fit Region (Signal window)	Conclusion
$\eta' \to \pi^+\pi^-\gamma$ **Channel:**		
$B^0 \to K^{*0}\eta'$	0.28 ± 0.04 (0.10 ± 0.02)	Negligible
$B_s^0 \to \phi\phi$	1.7 ± 0.3 (0.4 ± 0.1)	Negligible
$\Lambda_b^0 \to pK^-\phi$	19 (5)	Considered in systematics
$\Lambda_b^0 \to pD^0\pi^-$	7.2 ± 1.8 (0.8 ± 0.6)	Negligible
$\Lambda_b^0 \to 4h + \pi^0$	133 ± 50	Included in fit
Combinatorial		Modelled with falling exponential
$\eta' \to \pi^+\pi^-\eta$ **Channel:**		
Combinatorial		Modelled with falling exponential

Table 4.22: A summary of all the potential sources of background discussed in this section. No uncertainty is assigned for the decay $\Lambda_b^0 \to pK^-\phi$ because the prediction is based on an assumed branching fraction, therefore the knowledge of the branching fraction will dominate over the statistical uncertainty.

4.4 Background Investigations

This section describes the investigations that have been performed for the various backgrounds that could potentially be present in the rare channel fit regions after the full selection has been applied. Section 4.4.1 describes possible backgrounds that involve the mis-identification of a particle and Section 4.4.2 describes backgrounds where one particle in a decay chain is not reconstructed. A summary of all backgrounds is shown in Table 4.22.

Throughout this section predicted yields of potential backgrounds are calculated, by normalising to the control channel, as

$$N_{Background} = N_C \times \frac{\mathcal{B}_{Background}}{\mathcal{B}_C} \times \frac{\epsilon_{Background}}{\epsilon_C} \times \frac{f_{Background}}{f_u}, \tag{4.25}$$

where $N_{Background}(N_C)$ is the predicted(measured) yield of the background(control) channel, $\mathcal{B}_{Background}(\mathcal{B}_C)$ is the branching fraction of the background(control) channel, $\varepsilon_{Background}(\varepsilon_C)$ is the overall selection efficiency for the background(control) channel and $\frac{f_{Background}}{f_u}$ is the ratio of background to B^+ fragmentation fractions.

4.4.1 Mis-ID Backgrounds

One possible source of mis-ID background in the rare channels is $B^0 \to (K^{*0} \to K^-\pi^+)\eta'$ decays where the π^+ from the K^{*0} is mis-identified as a proton. To investigate this decay the full selection has been applied to a sample of fully simulated $B^0 \to K^{*0}\eta'$ events in order to obtain selection efficiencies and calculate a predicted yield. For the decay $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$, Equation 4.22 predicts a yield of 0.10 ± 0.02 $B^0 \to K^{*0}\eta'$ events will be found in the signal mass window, and 0.28 ± 0.04 in the full fit region. The full fit region is defined as $5200\,\text{MeV} < M(pK\eta') < 5950\,\text{MeV}$ and, for this purpose, the signal window is defined as $|M(pK\eta') - M(\Lambda_b^0)| < 60.0\,\text{MeV}$. This is a negligible number of events compared to the ≈ 430 combinatorial background events present in the full fit region. A similarly negligible number of events is assumed in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$ channel because it would require the same mis-identification to take place and the particle ID cuts are similar between channels.

Another possible background in the $\eta' \to \pi^+\pi^-\gamma$ channel involving mis-identified particles is the decay $B_s^0 \to (\phi \to K^+K^-)(\phi \to \pi^+\pi^-\pi^0)$, where one of the photons from the π^0 is not reconstructed and the K^+ is mis-identified as a proton; missing a photon from the π^0 causes the reconstructed ϕ mass to peak very close to the η' mass. Again, fully simulated MC samples are used to calculate the predicted number of these events surviving the full selection; only 1.70 ± 0.25 events are predicted to be present in the mass fit region and 0.4 ± 0.1 in the signal region. This level of background is also negligible compared to the level of combinatorial background present.

4.4.2 Partially Reconstructed Backgrounds

Partially reconstructed backgrounds (PRB) occur when one or more particles in a decay chain are not reconstructed. This most commonly causes a broad and shifted peak in the spectrum of signal candidates. It is most common for neutral particles to be missed as their reconstruction relies on the association of a calorimeter cluster with a track vertex. In the $\eta' \to \pi^+\pi^-\gamma$ rare channel the level of background is significantly higher in the Lower Sideband (LSB) compared to the Upper Sideband (USB)[18]. This suggests the presence of PRB.

Many sources of PRB have been considered. The most likely candidates are five body Λ_b^0 decays involving a π^0, such as $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$, where only one of the photons from the π^0 is reconstructed. Decays of this type can proceed through various two body intermediate resonances such as $\Lambda_b^0 \to pK^-\pi^+(\rho^- \to \pi^-\pi^0)$, $\Lambda_b^0 \to p\pi^-\pi^+(K^{*-} \to K^-\pi^0)$ and $\Lambda_b^0 \to (\Delta \to p\pi^0)\pi^+\pi^-K^-$. In other analyses using the similar channel $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ (see Ref. [67]) a significant number of these decays have been seen as partially reconstructed background where a π^0 is missed after the full selection.

To investigate the plausibility of these backgrounds further, the `RapidSim` package is used to generate large smeared samples of these decays (see section 3.1 for further description of `RapidSim`) [94]. All of the kinematic cuts from stripping, the η' mass window, $M(\pi^+\pi^-) > 510\,\text{MeV}$ cut and the specific vetoes are applied to these samples. The efficiencies of these cuts are shown in Table 4.23. It can be seen that the efficiency of the decay $\Lambda_b^0 \to pK^-\pi^+(\rho^- \to \pi^-\pi^0)$ is twice that of the other decays. This makes sense as the signal $\eta' \to \pi^+\pi^-\gamma$ decay usually proceeds through a ρ^0 resonance so this background will be kinematically similar. Given the

[18]The sidebands are the regions of the mass distribution inside the full fit window which are not blind. Explicitly, in the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel the LSB is defined as $5200\,\text{MeV} < M(pK\eta') < 5494(5444)\,\text{MeV}$. The USB is defined as $5744(5794)\,\text{MeV} < M(pK\eta') < 5950$ in the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

Decay	Efficiency
$\Lambda_b^0 \to pK^-\pi^+(\rho^- \to \pi^-\pi^0)$	$(0.3979 \pm 0.0039)\%$
$\Lambda_b^0 \to p\pi^-\pi^+(K^{*-} \to K^-\pi^0)$	$(0.2117 \pm 0.0028)\%$
$\Lambda_b^0 \to (\Delta \to p\pi^0)\pi^+\pi^-K^-$	$(0.2385 \pm 0.0030)\%$

Table 4.23: Efficiencies of all kinematic cuts, η' mass window, $M(\pi^+\pi^-) > 510\,\text{MeV}$ requirement and specific vetoes on `RapidSim` samples of the stated decays. In each case the η' candidate is constructed from the π^+, π^- and one of the photons from the π^0.

abundance of these decays seen in other analyses and the efficiencies presented in Table 4.23, it is believed that the η' window and kinematic cuts will not completely remove these backgrounds.

In order to try to compare the shape of these decays to what is seen in data the reconstructed Λ_b^0 mass is plotted for events passing the full selection and falling in the η' sidebands ($858\,\text{MeV} < M(\pi^+\pi^-\gamma) < 880\,\text{MeV}$ or $1020\,\text{MeV} < M(\pi^+\pi^-\gamma) < 1058\,\text{MeV}$). A fit to these data is then performed using a model comprising an exponential for combinatorial background and a bifurcated Gaussian for PRB. A bifurcated Gaussian is a standard Gaussian function with different widths, σ_R and σ_L, on either side of the mean. Explicitly, it is given by

$$G_b\left(x;\mu,\sigma\right) = \frac{1}{N}\exp\left(\frac{-\left(x-\mu\right)^2}{2\sigma^2}\right), \tag{4.26}$$

where N is the normalisation factor, x is the parameter to which the fit is being performed (in this case Λ_b^0 mass), μ is the mean of the Gaussian and $\sigma = \sigma_L$ for $x < \mu$ and $\sigma = \sigma_R$ for $x > \mu$. The parameters of the PRB shape are fixed from a fit to simulated $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ decays; this can be seen in Figure 4.31. The results of this fit to data can be seen in Figure 4.32.

Firstly, looking at the data points alone, it is clear there is significantly more background at M($pK^-\ \eta'$) $<\Lambda_b^0$ mass (5619.51 MeV) than M($pK^-\ \eta'$) >M(Λ_b^0), which is consistent with a PRB where a photon is missed. Secondly, this model provides a good quality fit to the sideband data with a PRB yield of 40 ± 15 events. For these

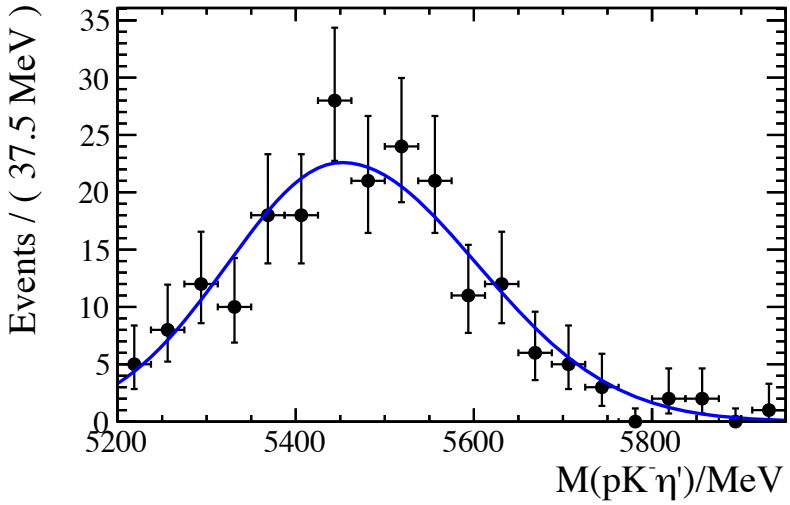


Figure 4.31: Fit performed to $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ MC that passes the full selection. The fit function is a bifurcated Gaussian.

reasons it is concluded that there is PRB from five body Λ_b^0 decays present after the full selection. A PDF will be included to account for these decays in the final mass fit.

To try and estimate how many five body Λ_b^0 background events we can expect in the final fit, the PRB yield from the fit in Figure 4.31 is scaled by the estimated fraction of five body events selected by requiring events to fall in the η' sidebands. In the simulated $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ decays 30% of events fall in the η' sidebands, therefore the estimated yield of Λ_b^0 to five body background in the full data samples is 133 ± 50 events. Although this is a very rough estimate, it shows that this is likely to be the dominant source of non-combinatorial background.

A subset of these five body Λ_b^0 decays could proceed through three body resonances. One specific decay is $\Lambda_b^0 \to pK^-\phi$ where $\phi \to \pi^+\pi^-\pi^0$ and a photon from the π^0 is not reconstructed. Figure 4.33 shows a comparison of the reconstructed M($\pi^+\pi^-\gamma$) for simulated $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) signal MC and $\Lambda_b^0 \to pK^-\phi$ background with the full selection applied; this shows the η' mass window would not remove very much of this background. This decay is an "unobserved" decay so there is no branching fraction measurement but it has been seen as mis-ID background

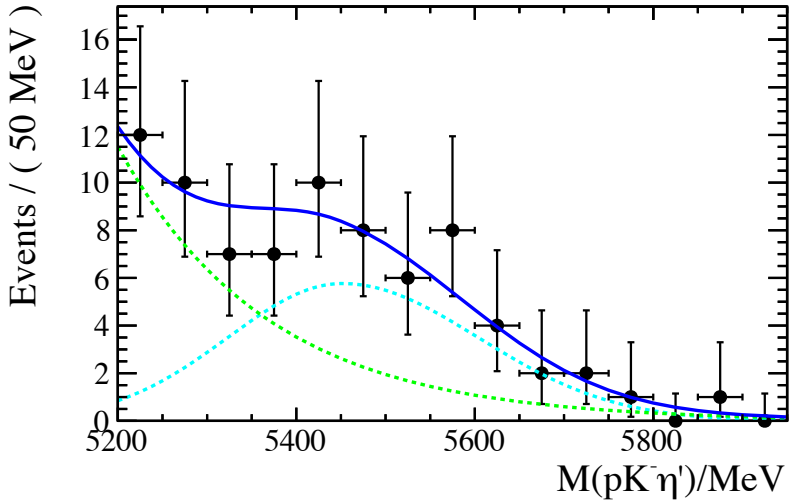


Figure 4.32: Fit performed to data from the η' sidebands. The green dashed line is combinatorial background and the cyan dashed line is partially reconstructed 5 body Λ_b^0 background.

in measurements of CP violation in $B_s^0 \to \phi\phi$ decays, where both ϕ particles are reconstructed through the decay $\phi \to K^+K^-$ but one K^+ is a mis-ID proton [122]. Despite the use of particle ID cuts in both the stripping and MVA selection there are 103 $\Lambda_b^0 \to pK^-\phi$ events observed by this analysis; this suggests a high branching fraction for a charmless B decay. Therefore, for the purposes of a predicted yield calculation the branching fraction is assumed to be 1×10^{-5}; this will obviously not give an accurate predicted yield but the only purpose of the calculation is to determine whether there will be a significant number of events passing the selection. The selection efficiency is calculated by applying the selection to a sample of fully simulated $\Lambda_b^0 \to pK^-\phi$ ($\phi \to \pi^+\pi^-\pi^0$) events, but it should be noted that phase space Monte Carlo is used and this decay could proceed through a variety of Λ resonances. Using Equation (4.25), and taking into account $\mathcal{B}$ ($\phi \to \pi^+\pi^-\pi^0$) = $(15.32 \pm 0.32)\%$ [1], the predicted yield based on this $\mathcal{B}$ assumption is 19 events in the mass fit region and 5 events in the signal region; background from this decay will need to be considered as a source of systematic uncertainty.

Another PR five body Λ_b^0 decay that could be present in the $\eta' \to \pi^+\pi^-\gamma$ rare

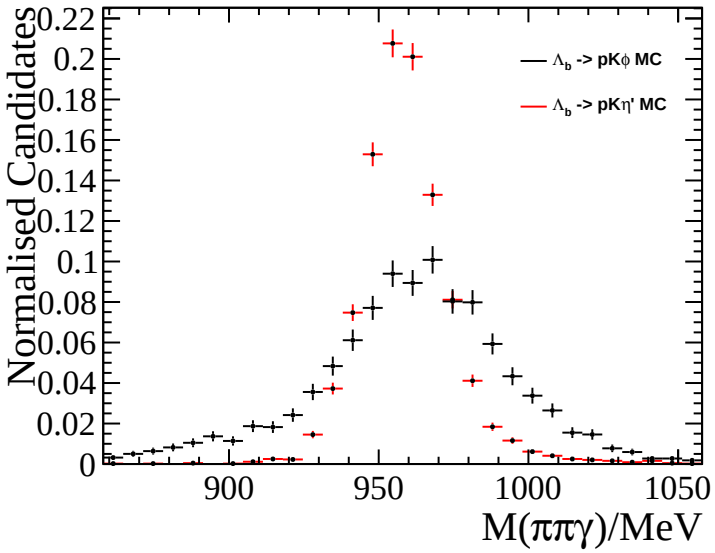


Figure 4.33: A comparison of the reconstructed M($\pi^+\pi^-\gamma$) mass for $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) MC and $\Lambda_b^0 \to pK^-\phi$ ($\phi \to \pi^+\pi^-\pi^0$) MC

channel is from the decay $\Lambda_b^0 \to pD^0\pi^-$ ($D^0 \to K^-\pi^+\pi^0$) where one of the daughter photons from the π^0 is not reconstructed. Using a sample of fully simulated $\Lambda_b^0 \to pD^0\pi^-$ ($D^0 \to K^-\pi^+\pi^0$) events and the combined branching fraction for this decay (8.96×10^{-5}) it is predicted that 7.2 ± 1.8 events are present in the full fit region but only 0.87 ± 0.62 in the signal region. Therefore, this background is considered to be negligible.

In the $\eta' \to \pi^+\pi^-\eta$ channel PRBs are considered to be less of an issue because two photons are required to have an invariant mass within 50 MeV of the η mass, which is a long way from the π^0 mass. No evidence for PRBs has been seen in this channel after the full selection.

CHAPTER 5

Results of the Search for $\Lambda_b^0 \to pK\eta'$

This chapter describes the results of the search for the decay $\Lambda_b^0 \to pK\eta'$. Firstly an overview of the fit strategy used to extract all signal yields and the branching fraction of $\Lambda_b^0 \to pK\eta'$ is described in Section 5.1. The signal yield results are then described in Section 5.2, along with the associated systematic uncertainties and the statistical significance of the results in Sections 5.3 and 5.4. This is followed by the results of the efficiency corrections in Section 5.5, and the systematic uncertainties on the branching fraction in Section 5.6. The results of the measurement of $\mathcal{B}\left(\Lambda_b^0 \to pK\eta'\right)$ are presented Section 5.7. Finally, in Section 5.8 the two body invariant mass distributions and Dalitz plots are investigated in order to understand any resonant structure.

5.1 Fit Strategy

The general fit strategy is to perform an extended unbinned maximum likelihood fit to the reconstructed invariant masses of the Λ_b^0/B^+ candidates in both rare channels and the control channel simultaneously. Both years of data will be merged because MC shows consistent signal shapes between years of data taking. In the control channel a 2D fit will be performed to the M(K^+ η') and M($\pi^+\pi^-\gamma$) variables, whereas in the rare channels the fit will only be performed to M(pK^- η'). In all cases the Λ_b^0/B^+ invariant mass is calculated using DTF with the η' mass fixed to the known value [1]. The η' mass is not subject to a fit in the rare channels because, as shown in Figure 5.1, the majority of the background involves a real η'. Therefore, fitting to the η' mass spectra as well would add extra complication for little gain in discrimination power.

In general for each channel the PDF used is given by

$$\mathcal{P} = \frac{n_s}{n_s + n_b}\mathcal{F}_s(x;\vec{\theta}) + \frac{n_b}{n_s + n_b}\mathcal{F}_b(x;\vec{\theta}), \tag{5.1}$$

where n_s and n_b are the expected numbers of signal and background events respectively; x is the variable to which the fit is performed; $\mathcal{F}(x;\vec{\theta})$ is the fit function used to describe the distribution of either signal or background events and $\vec{\theta}$ are the parameters of the fit model. In the case of the control channel, where the fit is performed to both the $M(K^+\eta')$ and $M(\pi^+\pi^-\gamma)$ variables, the overall PDF is simply the product of the individual PDFs for each invariant mass distribution. The product of PDFs is used because calculating $M(K^+\eta')$ using DTF, means there is negligible correlation between the $M(K^+\eta')$ variable and the $M(\pi^+\pi^-\gamma)$ variable which is not calculated using DTF.

The extended likelihood function for each channel is then given by

$$\mathcal{L} = e^{-(n_s+n_b)} \frac{(n_s+n_b)^N}{N!} \prod_i^N \mathcal{P}(x_i;\vec{\theta}), \tag{5.2}$$

where N is the total observed number of candidates. The term in front of the product accounts for the Poisson fluctuations on the observed number of events N. In the absence of constant terms, the Negative Log Likelihood (NLL) for each channel is then given by,

$$-\ln\mathcal{L} = n_s + n_b - N\ln(n_s+n_b) - \sum_i^N \ln\Big(\mathcal{P}(x_i;\vec{\theta})\Big). \tag{5.3}$$

The overall simultaneous likelihood is then the product of the likelihood functions for each channel (sum of log likelihoods). The advantage of using a simultaneous fit is two fold. Firstly, it allows the use of the control channel to constrain parameters in the rare channel. Secondly, it allows the ratio of branching fractions, $\mathcal{B}(\Lambda_b^0 \to pK\eta')/\mathcal{B}(B^+ \to K^+\eta')$, to be extracted directly from the fit.

The full fitting procedure used to extract the signal yields, perform the efficiency corrections and measure the ratio of branching fractions is as follows:

- A first fit is performed to extract the rare channel signal yields. This will be refered to as the "Yield Fit".
- All sources of systematic uncertainty on the signal yields are considered and the statistical significance of the observed signal yields is calculated, along with the combined significance.
- In the event that the statistical significance of the signal in either rare channel is $> 3.0\sigma$, *sWeights* will be calculated using the "Yield Fit" in order to perform corrections for the efficiency variation across the phase space of the decay (see Section 4.3). A consequence of the necessity to extract *sWeights* is that *all*

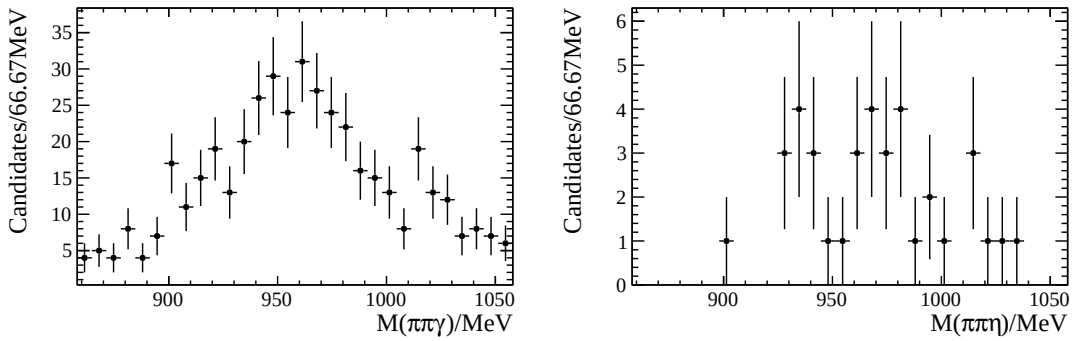


Figure 5.1: Left (Right): The η' mass spectra in the Λ_b^0 sidebands for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

yields in the fit need to be free and unconstrained [111]. This places limitations on the complexity of fit possible whilst still obtaining a stable fit. In any channel where the observed signal yield has a statistical significance $< 3.0\sigma$ the phase space integrated efficiency will be used and a systematic uncertainty will be assigned for the phase space variation.

- The calculated efficiencies will then be used to modify the "Yield Fit" such that the ratio of branching fractions is a floating parameter in the fit instead of the control channel yield; the three yield parameters N_γ, N_η and N_C which were present in the "Yield Fit" are replaced with N_γ, N_η and R, where R is the ratio of branching fractions given in Equation (4.1). This second fit is known as the "Ratio Fit". The modification is performed by parameterising N_C as a function of R, as shown in Equation (5.4),

$$N_c = \frac{1}{R}\left(\frac{N_\gamma \epsilon_c}{\epsilon_\gamma}\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_\gamma + \frac{N_\eta \epsilon_c}{\epsilon_\eta}\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_\eta\right) \times \frac{\mathcal{B}_\gamma}{\mathcal{B}_\gamma + \mathcal{B}_\eta}. \tag{5.4}$$

- In the event that the combined signal significance of both channels is $< 3\sigma$, the profile likelihood ratio for the ratio of branching fractions will be integrated in the physical region to place an upper limit on $\mathcal{B}$ ($\Lambda_b^0 \to pK\eta'$) at the 90% confidence level.

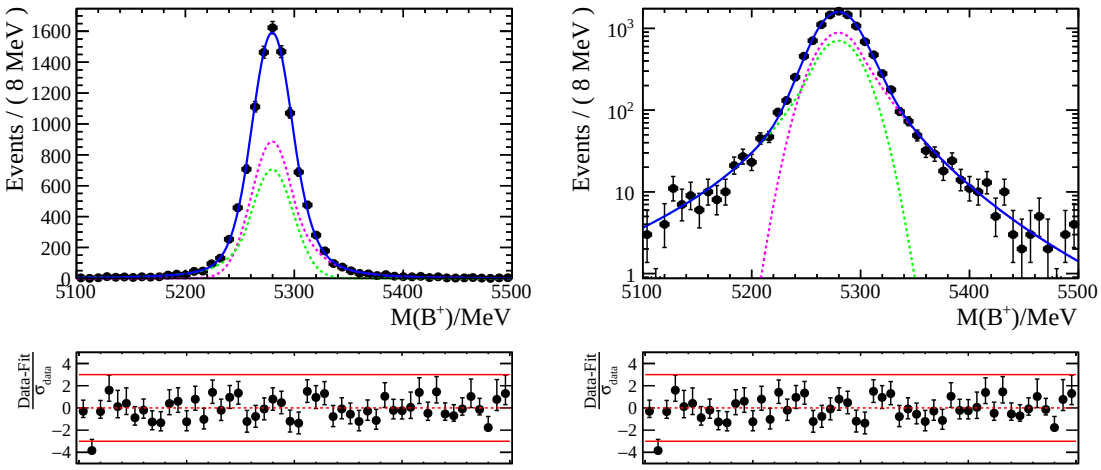


Figure 5.2: Left (Right): Fit to the B^+ mass spectrum for control channel MC using a DCB function plotted with a linear (log) scale. The green and magenta dashed lines show the individual CB functions. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$.

5.1.1 Signal Parameterisation

The Λ_b^0, B^+ and η' signal shapes are parameterised as the sum of two CB functions (Double Crystal Ball (DCB) function); the definition of a CB function is given in Equation (4.6). All of the shape parameters of the DCB functions are determined by performing fits to MC. In the control channel the fit to MC, which is shown in Figures 5.2 and Figures 5.3, is performed in 2D for the M($\pi^+\pi^-\gamma$) and M($K^+\,\eta'$) variables. This is done to match the fit to data. The fits to $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$ and $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ MC are shown in Figures 5.4 and 5.5. The numerical results of these fits are shown in Table 5.1.

When the fit is performed to data the tail parameters $(n_R, n_L, \alpha_R, \alpha_L)$ and the fit fractions f, are all fixed to the values shown in Table 5.1. In the control channel the width, σ, of the B^+ mass distribution is wider in data than MC. There are enough events in the control channel to leave σ as a free parameter but in the rare channels, where the same discrepancy is likely to be present, it is unlikely (even in the most optimistic scenario) that there will be enough signal events to float this width.

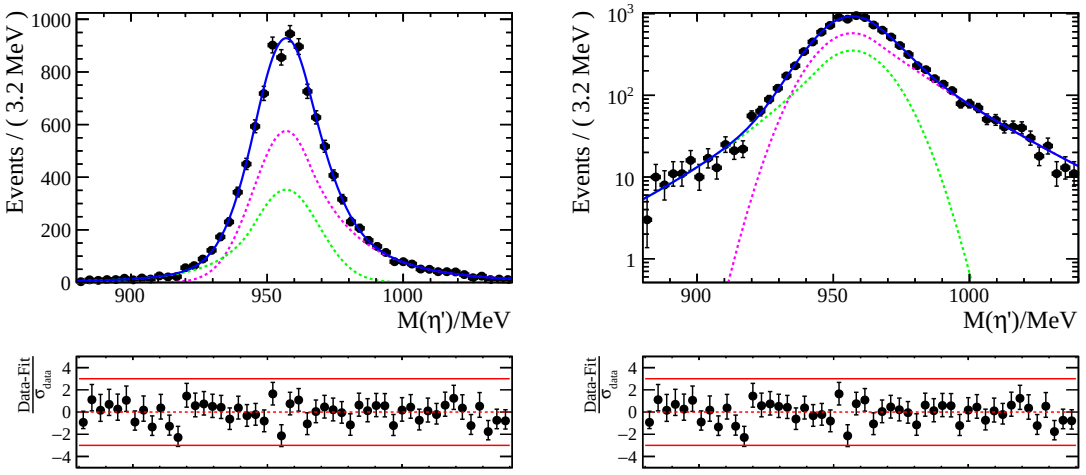


Figure 5.3: Left (Right): Fit to the η' mass spectrum for control channel MC using a DCB function plotted with a linear (log) scale. The green and magenta dashed lines show the individual CB functions. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$.

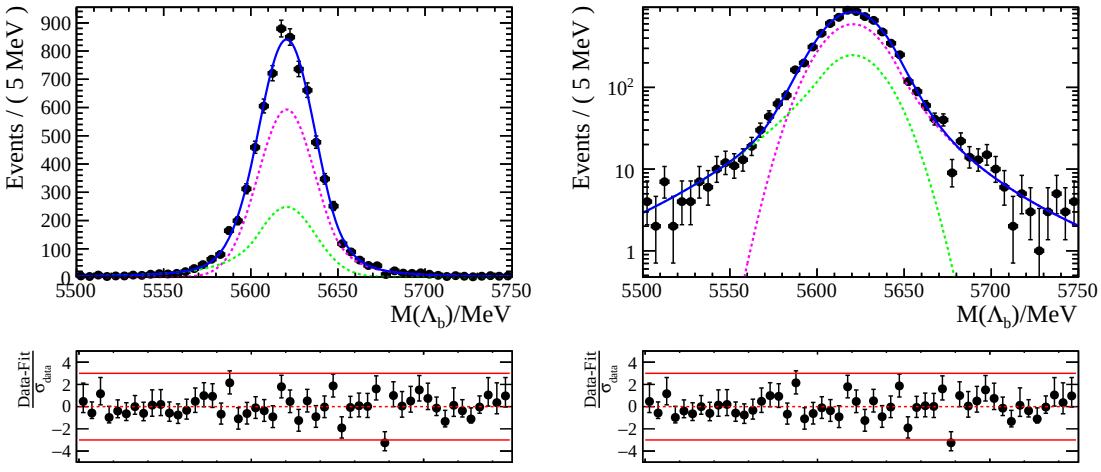


Figure 5.4: Left (Right): Fit to the Λ_b^0 mass spectrum for $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) MC using a DCB function plotted with a linear (log) scale. The green and magenta dashed lines show the individual CB functions. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$.

Table 5.1: Results of fits to MC. $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$ refer to the two rare channels whereas "Control B^+" and "Control η'" refer to the B^+ and η' signal shapes for the control channel. $(n_R, n_L, \alpha_R, \alpha_L)$ are the tail parameters of the CB function defined in Equation (4.6), where the subscript $L(R)$ refers to the CB function with the lower(upper) side tail. The fit fraction f defines the relative contribution of each CB function to the DCB function. The uncertainties are statistical uncertainties obtained from the likelihood fit.

Parameter	Value
$\mu(\eta' \to \pi^+\pi^-\eta\ )$	$5618.12 \pm 0.45\,\text{MeV}$
$\mu(\eta' \to \pi^+\pi^-\gamma\ )$	$5620.58 \pm 0.26\,\text{MeV}$
$\mu(\text{Control } B^+)$	$5279.72 \pm 0.28\,\text{MeV}$
$\mu(\text{Control } \eta')$	$956.99 \pm 0.26\,\text{MeV}$
$\sigma(\eta' \to \pi^+\pi^-\eta\ )$	$26.19 \pm 0.60\,\text{MeV}$
$\sigma(\eta' \to \pi^+\pi^-\gamma\ )$	$16.49 \pm 0.27\,\text{MeV}$
$\sigma(\text{Control } B^+)$	$19.17 \pm 0.32\,\text{MeV}$
$\sigma(\text{Control } \eta')$	$12.15 \pm 0.30\,\text{MeV}$
$\alpha_L(\eta' \to \pi^+\pi^-\eta\ )$	-1.08 ± 0.14
$\alpha_L(\eta' \to \pi^+\pi^-\gamma\ )$	-1.58 ± 0.14
$\alpha_L(\text{Control } B^+)$	-1.05 ± 0.11
$\alpha_L(\text{Control } \eta')$	-0.68 ± 0.12
$\alpha_R(\eta' \to \pi^+\pi^-\eta\ )$	0.99 ± 0.16
$\alpha_R(\eta' \to \pi^+\pi^-\gamma\ )$	1.05 ± 0.26
$\alpha_R(\text{Control } B^+)$	1.23 ± 0.16
$\alpha_R(\text{Control } \eta')$	0.97 ± 0.19
$n_L(\eta' \to \pi^+\pi^-\eta\ )$	10.78 ± 5.34
$n_L(\eta' \to \pi^+\pi^-\gamma\ )$	3.09 ± 0.50
$n_L(\text{Control } B^+)$	5.16 ± 0.75
$n_L(\text{Control } \eta')$	10.08 ± 6.83
$n_R(\eta' \to \pi^+\pi^-\eta\ )$	6.54 ± 2.85
$n_R(\eta' \to \pi^+\pi^-\gamma\ )$	4.01 ± 1.18
$n_R(\text{Control } B^+)$	3.14 ± 0.45
$n_R(\text{Control } \eta')$	5.82 ± 2.31
$f(\eta' \to \pi^+\pi^-\eta\ )$	0.37 ± 0.10
$f(\eta' \to \pi^+\pi^-\gamma\ )$	0.31 ± 0.12
$f(\text{Control } B^+)$	0.44 ± 0.09
$f(\text{Control } \eta')$	0.36 ± 0.09

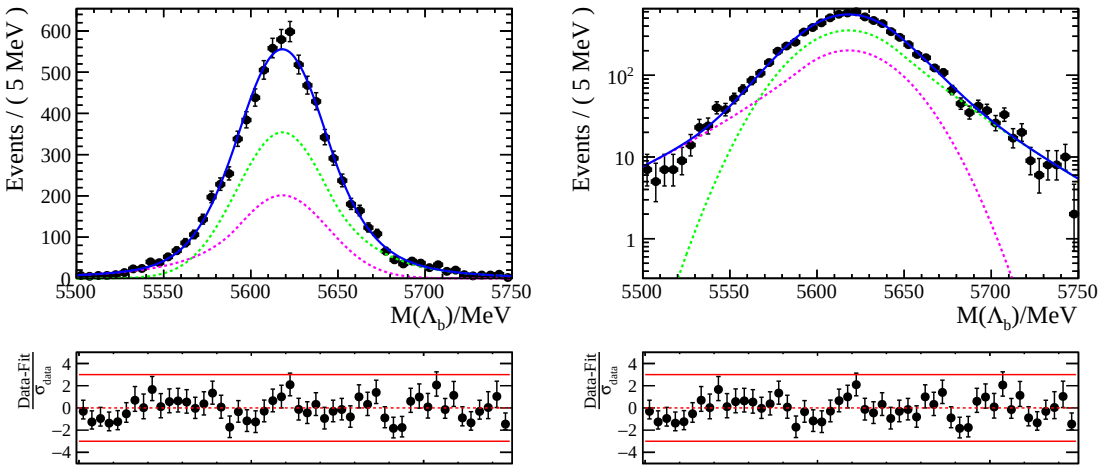


Figure 5.5: Left (Right): Fit to the Λ_b^0 mass spectrum for $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta)$ MC using a DCB function plotted with a linear (log) scale. The green and magenta dashed lines show the individual CB functions. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$

Therefore, a data/MC correction factor σ_{scale} is introduced such that the fitted width in data is the product $\sigma = \sigma_{MC} \times \sigma_{scale}$ where σ_{MC} is the value obtained in the fit to MC. σ_{scale} is then left free to vary but shared between all channels; the high statistics in the control channel is used to constrain the rare channels. To further improve the stability of the fit, the mass difference between the B^+ and Λ_b^0 is also fixed to the latest LHCb measurement such that $\delta_m = \mu(\Lambda_b^0) - \mu(B^+) = 339.81 \pm 0.72\,\text{MeV}$ [123].

5.1.2 Background Modelling

In the control channel only combinatorial background is present, which is modelled with a second order Chebychev polynomial in both M($\pi^+\pi^-\gamma$) and M($K^+\,\eta'$) variables. The definition of a Chebychev polynomial is given in Equations. (4.7) and (4.8). The two shape parameters are free to vary.

In the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma)$ channel, as indicated in Table 4.22, there are two background components to be considered in the nominal fit model: combinatorial and partially reconstructed background from $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ decays. The shape

Parameter	Value
μ	$5452 \pm 24\,\text{MeV}$
σ_L	$128.5 \pm 21.1\,\text{MeV}$
σ_R	$152.2 \pm 15.3\,\text{MeV}$

Table 5.2: The parameters of the bifurcated Gaussian used to model the $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ background shape.

used for $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ decays is a bifurcated Gaussian with the parameters fixed to the results of the fit to MC shown in Figure 4.31 and Table 5.2. Combinatorial background is modelled with a falling exponential function where the shape parameter is free to vary.

In the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) rare channel there are only $\mathcal{O}(100)$ events passing the full selection. As mentioned in Section 4.4 there is no evidence for any background other than combinatorial background, which is modelled with a falling exponential function.

5.1.3 Fit Model Summary

Overall the fit has 45 parameters, of which 27 are fixed to MC and the mean of the two Λ_b^0 signal functions are constrained by the control channel. Therefore, the fit has 17 free parameters. A summary of the models used and the free parameters is given in Table 5.3.

5.1.4 Fit Validation

Pseudoexperiments are used to test the stability of the simultaneous fit over a range of different $\mathcal{B}$ ($\Lambda_b^0 \to pK\eta'$) assumptions between 1×10^{-7} and 1×10^{-5}. This large range of assumptions is tested because there is no prediction for $\mathcal{B}$ ($\Lambda_b^0 \to pK\eta'$), therefore the stability of the fit needs to be tested for all plausible scenarios. For

Component	Model	Free parameters
$\underline{B^+ \to K^+\eta'}$ **Channel:**		
Signal	Sum of two CBs	$\mu(B^+), \mu(\eta')$
		$\sigma(\eta')$
		$N_{Signal}(Control)$
Combinatorial Background	Chebychev polynomials	$P1(B^+), P1(\eta')$
		$P2(B^+), P2(\eta')$
		$N_{comb.back}(Control)$
$\underline{\eta' \to \pi^+\pi^-\gamma}$ **Rare Channel:**		
Signal	Sum of two CBs	$N_{Signal}(\eta' \to \pi^+\pi^-\gamma)$
PR Background $\Lambda_b^0 \to 4h + \pi^0$	Bifurcated Gaussian	$N_{PRback}(\eta' \to \pi^+\pi^-\gamma)$
Combinatorial Background	Exponential	$K(\eta' \to \pi^+\pi^-\gamma)$
		$N_{comb.back}(\eta' \to \pi^+\pi^-\gamma)$
$\underline{\eta' \to \pi^+\pi^-\eta}$ **Rare Channel:**		
Signal	Sum of two CBs	$N_{Signal}(\eta' \to \pi^+\pi^-\eta)$
Combinatorial Background	Exponential	$K(\eta' \to \pi^+\pi^-\eta)$
		$N_{comb.back}(\eta' \to \pi^+\pi^-\eta)$
Shared Parameters	Data/MC width correction	σ_{scale}

Table 5.3: A summary of the simultaneous maximum likelihood fit used to extract signal yields. The N parameters are event yields; *Signal* is the B^+/Λ_b^0 signal yield, *PRback* is the PRB yield and *comb.back* is the combinatorial background yield. μ is the mean of the CB functions, σ is the width of the CB functions, $P1$ and $P2$ are the two parameters describing the shape of the second order Chebychev polynomials and K is the constant of the falling exponential functions.

$\Lambda_b^0 \to pK\eta'$ $\mathcal{B}$ Assumption	$\eta' \to \pi^+\pi^-\gamma$ Generated Yield	$\eta' \to \pi^+\pi^-\eta$ Generated Yield
1×10^{-7}	1	0
2×10^{-7}	2	0
3×10^{-7}	3	1
4×10^{-7}	4	1
5×10^{-7}	5	2
6×10^{-7}	6	2
7×10^{-7}	7	2
8×10^{-7}	8	3
9×10^{-7}	9	3
1×10^{-6}	10	4
2×10^{-6}	20	8
3×10^{-6}	30	12
4×10^{-6}	40	16
5×10^{-6}	50	20
6×10^{-6}	61	24
7×10^{-6}	71	29
8×10^{-6}	81	33
9×10^{-6}	91	37
1×10^{-5}	101	41

Table 5.4: Generated yield in each of the rare channels for the $\mathcal{B}$ assumptions tested with pseduoexperiments. These are predicted yields based on phase space integrated efficiencies.

each of the BF assumptions tested the phase space integrated efficiencies detailed in Section 4.3.6 are used to calculate a predicted yield in each of the rare channels. These predicted yields can be found in Table 5.4. For each $\mathcal{B}$ assumption 1000 pseudoexperiments are generated from the nominal fit model with these rare channel yields. For the generation of the pseudoexperiments all other free parameters of the fit model are fixed to the fit results from data (see Table 5.6)[1]. Each of the generated pseudoexperiments is then subject to the nominal mass fit with all 17 free parameters unconstrained. It is found that 100% of the fits converge and have fully accurate error matrices.

The linearity of the fit is tested by plotting the mean fit result for each of the rare channel yields against the generated yield (see Figure 5.6) and a linear minimum χ^2 fit is applied, the results of which are shown in Table 5.5. These plots and results

[1]The signal yields were still blind when the pseudoexperiments were performed

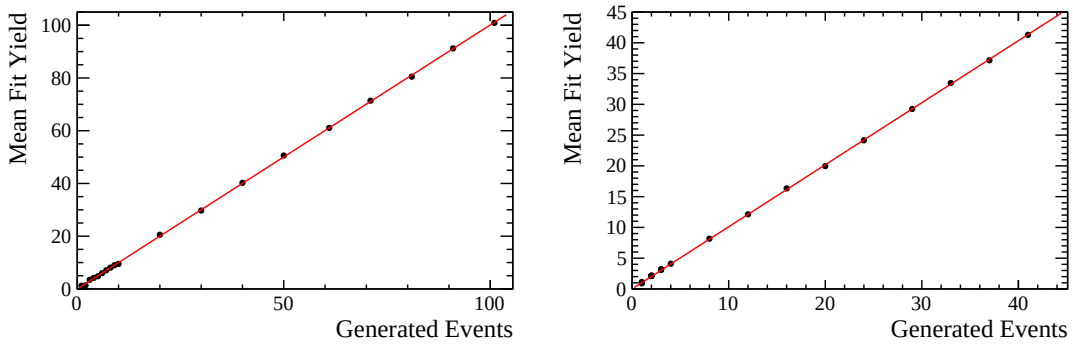


Figure 5.6: Left(Right): Fit linearity for each of the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) yields.

Parameter	$\eta' \to \pi^+\pi^-\gamma$ Result	$\eta' \to \pi^+\pi^-\eta$ Result
χ^2/ndf	1.42	1.35
Gradient	1.001 ± 0.003	1.008 ± 0.003
Intercept	0.006 ± 0.086	0.02 ± 0.04

Table 5.5: Results of the linear minimum χ^2 fit applied to the linearity plots shown in Figure 5.6

show excellent linearity for a wide range of possible signal yields.

The pulls of the fit are also studied for each of the rare channel yields; Figure 5.7 shows the mean and RMS of the pull distribution for each BF assumption. There is clearly some bias at very low signal yields (<4 events) in the $\eta' \to \pi^+\pi^-\eta$ channel; this is because there are also very few background events in this channel. With low statistics a maximum likelihood fit is intrinsically biased; in the event of a signal yield < 4 events any bias will be corrected for.

5.2 Signal Yield Results

The projection of the simultaneous fit in the control channel, which has always been unblind, is shown in Figure 5.8. The signal yield in the control channel is 11848 ± 131 events. To further investigate the quality of this fit, projections in the B^+ and η' signal windows are shown in Figure 5.9; both the signal window projections and the

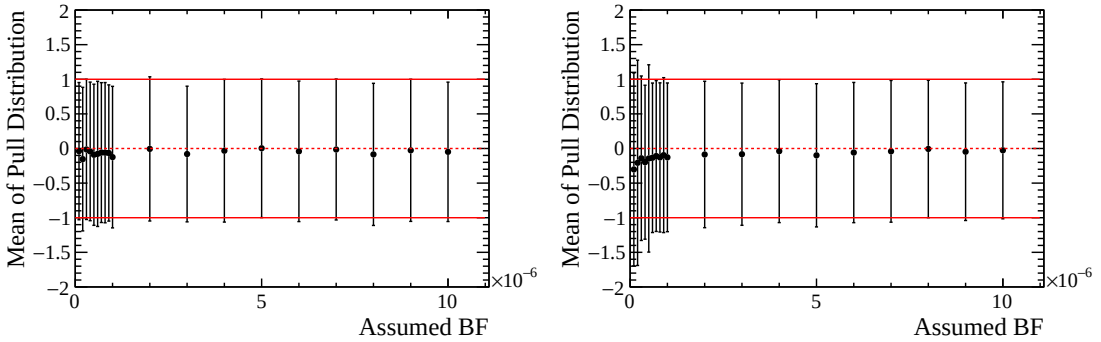


Figure 5.7: Left(Right): Representation of the pull distributions at each $\mathcal{B}$ assumption for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel. The marker is the mean of the pull distribution and the error bars represent the RMS of the pull distribution.

full fit projections show a high quality of fit. The signal windows are defined as within $\pm 3\sigma$ of the B^+/η' mean μ. In the case of the η' mass distribution σ is a free parameter of the fit; the value of which is given in Table 5.6. In the case of the B^+ mass distribution σ is the product $\sigma_{scale} \times \sigma_{MC}$, where σ_{scale} is also given in Table 5.6 and σ_{MC} is the result from the fit to MC which is given in Table 5.1. The projection of the B^+ (η') mass distribution and fit in the η' (B^+) signal window involves plotting only data which falls in the corresponding mass window and projecting the fit result in the same window. Within the signal window, the signal to background ratio is higher than the full fit region; this is a further test of the fit quality.

After the selection was frozen, the efficiency maps had been created and the fit was validated, the rare channel mass fits were unblinded. The rare channel projections of the simultaneous mass fit are shown in Figures 5.10 and 5.11. These projections show clear and abundant signal for the decay $\Lambda_b^0 \to pK\eta'$ in both η' decay channels; the signal yield in the $\eta' \to \pi^+\pi^-\gamma$ channel is 117 ± 15 events and the signal yield in the $\eta' \to \pi^+\pi^-\eta$ channel is 45 ± 8 events. The results for all free parameters of the simultaneous fit are shown in Table 5.6.

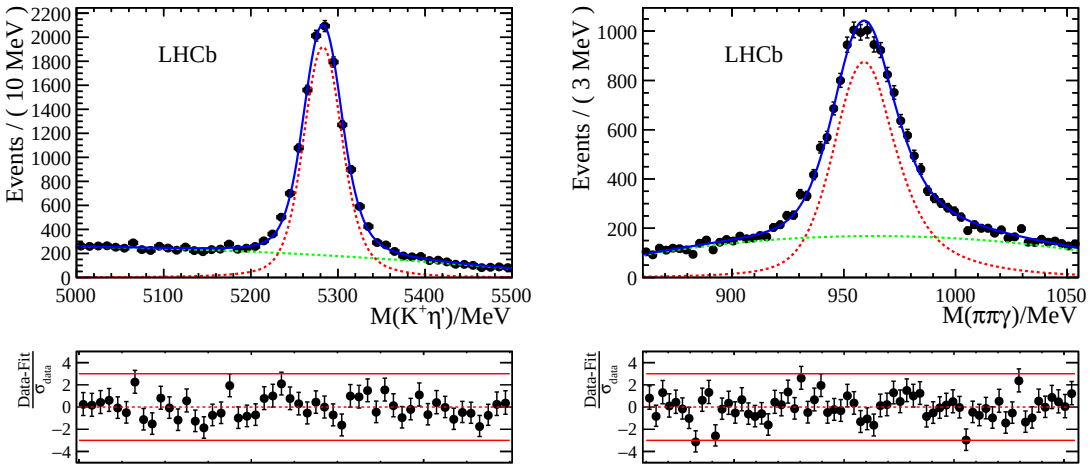


Figure 5.8: Left(Right): Results of the simultaneous fit to data in the control channel for the M(K^+ η') (M($\pi^+\pi^-\gamma$)) observable. The red dashed line represents the signal PDF and the green dashed line is combinatorial background. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$

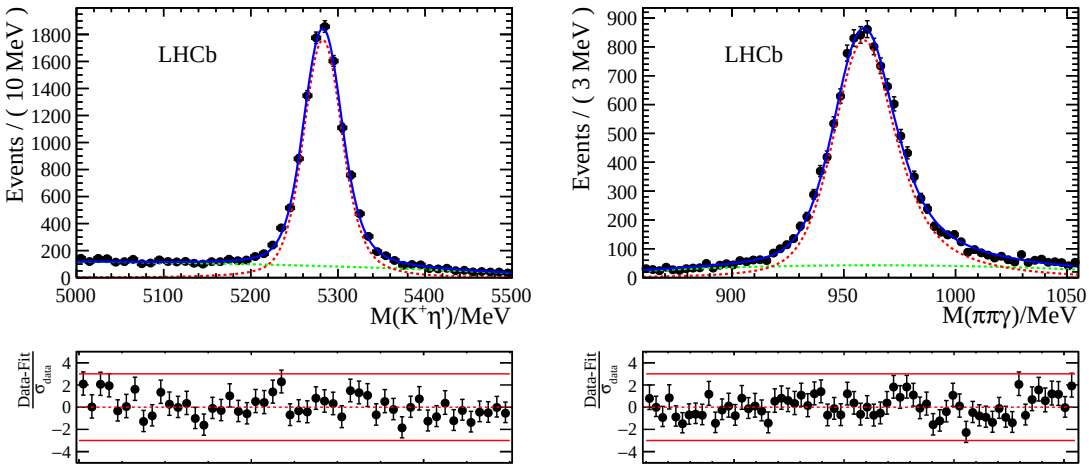


Figure 5.9: Left(Right): Projections of the control channel PDF in the η' (B^+) signal window. The red dashed line represents the signal PDF and the green dashed line is combinatorial background. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$

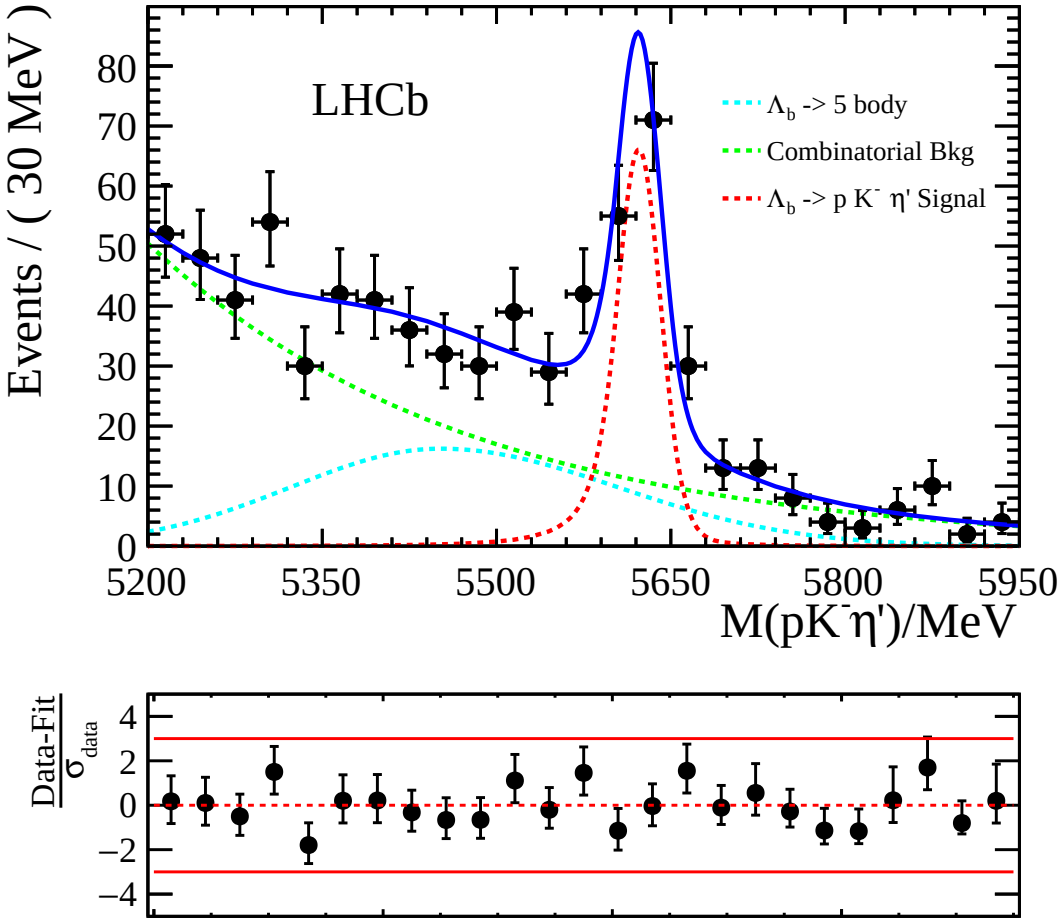


Figure 5.10: The fit performed to the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) channel data. The green line is combinatorial background, the cyan line is partially reconstructed background from $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ decays and the red line is signal. The overall fit function is shown by the solid blue line. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$

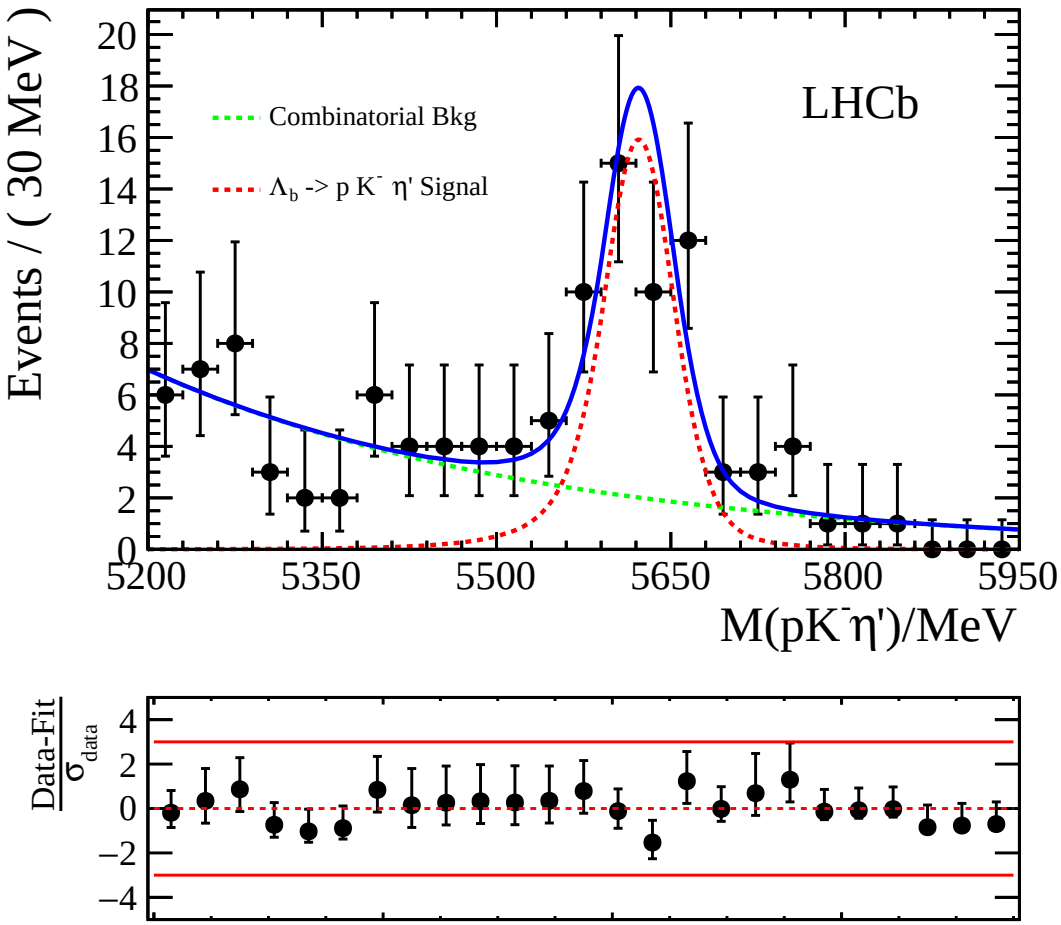


Figure 5.11: The fit performed to the $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) channel data. The green line is combinatorial background and the red line is signal. The overall fit function is shown by the solid blue line. The lower panel is the pull distribution $\frac{\text{Data}-\text{Fit}}{\sigma_{\text{data}}}$

Parameter	Value
$N_{\text{signal}}(B^+ \to K^+\eta'\ )$	11848 ± 131
$N_{\text{signal}}(\eta' \to \pi^+\pi^-\gamma\ )$	117 ± 15
$N_{\text{signal}}(\eta' \to \pi^+\pi^-\eta\ )$	45 ± 8
$N_{\text{PRBack.}}(\eta' \to \pi^+\pi^-\gamma\ )$	186 ± 40
$N_{\text{Comb.Back.}}(\eta' \to \pi^+\pi^-\gamma\ )$	432 ± 41
$N_{\text{Comb.Back.}}(\eta' \to \pi^+\pi^-\eta\ )$	70 ± 10
$N_{\text{Comb.Back.}}(\text{Control})$	9605 ± 123
$K(\eta' \to \pi^+\pi^-\eta)$	-0.00294 ± 0.00070
$K(\eta' \to \pi^+\pi^-\gamma)$	-0.00363 ± 0.00032
$P1(\text{Control } B^+)$	-0.645 ± 0.016
$P2(\text{Control } B^+)$	-0.121 ± 0.019
$P1(\text{Control } \eta')$	0.06 ± 0.02
$P2(\text{Control } \eta')$	-0.271 ± 0.021
$\mu(\text{Control } \eta')$	959.06 ± 0.18
$\mu(\text{Control } B^+)$	5282.73 ± 0.27
$\sigma(\text{Control } \eta')$	13.44 ± 0.17
σ_{scale}	1.152 ± 0.013

Table 5.6: The final results of the simultaneous fit to data. $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$ refer to the two rare channels whereas Control B^+ and Control η' refer to the B^+ and η' mass distributions for the control channel. K is the constant of the exponential used to model combinatorial background and $P1/P2$ are the coefficients of the Chebychev polynomials used to model combinatorial background in the control channel.

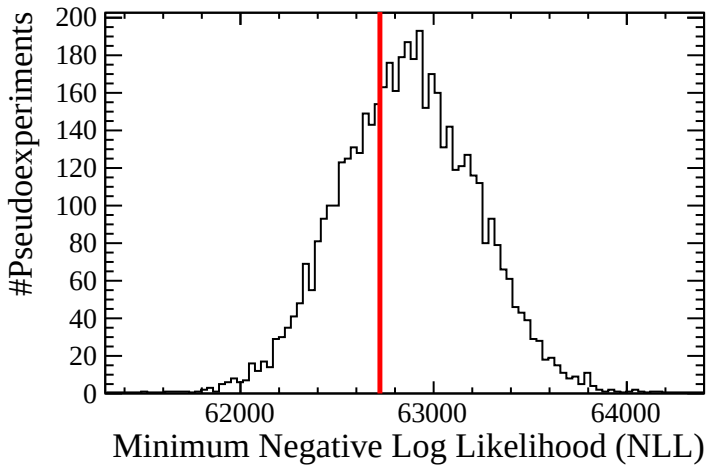


Figure 5.12: The distribution of minimum NLL values obtained when the pseudoexperiments described in Section 5.1.4 are run with the yields observed in data. The vertical red line shows the minimum NLL value obtained from the fit to data.

5.2.1 Goodness Of Fit

In order to assess the quality of the fit, the pseudoexperiments described in Section 5.1.4 are performed with the signal yields observed in data. The value of the NLL observed from the fit to data is then compared to the distribution of the NLL values obtained from the pseudoexperiments; this comparison can be seen in Figure 5.12. The value obtained from data is comfortably within the distribution of expected values obtained from pseudoexperiments, which confirms a good quality of fit.

5.2.2 Checks for Non-Resonant Background

Further checks are performed to ensure that the observed signals proceed through an η' resonance as expected; the non-resonant decays $\Lambda_b^0 \to pK^-\pi^+\pi^-\gamma$ and $\Lambda_b^0 \to pK^-\pi^+\pi^-\eta$ are allowed but expected to be suppressed by both the physics and selection. Nevertheless a fit is performed to the background-subtracted[2] $M(\pi^+\pi^-\gamma)$ and

[2]The sPlot method is used to perform the background subtraction

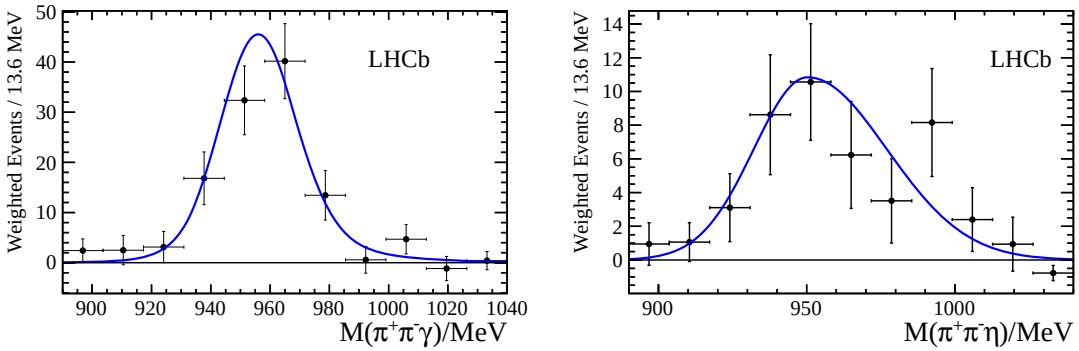


Figure 5.13: Left (Right): Background subtracted η' mass distributions in the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) decay channel. The solid blue line is the total fit function.

$M(\pi^+\pi^-\eta)$ distributions, which includes a signal component taken from MC and a linear background shape to account for the possibility of non-resonant background from the aforementioned decays. In both channels a first order Chebychev polynomial is used to model the non-resonant component, in which the gradient is free to vary. In the $\eta' \to \pi^+\pi^-\gamma$ channel a DCB function is used for the signal component, with all parameters except the signal yield and mean fixed to the values obtained in a fit to MC. In the $\eta' \to \pi^+\pi^-\eta$ channel, the signal shape is affected by the $\pm 50\,\text{MeV}$ cut on the η mass used in stripping. Therefore, the $M(\pi^+\pi^-\eta)$ shape is modelled with a bifurcated Gaussian function, where the yield and mean are free to vary but all other parameters are fixed to values obtained from a fit to MC. The results of these fits are shown in Figure 5.13. In both channels the yield of the non-resonant background component is consistent with zero and the yield of the signal component is consistent with that extracted from the nominal mass fit. This confirms that there is no contamination from non-resonant background. Furthermore, the fit result for the mean of the DCB in the $\eta' \to \pi^+\pi^-\gamma$ channel is $956 \pm 3\,\text{MeV}$ which is consistent with the world average η' mass ($957.78 \pm 0.06\,\text{MeV}$) [1]. In the $\eta' \to \pi^+\pi^-\eta$ channel the mean of the bifurcated Gaussian is consistent with the expectation from MC.

As a further check in the $\eta' \to \pi^+\pi^-\gamma$ channel, the ρ^0 polarisation angle is studied. It has been seen that the decay $\eta' \to \pi^+\pi^-\gamma$ proceeds through a $\rho^0 \to \pi^+\pi^-$ resonance nearly 100% of the time [124]. Consequently, the decay topology of this

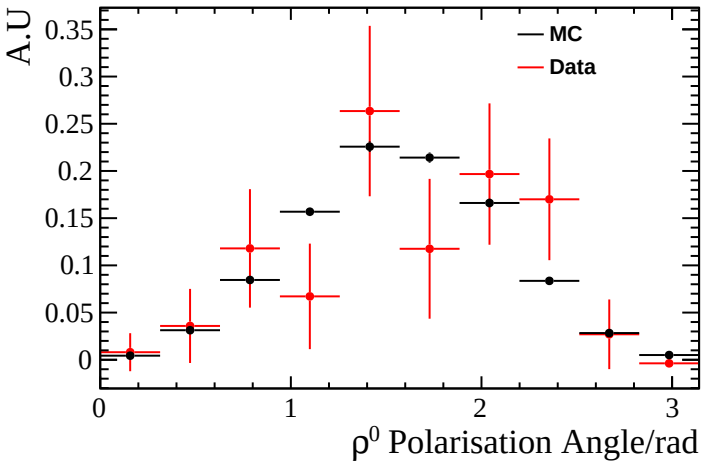


Figure 5.14: A comparison of the ρ^0 polarisation angle, in the $\eta' \to \pi^+\pi^-\gamma$ decay channel, between MC and data.

$pseudoscalar \to vector + photon$ decay is quite unique. The ρ^0 polarisation angle is defined as the angle between the π^+ meson in the ρ^0 rest frame and the Lorentz boost from the η' rest frame to the ρ^0 rest frame; a comparison of the ρ^0 polarisation angle between MC and background subtracted data can be seen in Figure 5.14. This shows good agreement between data and MC, with the simulation providing a good description of the overall shape of data. In non-resonant decays to $\pi^+\pi^-\gamma$, the yield would be independent of the ρ^0 polarisation angle which would lead to a flat distribution; this is not seen in background subtracted data. This check gives further confidence that there is no appreciable contamination from non-resonant background in the $\eta' \to \pi^+\pi^-\gamma$ decay channel.

5.2.3 Control Channel Expected Yield Comparison

In a similar manner to the comparisons performed in Section 4.2.2.1, a comparison of the expected control channel yield and observed yield after the full selection is made. The expected number of events in the control channel, for both years of data

combined, is given by

$$N_{fullsel}^{exp} = 2 \times f_u \times \mathcal{B}(B^+ \to K^+\eta') \times \mathcal{B}(\eta' \to \pi^+\pi^-\gamma) \times \left(\mathcal{L}_{12}\sigma_{12}(b\bar{b})\varepsilon_{12} + \mathcal{L}_{11}\sigma_{11}(b\bar{b})\varepsilon_{11}\right), \tag{5.5}$$

where f_u is the probability of a b-quark hadronising with a u-quark to form a B^+ meson; $\mathcal{L}_y$,$\sigma(b\bar{b})_y$ and ε_y are the luminosity, $b\bar{b}$ cross section and efficiency for a given year y of data taking respectively. As the combined efficiencies described in Section 4.3 are the weighted average of 2012 and 2011 efficiencies, where the weights are the product of luminosity and $b\bar{b}$ cross section for a given year of data taking, the expected number of events in the control channel can also be expressed as

$$N_{fullsel}^{exp} = 2 \times f_u \times \mathcal{B}(B^+ \to K^+\eta') \times \mathcal{B}(\eta' \to \pi^+\pi^-\gamma) \times \varepsilon_{comb} \times \left(\mathcal{L}_{12}\sigma_{12}(b\bar{b}) + \mathcal{L}_{11}\sigma_{11}(b\bar{b})\right), \tag{5.6}$$

where ε_{comb} is the combined 2012 and 2011 efficiency. Taking the value of ε_{comb} given in Table 4.20 and using the same values as the pre-selection comparison for all other parameters (see Table 4.5) the expected number of $B^+ \to K^+\eta'$ events after the full selection is $N_{fullsel}^{exp} = 16079 \pm 1774$. This is consistent with the number of observed events, $N_{obs} = 11848 \pm 131$, at the level of 2.5σ. However, the ratio of expected to observed events is consistent with the ratio seen in the pre-selection comparison; in the pre-selection comparison the ratio is 1.37 ± 0.15 and in this comparison the ratio is also 1.38 ± 0.15. This provides confidence that the efficiencies of the BDT, PID and $M(\pi^+\pi^-)$ cuts are accurate.

5.3 Systematic Uncertainties on Signal Yield

This section describes the sources of systematic uncertainty on the rare channel signal yields reported in Section 5.2 and their determination. The sources of systematic uncertainty are described in Sections 5.3.1- 5.3.4 and a summary of all systematic

Channel:	$\eta' \to \pi^+\pi^-\gamma$	$\eta' \to \pi^+\pi^-\eta$	Control
Fixed Fit Parameters	7.7	1.3	102
Fit Model Choice	6.1	1.2	283
$\Lambda_b^0 \to pK^-\phi$ Background	0.5	0.04	0.5
σ_{scale}	2.4	1.1	15
Total	10.2	2.0	301

Table 5.7: A summary of all the systematic uncertainties on the measured signal yields. All uncertainties are given as an absolute number of events.

uncertainties is given in Table 5.7. The total systematic uncertainty for each channel is all the individual systematics components summed in quadrature. The final result for the signal yields of the decay $\Lambda_b^0 \to pK\eta'$ are:

$$N(\eta' \to \pi^+\pi^-\gamma) = 117 \pm 15(\text{stat.}) \pm 10(\text{sys.}), \tag{5.7}$$

$$N(\eta' \to \pi^+\pi^-\eta) = 45 \pm 8(\text{stat.}) \pm 2(\text{sys.}). \tag{5.8}$$

5.3.1 Fixed Fit Model Parameters

Many parameters of the signal shapes are fixed to values determined with fits to MC. Therefore the uncertainties on these parameters from the fits to MC, the results of which can be found in Table 5.1, need to be propagated to a systematic uncertainty on the signal yield. This is done by repeating the "Yield Fit" 1000 times, each time sampling the fixed parameters from a Gaussian with a mean equal to the nominal value and width equal to the uncertainty from the fit to MC. The RMS of the distributions of rare channel signal yields will be taken as the systematic uncertainty. The correlations between the fixed parameters need to be preserved in the sampling process, otherwise the systematic uncertainty could be overestimated. This is done with the use of the covariance matrices, $\mathbf{\Omega}$, from the fits to MC. Any symmetric positive definite matrix can be written as the product of a lower triangular matrix,

$\mathbf{L}$, and its conjugate transpose,

$$\mathbf{\Omega} = \mathbf{L}\mathbf{L}^{\top}. \tag{5.9}$$

The lower diagonal matrix $\mathbf{L}$ is determined with the use of a Cholesky decomposition. A set of correlated random variables, $\vec{C}$, with variances equal to those entering the covariance matrix, can then be produced by taking the product,

$$\vec{C} = \mathbf{L}\vec{X}, \tag{5.10}$$

where $\vec{X}$ is a vector of uncorrelated unit normal random variables (sampled from a Gaussian with a mean of 0 and width of 1). Sets of correlated fixed fit parameters, $\vec{F}$, are therefore sampled by calculating

$$\vec{F} = \mathbf{L}\vec{X} + \vec{\mu}, \tag{5.11}$$

where $\vec{\mu}$ are the nominal values of the fixed fit parameters.

5.3.2 $\Lambda_b^0 \to pK^-\phi$ Background

There is potential for background from the decay $\Lambda_b^0 \to pK^-\phi$ ($\phi \to \pi^+\pi^-\pi^0$) in the $\eta' \to \pi^+\pi^-\gamma$ rare channel. However, as the shape of this background is very similar to the shape of the $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ background it is not possible to achieve a stable and unbiased fit with both components included. Therefore, as this background is in fact a subset of the $\Lambda_b^0 \to pK^-\pi^+\pi^-\pi^0$ background, it was not included in the nominal fit as a distinct component and a systematic uncertainty is assigned for its non-inclusion. The contribution associated with this source of systematic uncertainty is estimated by reperforming the fit with a component for $\Lambda_b^0 \to pK^-\phi$ included and the change in signal yield is assigned as a systematic uncertainty.

Channel:	$\eta' \to \pi^+\pi^-\gamma$	$\eta' \to \pi^+\pi^-\eta$	Control
Model I:	0.8	1.2	0.8
Model II:	6.1	0.2	283
Model III:	1.1	0.01	0.5

Table 5.8: A summary of the changes in signal yields for each of the alternative fit models considered. All values are given in units of Events.

5.3.3 Choice of Fit Model

A systematic uncertainty can occur due to the choice of fit model used; the true shape of a distribution may not exactly match the distribution with which it has been modelled. To assess the size of this systematic uncertainty the fit is repeated with alternative models:

- **Model I:** The exponential combinatorial background component in the rare channels is replaced with a second order Chebychev polynomial.
- **Model II:** The DCB signal parameterisations in all channels are replaced by Hypatia functions [125].
- **Model III:** The bifurcated Gaussian partially reconstructed background shape in the $\eta' \to \pi^+\pi^-\gamma$ rare channel is replaced with a `RooKeysPDF` kernel density estimation [126].

The difference in signal yield with respect to the nominal fit model has been determined for each model change; these differences are shown in Table 5.8. The systematic uncertainty is assigned as the largest change in the signal yield.

5.3.4 $\sigma_{\mathbf{scale}}$

Another systematic uncertainty can arise from the assumption that the σ_{scale} Data/MC width correction factor is the same in the control channel and both rare channels.

It is possible due to the presence of an extra track in both rare channels and an extra photon in the $\eta' \to \pi^+\pi^-\eta$ rare channel that this assumption is not completely correct. To assess an uncertainty due to this assumption not holding, the fit is reperformed with σ_{scale} fixed to $\frac{\sigma_\eta}{\sigma_c} \times \sigma_{scale}^{nominal}$ where $\sigma_{c,\eta}$ is the width of signal PDF in the $B^+ \to K^+\eta'$ ($\eta' \to \pi^+\pi^-\gamma$) and $\Lambda_b^0 \to pK\eta'$ ($\eta' \to \pi^+\pi^-\eta$) channels respectively. The change in the signal yields with respect to the nominal fit is assigned as a systematic uncertainty.

5.4 Signal Significance

The statistical significance of the signal yields observed is assessed using the profile likelihood ratio, which is given by

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\nu}})}{\mathcal{L}(\hat{\mu}, \hat{\nu})}, \tag{5.12}$$

where $\mathcal{L}(\hat{\mu}, \hat{\nu})$ is the value of the likelihood function at its minimum with respect to all parameters; the parameter of interest μ and the nuisance parameters ν are free to vary and take the values $(\hat{\mu}, \hat{\nu})$ which minimise the likelihood function. $\mathcal{L}(\mu, \hat{\hat{\nu}})$ is the value of the likelihood function with μ fixed and the nuisance parameters ν allowed to float, this means the nuisance parameters take new values $(\hat{\hat{\nu}})$ which minimise the likelihood function for a given value of μ.

Wilks' theorem states that $-2\ln(\lambda(\mu))$ will follow a χ^2 distribution where the number of degrees of freedom is equal to the difference in dimensionality between $\mathcal{L}(\mu, \hat{\hat{\nu}})$ and $\mathcal{L}(\hat{\mu}, \hat{\nu})$ [110]. In this case the difference in dimensionality is one, because only one signal yield is fixed at a time. A consequence of this is that the signal significance of each yield, in units of Gaussian standard deviations σ, is given by

$$\sigma = \sqrt{-2\ln\lambda(0)}. \tag{5.13}$$

Channel	Without Systematics	With Systematics
$\eta' \to \pi^+\pi^-\gamma$	10.4σ	9.7σ
$\eta' \to \pi^+\pi^-\eta$	7.3σ	7.1σ
Combined	12.7σ	12.0σ

Table 5.9: A summary of signal significances, both with and without systematic uncertainties, for each rare channel and both combined.

However, the systematic uncertainties on the signal yields need to be accounted for. This is done by convolving the profile likelihood ratio with a Gaussian with a mean of zero and width equal to the systematic uncertainty on the signal yield.

For the purposes of assessing a combined signal significance of both decay channels the null hypothesis is a total $\Lambda_b^0 \to pK\eta'$ signal yield of 0 events. Therefore, in order to assess the combined signal significance the fit is parameterised such that the total signal yield (the sum of the two rare channel yields) is a free parameter. This allows the profile likelihood ratio to be determined as a function of the total signal yield. The systematic uncertainty on the total yield is calculated as the systematic uncertainties on the individual yields summed in quadrature.

The profile likelihood ratios as a function of the individual rare channel yields and the total yield are shown in Figure 5.15. The resulting signal significances are given in Table 5.9. The combined significance after accounting for systematic uncertainties is 12.0σ, which is equivalent to a p-value of 1.78×10^{-33}. This is undoubtedly the first observation of $\Lambda_b^0 \to pK\eta'$.

5.5 Efficiency Corrections

As the signal significance is greater than 3σ in both rare decay channels the efficiencies are corrected for the variation over the phase space of the decay. This is done

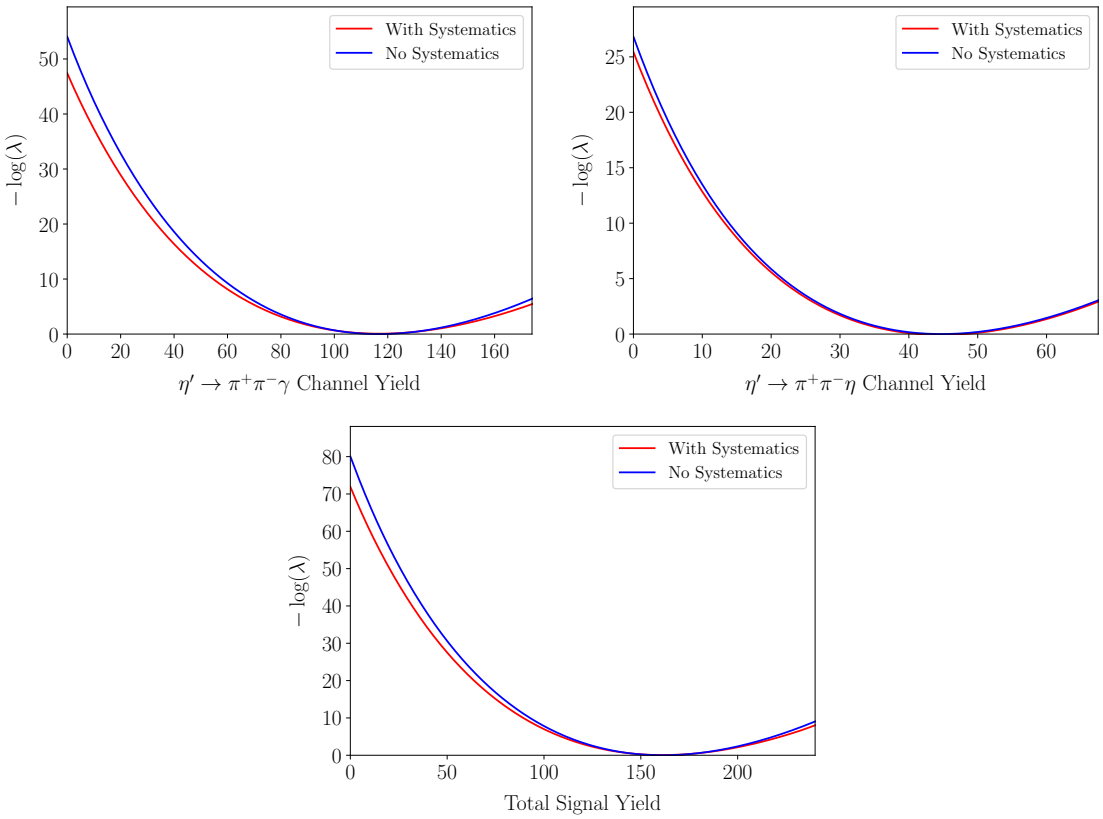


Figure 5.15: Scans of $-\ln\lambda$ for the two rare channel yields and the total signal yield both with and without systematic uncertainties included. The point at which the likelihood function crosses the yield=0 axis is equal to $\sigma^2/2$

Channel	Corrected Efficiency
$\eta' \to \pi^+\pi^-\gamma$	0.0200%
$\eta' \to \pi^+\pi^-\eta$	0.0157%

Table 5.10: The values of the phase space corrected rare channel efficiencies. The uncertainties on these values are discussed in section 5.6.1.

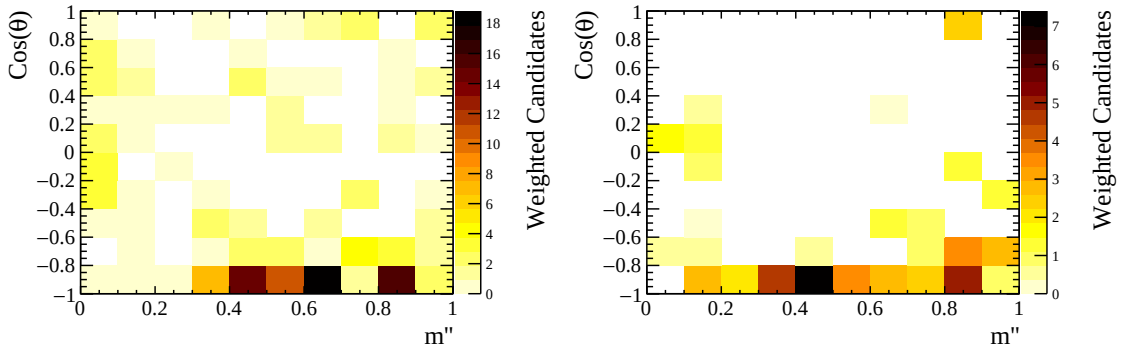


Figure 5.16: Left(Right):The distribution of background subtracted $\Lambda_b^0 \to pK\eta'$ events as a function of $\cos(\theta_{\eta' p})$ and m'' in the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

by calculating the corrected efficiency as

$$\bar{\varepsilon} = \frac{\sum_i^N w_i}{\sum_i^N \frac{w_i}{\varepsilon_i}}, \tag{5.14}$$

where w_i is the $sWeight$ of event i and ε_i is the efficiency of event i taken from the total efficiency maps shown in Figure 4.29. The values of the corrected rare channel efficiencies are given in Table 5.10.

The phase space corrected efficiencies are 19% and 33% higher than the phase space integrated efficiencies in the $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$ channels respectively. The background-subtracted distributions of $\Lambda_b^0 \to pK\eta'$ signal events are shown in Figure 5.16, which show a concentration of signal events at low values of $\cos(\theta_{\eta' p})$. Signal events are expected to occupy this area of the phase space when they proceed through a $\Lambda^* \to pK^-$ resonance (this is discussed further in section 5.8). This region of the phase space also has the highest signal efficiency (see Figure 4.29), which explains why the corrected efficiencies are considerably larger than the phase space integrated efficiencies.

5.6 Systematic Uncertainties on Branching Fraction

When the branching fraction is determined using the "Ratio Fit" there are several parameters, both in the fit model and Equation (4.1), which are fixed to their central values. The systematic uncertainties on these parameters need to be taken into account. Following the procedure used to evaluate the systematic uncertainty on the signal yields due to the fixed fit model parameters in Section 5.3, the "Ratio Fit" is repeated 1000 times whilst sampling the fixed parameters from Gaussian distributions with width equal to their systematic uncertainty. The RMS of the resulting branching fraction ratio distribution is then taken as the systematic uncertainty. The fixed parameters which are varied are as follows:

- All of the fit model shape parameters that are fixed to MC: the uncertainties on these parameters are shown in Table 5.1 and a description of how they are varied whilst preserving correlations is given in Section 5.3.
- The M(Λ_b^0)-M(B^+) value, used to constrain the means of the signal shapes.
- The individual branching fractions $\mathcal{B}\,(\eta' \to \pi^+\pi^-\gamma\,)$ and $\mathcal{B}\,(\eta' \to \pi^+\pi^-\eta\,)$; the uncertainties on these values are taken from their PDG averages [1].
- The ratios of fragmentation fractions, $\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_{\gamma,\eta}$, the uncertainties on which are given in Section 4.1.1. As the uncertainties on the fragmentation fractions are 100% correlated, the same change in the fragmentation fraction will be used for both $\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_{\gamma}$ and $\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_{\eta}$ in order to preserve the correlation.
- The ratios of efficiencies; these are subject to several sources of systematic uncertainties which are discussed in Section 5.6.1.

The systematic uncertainties on the branching fraction ratio due to the possible $\Lambda_b^0 \to pK^-\phi$ background, choice of fit model and the σ_{scale} parameter are assessed in the same manner as for the signal yield (see Section 5.3.2 – 5.3.4). However, the

	Δ R
Model I:	0.0012
Model II:	0.0015
Model III:	0.0006

Table 5.11: A summary of the changes in branching fraction ratio for each of the alternative fit models considered.

Systematic	Absolute Value	Relative Uncertainty
Fixed Parameters (incl. Efficiency Ratios)	0.0130	10.8%
$\Lambda_b^0 \to pK^-\phi$ background	0.0004	0.3%
σ_{scale}	0.0014	1.2%
Choice of Fit Model	0.0015	1.3%
Vetoes	0.0010	0.8%
Total	0.0131	10.9%

Table 5.12: A summary of all systematic uncertainties on the branching fraction ratio. Both the absolute values and the percentage uncertainties are shown.

"Ratio Fit" is instead used and the change in branching fraction ratio is considered in each case. A summary of the changes in branching fraction ratio for each alternative fit model considered is given in Table 5.11.

An additional source of systematic uncertainty on the ratio of branching fractions arises from the mass vetoes used in the $\eta' \to \pi^+\pi^-\gamma$ rare channel, which are described in Section 4.2.5. As certain regions of the phase space may be completely depleted by these vetoes, there may be no knowledge of the signal efficiency in these regions of the phase space. Therefore, the width of each of the mass vetoes is increased by 50% and the corrected value of ε_γ is recalculated. Any change in the ratio of branching fractions when the fit is reperformed with the increased veto size and the recalculated ε_γ value is assigned as a systematic uncertainty.

A summary of all systematic uncertainties on the branching fraction ratio is given in Table 5.12. The systematic uncertainties on the efficiency ratios are discussed in detail in section 5.6.1.

Systematic	$\eta' \to \pi^+\pi^-\gamma$	$\eta' \to \pi^+\pi^-\eta$
L0 Trigger	1.6%	3.1%
Finite MC statistics	3.2%	5.8%
Phase Space Binning	0.4%	2.2%
BDT Selection	3.5%	2.7%
PID Efficiency	2.5%	2.5%
Photon Efficiency Correction	0.1%	4.0%
Total	5.6%	8.9%

Table 5.13: A summary of all the systematics on the efficiency ratios

5.6.1 Efficiency Ratio Systematics

It is possible for systematic uncertainties on the efficiency ratio to occur due to the mis-modelling in MC of quantities used by the selection. The majority of these uncertainties will cancel because the equivalent mis-modelling will also be present in the control channel. The following sections describe the systematics assigned to account for any uncertainties that do not fully cancel. A summary of all systematic uncertainties on the efficiency ratios is presented in Table 5.13.

5.6.1.1 L0 Trigger

It is known that there is a mis-calibration of transverse energy measurements made by the hadronic calorimeter in MC. As the `L0Hadron` trigger requirements rely heavily on E_{T} thresholds, the L0 trigger efficiencies determined from MC may not be accurate. Data calibration samples are available which, in theory, facilitate the determination of the `L0Hadron` trigger efficiency without using the mis-calibrated MC. However, studies of these calibration samples have shown that the overlap between different tracks is not fully accounted for; it is possible for more than one track to hit the same cell of the hadronic calorimeter which affects the probability of the trigger being fired. Consequently, these calibration samples are used to assess a systematic uncertainty rather than correct the `L0Hadron` trigger efficiency. The efficiencies are recalculated using the data calibration samples and the differences relative to the

nominal efficiency ratios are taken as systematic uncertainties, the values of which are shown in Table 5.13.

5.6.1.2 Finite MC statistics

The statistical uncertainties on the average efficiency in the control channel and the uncertainties as a function of the phase space variables are presented in Section 4.3. In order to propagate these uncertainties to the ratio of branching fractions, 1000 new efficiency maps are generated by sampling each bin from a Gaussian with mean equal to the values in the nominal efficiency map and width equal to the uncertainty due to finite MC stats. The corrected efficiency is then recalculated for each sampled efficiency map. The value of the control channel efficiency is also sampled from a Gaussian in the same manner. The standard deviation of the distributions of the resulting efficiency ratios are taken as the systematic uncertainties due to finite MC statistics, the values of which are shown in Table 5.13.

5.6.1.3 Phase Space Binning

The use of a binned approach when the efficiency maps are created introduces uncertainty from the assumption that the efficiency is smoothly varying across each bin. The size of this uncertainty is reduced with the use of cubic spline interpolation between bins but a systematic uncertainty still needs to be evaluated. This is done by reducing the number of bins used along each axis of the efficiency maps from 10 to 7 and recalculating the efficiency ratios. The relative difference to the nominal efficiency ratio is taken as the systematic uncertainty; the values of this for each channel are shown in Table 5.13.

5.6.1.4 BDT Selection

Systematic uncertainty can be introduced on the efficiency of the BDT selection if variables entering the BDT are not correctly modelled. Although the control channel was used to check the modelling of all variables entering the BDT to try to eliminate this (see Figure 4.2), comparisons of the BDT distribution in control channel MC and *sWeighted* data show small residual discrepancies. To evaluate this uncertainty the nominal BDT cuts are applied to both $B^+ \to K^+\eta'$ MC and *sWeighted* data and the difference in efficiency is taken as the uncertainty on the BDT efficiency. The same method is used for the rare channel BDT efficiencies, but using the rare channel BDT cuts. To estimate the correlation between the control and rare channel BDT output, the control channel BDT is applied to rare channel MC and the output of the control channel and rare channel BDT is plotted event-by-event, from which the correlation coefficient can be calculated. Using this correlation, the efficiency uncertainties are propagated to uncertainties on the efficiency ratios. The correlation coefficients are 0.86 and 0.74 for the $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$ rare channels respectively, and the systematic uncertainties assigned are shown in Table 5.13.

5.6.1.5 PID Efficiency

The uncertainty on the PID efficiency due to the limited size of both the calibration and reference samples is already accounted for in the statistical uncertainties quoted in Section 4.3.3 and these are propagated to a systematic uncertainty on the efficiency ratios as described in Section 5.6.1.2. However, another source of uncertainty arises from the fact a binned approach is used. It is assumed that the efficiency within in each bin of the calibration sample is smoothly varying, but this may not be the case, particularly for protons where large bins have to be used due to the low statistics of the calibration samples (see Figure 4.24). To evaluate a systematic uncertainty due to this assumption, every bin boundary is shifted in the direction

of both increasing and decreasing p/ η/`nTracks` by half the width of the average bin. The efficiency ratios are recalculated and the larger of the two changes in the efficiency ratios, with respect to the nominal binning schemes, is assigned as a systematic uncertainty.

5.6.1.6 Photon Efficiency Correction

The photon reconstruction efficiency correction factors used, as described in Section 4.3.5, have both statistical and systematic uncertainties. These are propagated to the efficiency ratios by varying the correction factors to $\varepsilon_{cor} + 1\sigma$ and $\varepsilon_{cor} - 1\sigma$, where σ is the sum of statistical and systematic uncertainites in quadrature, before recalculating the efficiency ratios. The average difference relative to the nominal efficiency ratios is taken as a systematic uncertainty.

5.7 Branching Fraction Results

The ratio of branching fractions $\mathcal{B}(\Lambda_b^0 \to pK\eta')/\mathcal{B}(B^+ \to K^+\eta')$ is given by Equation (4.1). The phase space corrected efficiencies for the rare decay channels are shown in Table 5.10; the control channel efficiency is shown in Table 4.20; the fragmentation fractions are given in Table 4.1 and the values of the η' branching fractions are taken from the PDG [1]. These are all used in the "Ratio Fit" to extract the ratio of branching fractions and the statistical uncertainty. The total systematic uncertainty is given in Table 5.12. The final value for the ratio of branching fractions is

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\eta')}{\mathcal{B}(B^+ \to K^+\eta')} = 0.120 \pm 0.013(\text{stat.}) \pm 0.013(\text{sys.}). \tag{5.15}$$

Using the PDG value for the control channel branching fraction, $\mathcal{B}(B^+ \to K^+\eta') = (7.06 \pm 0.25) \times 10^{-5}$, the final value for the branching fraction of the decay $\Lambda_b^0 \to$

$pK^-\eta'$ is

$$\mathcal{B}(\Lambda_b^0 \to pK\eta') = (8.48 \pm 0.88(\text{stat.}) \pm 0.97(\text{sys.})) \times 10^{-6}. \tag{5.16}$$

As a cross check, the branching fraction has been calculated individually for each η' decay channel; the results are consistent between decay channels. The details of this calculation can be found in Appendix D.

5.8 Resonant Structure

It is of interest to understand which resonances the decay $\Lambda_b^0 \to pK\eta'$ proceeds through. There is a plethora of Λ^* resonances which could contribute, such that the decay proceeds $\Lambda_b^0 \to (\Lambda^* \to pK^-)\eta'$. A recent amplitude analysis of $\Lambda_b^0 \to pK^- J/\psi$ revealed contributions from 13 Λ^* resonances, with the strongest contribution being from $\Lambda^*(1520)$ [16]. However, this $fermion \to fermion + pseudoscalar + vector$ decay is quite different to $\Lambda_b^0 \to pK\eta'$ which is $fermion \to fermion + pseudoscalar + pseudoscalar$. To investigate whether any resonant structure is present, the background-subtracted Dalitz plots of $\Lambda_b^0 \to pK\eta'$ signal events are studied; these are shown in Figure 5.17. These Dalitz plots show a concentration of signal events in the region $2\,\text{GeV}^2 < M(pK^-)^2 < 4\,\text{GeV}^2$, which is where an excess of events would be expected if the decay was proceeding through Λ^* resonances.

To investigate this resonant structure further, the background subtracted $M(pK^-)$ invariant mass distributions are shown in Figure 5.18. These show possible evidence for contributions from resonances between 1800 MeV and 1900 MeV. There is experimental evidence for five Λ^* resonances in this region ($\Lambda^*(1800)$, $\Lambda^*(1810)$, $\Lambda^*(1820)$, $\Lambda^*(1830)$, $\Lambda^*(1890)$) so their presence is not surprising [1]. However, the strength of these resonances relative to the $\Lambda^*(1520)$ is greater than seen in other baryonic decay channels [16,127]. In other baryonic decay channels the $\Lambda^*(1520)$ is the dom-

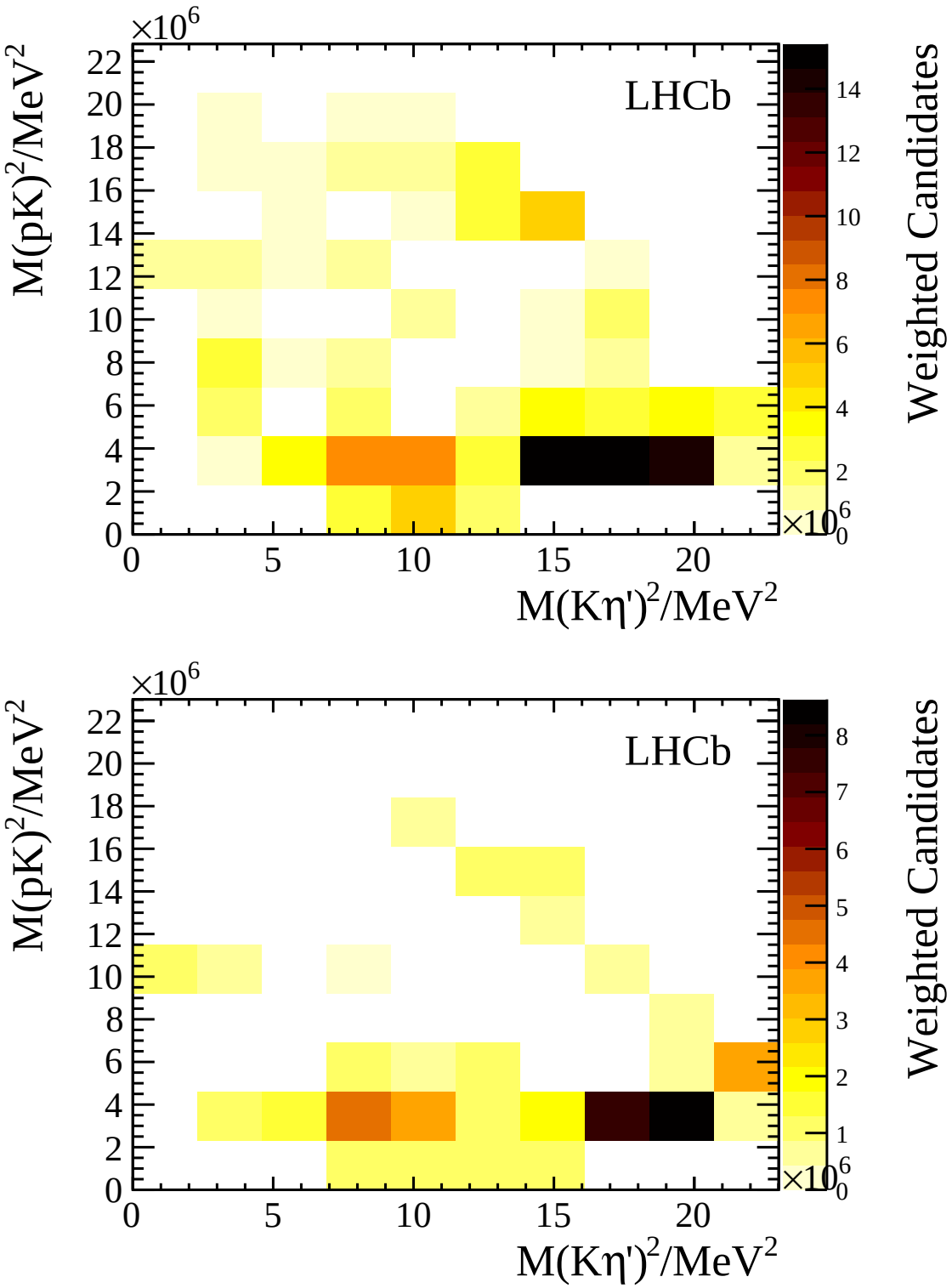


Figure 5.17: Top(Bottom): Dalitz plot distributions of background subtracted $\Lambda_b^0 \to pK\eta'$ events in the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) decay channel.

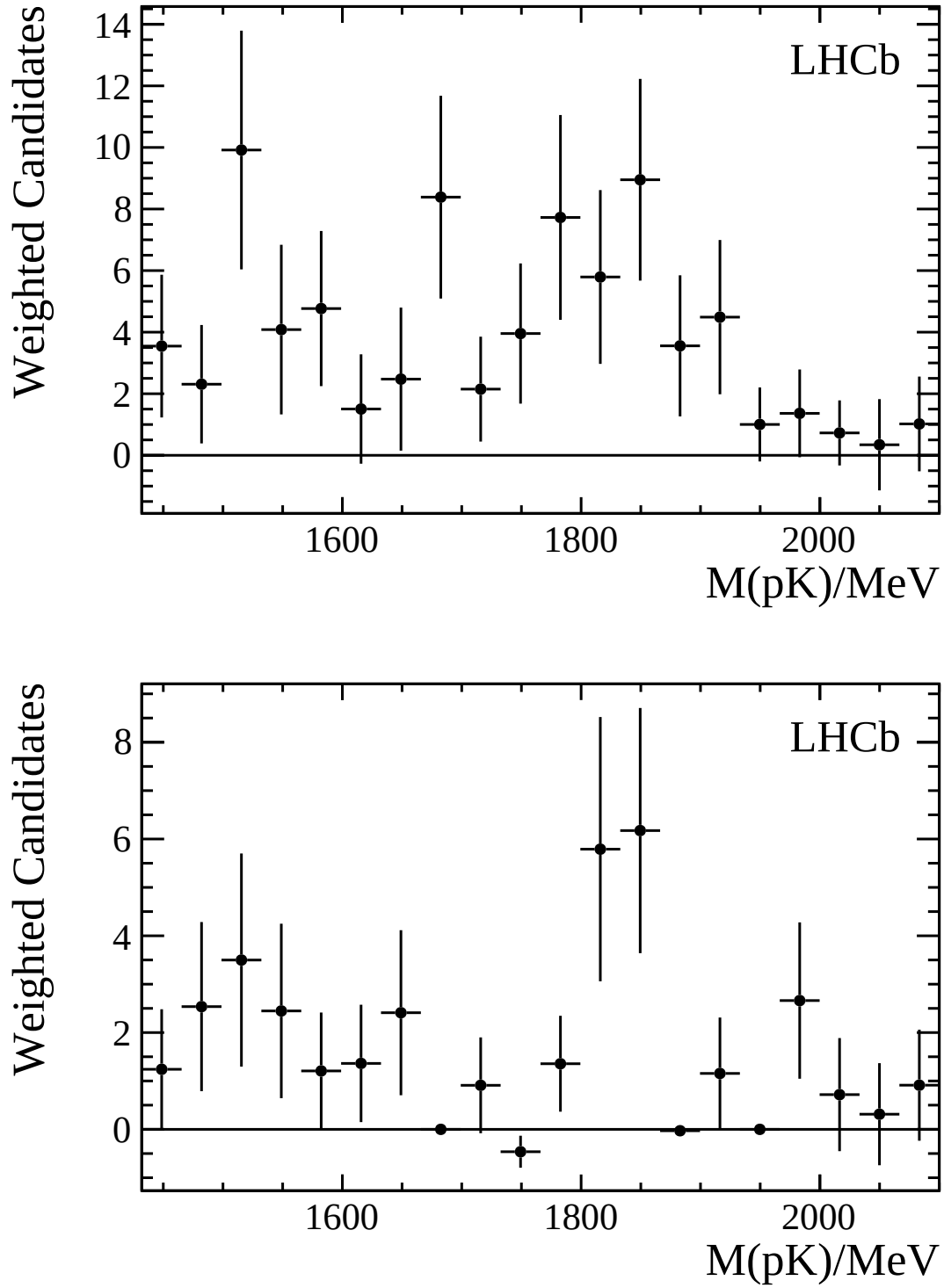


Figure 5.18: Top(Bottom): Background subtracted M(pK^-) distribution in the region where Λ^* resonances are expected for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

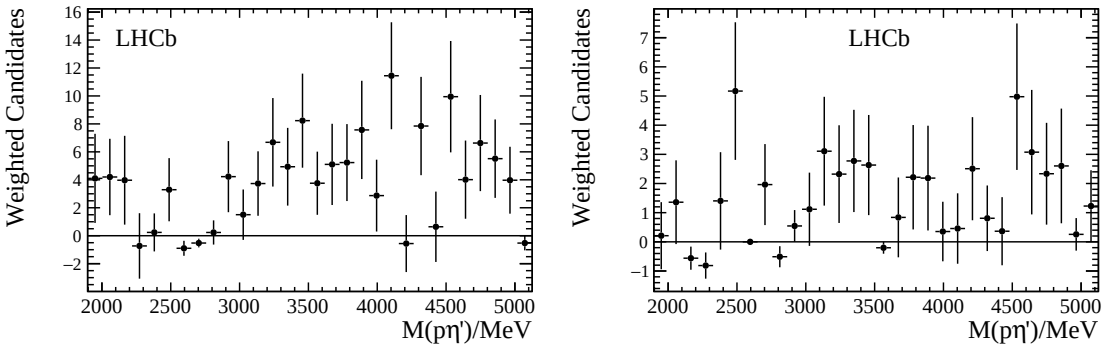


Figure 5.19: Left (Right): $M(p\eta')$ distribution for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) decay channel.

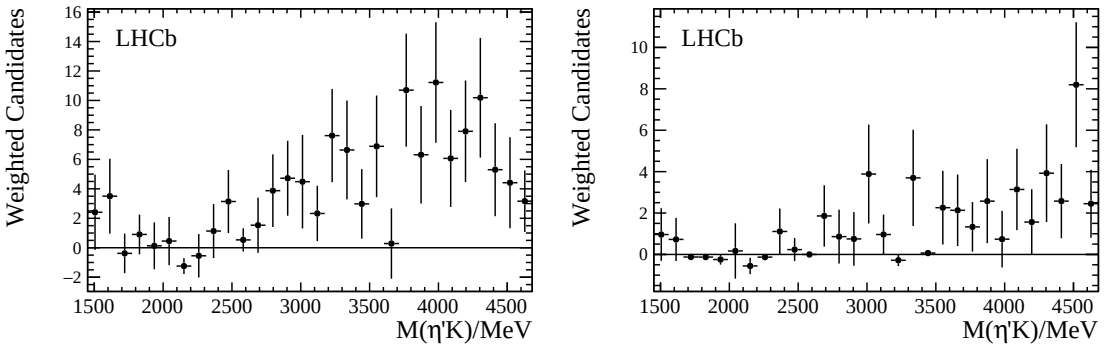


Figure 5.20: Left (Right): $M(\eta' K^-)$ distribution for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) decay channel.

inant resonance. An amplitude analysis would be required to fully understand and disentangle these resonances, but unfortunately much larger samples of signal events would be required for this.

The $M(p\eta')$ and $M(K^-\eta')$ invariant mass distributions are also studied and are shown in Figures 5.19 and 5.20. These show no particular resonant structure; the accumulation of candidates at high mass in the $M(\eta' K^-)$ distribution is expected when Λ^* resonances are present in the $M(pK^-)$ spectrum.

CHAPTER 6

Conclusions

A search has been performed for the decay $\Lambda_b^0 \to pK\eta'$ using the LHCb Run I dataset, which consists of $2\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 8\,\text{TeV}$ in 2012 and $1\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 7\,\text{TeV}$ in 2011. The search is performed in two decay channels to maximise sensitivity; the η' is reconstructed through both the channel $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$. An event selection, which mainly makes use of multivariate algorithms and the LHCb PID system, has been optimised to maximise sensitivity to this decay. In the $\eta' \to \pi^+\pi^-\gamma$ decay channel $117 \pm 15(\text{stat.}) \pm 10(\text{sys.})$ signal events are observed, and $45 \pm 8(\text{stat.}) \pm 2(\text{sys.})$ events are observed in the $\eta' \to \pi^+\pi^-\eta$ decay channel. The combined statistical significance of these yields is 12.0σ. This is therefore the first observation of a beauty baryon decaying to a final state involving an η' particle. The branching fraction of the decay $\Lambda_b^0 \to pK\eta'$ is

measured relative to the decay $B^+ \to K^+\eta'$ to be

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\eta')}{\mathcal{B}(B^+ \to K^+\eta')} = 0.120 \pm 0.013(\text{stat.}) \pm 0.013(\text{sys.}). \tag{6.1}$$

Using the world average value for the branching fraction of the decay $B^+ \to K^+\eta'$ [1], the branching fraction of $\Lambda_b^0 \to pK\eta'$ is measured to be

$$\mathcal{B}(\Lambda_b^0 \to pK\eta') = 8.48 \pm 0.88(\text{stat.}) \pm 0.97(\text{sys.}) \times 10^{-6}. \tag{6.2}$$

This is the first measurement at the LHC of the branching fraction of a charmless three body b-hadron decay, which has been reconstructed using neutral particles. Unfortunately, however, with no theoretical predictions for this branching fraction currently available it is not possible to make direct comparisons to the SM.

The resonant $m(pK^-)$ structure has been inspected and these distributions suggest there could be excited Λ^* resonances present in the region $1800 - 1900\,\text{MeV}$. The presence of these resonances is not surprising but the relative strength of these resonances in the $1800-1900\,\text{MeV}$ region compared to the $\Lambda^*(1520)$ resonance differs from what has been seen in other decay channels [16, 127]. In other decay channels the $\Lambda^*(1520)$ is the dominant resonance, but that does not appear to be the case here.

In order to understand fully the resonant structure of this decay an amplitude analysis would be required, but this would only be possible with significantly more than the 162 signal events observed in this analysis. With the future inclusion of the LHCb Run II dataset this may become possible, as $3.8\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ data has already been collected and up to another $2\,\text{fb}^{-1}$ could be added during 2018 data taking. An amplitude analysis could also be used to search for CPV in this decay channel, which would be of interest given that the first evidence for CPV in the baryonic sector was found in the charmless decay $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ [67] and large

CP asymmetries have been observed in charmless B meson decays to an $\eta^{(\prime)}$ final state [128, 129].

The inclusion of the Run II dataset would also reduce the statistical uncertainty on the branching fraction measurement. In order to reduce the systematic uncertainties a more precise measurement of the Λ_b^0 fragmentation fraction would be highly desirable, as the $\approx 8\%$ uncertainty on this quantity is one of the dominant sources of systematic uncertainty. Further improvements could also be made by using the channel $B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ as a normalisation channel for the corresponding $\Lambda_b^0 \to pK\eta'$ decay channel[1]. This would reduce the systematic uncertainty on the photon reconstruction efficiency corrections, in this channel, from 4.04% to the sub 1% level seen in the $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$ decay channel. More efficient generation of simulated samples would also be advantageous, as this would reduce the systematic uncertainties on the efficiency ratios due to the finite size of the simulated samples.

It would also be of interest to determine the branching fraction of the decay $\Lambda_b^0 \to pK\eta$. As discussed in Section 1.4, the branching fraction of B meson decays to an η' meson have been found to be significantly different to the same decay involving an η meson. It would therefore be of interest to understand whether this pattern continues in beauty baryon decays. Moreover, a measurement with similar precision to that presented here could be used in combination with the branching fraction of $\Lambda_b^0 \to pK\eta'$ to add valuable information on η'– η mixing. This could potentially lead to the first determination of the η'– η mixing angle, $\phi_{\rm p}$, using a baryonic decay and add constraints on the size of the gluon component in the η' wavefunction.

[1]It is not currently possible to reconstruct this channel in Run I data due to selection requirements which were applied during the centralised data processing campaign, but this is not the case for Run II data

REFERENCES

[1] C. Patrignani *et al.*, "Review of particle physics," *Chin. Phys.*, vol. C40, p. 100001, 2016. and 2017 update.

[2] J. J. Thomson, "Xl. cathode rays," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 44, no. 269, pp. 293–316, 1897.

[3] E.Rutherford, "Bakerian lecture: Nuclear constitution of atoms," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 97, no. 686, pp. 374–400, 1920. DOI:10.1098/rspa.1920.0040.

[4] J. Chadwick, "Possible Existence of a Neutron," *Nature*, vol. 129, p. 312, 1932.

[5] G. P. S. Occhialini and C. F. Powell, "Nuclear disintegrations produced by slow charged particles of small mass," *Nature*, vol. 159, pp. 186 EP –, 02 1947.

[6] G. D. Rochester and C. C. Butler, "Evidence for the existence of new unstable elementary particles," *Nature*, vol. 160, pp. 855 EP –, 12 1947.

[7] V. D. Hopper and S. Biswas, "Evidence concerning the existence of the new unstable elementary neutral particle," *Phys. Rev.*, vol. 80, pp. 1099–1100, Dec 1950.

[8] M. Gell-Mann, "A schematic model of baryons and mesons," *Physics Letters*, vol. 8, no. 3, pp. 214 – 215, 1964.

[9] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 2," in *Developments in the Quark Theory of Hadrons. VOL. 1. 1964 - 1978* (D. Lichtenberg and S. P. Rosen, eds.), pp. 22–101, 1964.

[10] Bloom, E. D. *et al.*, "High-energy inelastic $e - p$ scattering at 6° and 10°," *Phys. Rev. Lett.*, vol. 23, pp. 930–934, Oct 1969.

[11] B. Bjørken and S. Glashow, "Elementary particles and SU(4)," *Physics Letters*, vol. 11, no. 3, pp. 255 – 257, 1964.

[12] M. Kobayashi and T. Maskawa, "CP -violation in the renormalizable theory of weak interaction," *Progress of Theoretical Physics*, vol. 49, no. 2, pp. 652–657, 1973.

[13] F. Abe *et al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.

[14] S. Abachi *et al.*, "Observation of the top quark," *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, 1995.

[15] R. Aaij *et al.*, "A model-independent confirmation of the $Z(4430)^-$ state," *Phys. Rev.*, vol. D92, p. 112009, 2015.

[16] R. Aaij *et al.*, "Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays," *Phys. Rev. Lett.*, vol. 115, p. 072001, 2015.

[17] P.W. Higgs, "Broken symmetries, massless particles and gauge fields," *Physics Letters*, vol. 12, no. 2, pp. 132 – 133, 1964.

[18] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[19] 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.

[20] 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.

[21] G. Aad *et al.*, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments," *Phys. Rev. Lett.*, vol. 114, p. 191803, 2015.

[22] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963.

[23] M. Kobayashi and T. Maskawa, "CP-violation in the renormalizable theory of weak interaction," *Progress of Theoretical Physics*, vol. 49, no. 2, pp. 652–657, 1973.

[24] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π decay of the k_2^0 meson," *Phys. Rev. Lett.*, vol. 13, pp. 138–140, Jul 1964.

[25] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa matrix," *Phys. Rev. Lett.*, vol. 51, pp. 1945–1947, Nov 1983.

[26] R. Aaij *et al.*, "Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays," *Nature Physics*, vol. 11, p. 743, 2015.

[27] Aubert, B. *et al.*, "Improved measurement of $B^+ \to \rho^+\rho^0$ and determination of the quark-mixing phase angle α," *Phys. Rev. Lett.*, vol. 102, p. 141802, Apr 2009.

[28] Aubert, B. et al., "Measurement of the Branching Fraction, Polarization, and CP Asymmetries in $B^0 \to \rho^0\rho^0$ Decay, and Implications for the CKM Angle α," *Phys. Rev.*, vol. D78, p. 071104, 2008.

[29] R. Aaij *et al.*, "Measurement of *CP* violation in $B^0 \to J/\psi\, K^0_S$ decays," *Phys. Rev. Lett.*, vol. 115, p. 031601, 2015.

[30] R. Aaij *et al.*, "Measurement of the CKM angle γ from a combination of LHCb results," *JHEP*, vol. 12, p. 087, 2016.

[31] Y. Amhis *et al.*, "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016," 2016. updated results and plots available at `http://www.slac.stanford.edu/xorg/hflav/`.

[32] CKMfitter Group, J. Charles et al., "Updated results and plots available at: `http://ckmfitter.in2p3.fr/www/results/plots_ichep16/ckm_res_ichep16.html`," *Eur. Phys. J.*, vol. C41, pp. 1–131, 2005.

[33] "Measurement of the CKM angle γ from a combination of $B \to DK$ analyses," Jul 2017. `http://cds.cern.ch/record/2275866/`.

[34] N. Isgur, "Mass formula for nonets," *Phys. Rev. D*, vol. 13, pp. 122–124, Jan 1976.

[35] H. Fritzsch and J. Jackson, "Mixing of pseudoscalar mesons and m1 radiative decays," *Physics Letters B*, vol. 66, no. 4, pp. 365 – 369, 1977.

[36] C. Di Donato, G. Ricciardi, and I. I. Bigi, "$\eta - \eta'$ mixing: From electromagnetic transitions to weak decays of charm and beauty hadrons," *Phys. Rev. D*, vol. 85, p. 013016, Jan 2012.

[37] KLOE Collaboration, F. Ambrosino et al., "Measurement of the pseudoscalar mixing angle and η' gluonium content with the KLOE detector," *Physics Letters B*, vol. 648, no. 4, pp. 267 – 273, 2007.

[38] G. Li, Q. Zhao, and C.-H. Chang, "Decays of J/ψ and $\psi(2S)$ into vector and pseudoscalar mesons and the pseudoscalar glueball $q\bar{q}$ mixing," *Journal of Physics G: Nuclear and Particle Physics*, vol. 35, no. 5, p. 055002, 2008.

[39] Harland-Lang, L. A. *et al.*, "χ_c decays and the gluon content of the η', η mesons," *Phys. Lett.*, vol. B770, pp. 88–92, 2017.

[40] P. Kroll and K. Passek-Kumericki, "The $\eta^{(\prime)}$ gamma transition form factor and the gluon-gluon distribution amplitude," *J. Phys.*, vol. G40, p. 075005, 2013.

[41] Christ, N. H. *et al.*, "The η and η' mesons from Lattice QCD," *Phys. Rev. Lett.*, vol. 105, p. 241601, 2010.

[42] Dudek, Jozef J. *et al.*, "Isoscalar meson spectroscopy from lattice QCD," *Phys. Rev.*, vol. D83, p. 111502, 2011.

[43] R. Fleischer, R. Knegjens, and G. Ricciardi, "Exploring CP Violation and η'–η Mixing with the $B^0_{s,d} \to J/\psi\eta^{(\prime)}$ Systems," *Eur. Phys. J.*, vol. C71, p. 1798, 2011.

[44] R. Aaij *et al.*, "Study of η–η' mixing from measurement of $B^0_{(s)} \to J/\psi\,\eta^{(\prime)}$ decay rates," *JHEP*, vol. 01, p. 024, 2015.

[45] M. C. Chang *et al.*, "Measurement of $B^0 \to J/\psi\eta^{(\prime)}$ and Constraint on the $\eta - \eta'$ Mixing Angle," *Phys. Rev.*, vol. D85, p. 091102, 2012.

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

[47] R. Escribano, "$J/\psi \to PV$ decays and the quark and gluon content of the η and η'," *Eur. Phys. J.*, vol. C65, pp. 467–473, 2010.

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

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

[50] F. Ambrosino *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.

[51] V. V. Anisovich, D. V. Bugg, D. I. Melikhov, and V. A. Nikonov, "η'– η glueball mixing from photon - meson transition form-factors and decay ratio $D_s^+ \to \eta^{(\prime)}\ell\nu$," *Phys. Lett.*, vol. B404, pp. 166–172, 1997.

[52] F.-G. Cao and A. I. Signal, "Two analytical constraints on the η'– η mixing," *Phys. Rev.*, vol. D60, p. 114012, 1999.

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

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

[55] H.-Y. Cheng and B. Tseng, "Charmless B decays to η' and η," *Phys. Lett.*, vol. B415, pp. 263–272, 1997.

[56] A. Ali, J. Chay, C. Greub, and P. Ko, "Contribution of $s \to sgg$ through the QCD anomaly in exclusive decays $B^\pm \to \eta^{(\prime)}(K^\pm, K^{*\pm})$ and $B^0 \to \eta^{(\prime)}(K^0_{\rm S}, K^*)$," *Phys. Lett.*, vol. B424, pp. 161–174, 1998.

[57] R. Aaij *et al.*, "Search for the $\Lambda_b^0 \to \Lambda\eta$ and $\Lambda_b^0 \to \Lambda\eta'$ decays with the LHCb detector," *JHEP*, vol. 09, p. 006, 2015.

[58] G. Duplancic and B. Melic, "Form factors of B, $B_s^0 \to \eta^{(\prime)}$ and D, $D_s^+ \to \eta^{(\prime)}$ transitions from QCD light-cone sum rules," *JHEP*, vol. 11, p. 138, 2015.

[59] P. Ball and G. W. Jones, "$B \to \eta^{(\prime)}$ form factors in QCD," *Journal of High Energy Physics*, vol. 2007, no. 08, p. 025, 2007.

[60] P. Ball, "$B \to \pi$ and $B \to K$ transitions from QCD sum rules on the light cone," *JHEP*, vol. 09, p. 005, 1998.

[61] P. Ball and R. Zwicky, "New results on B to π, K, η decay form factors from light-cone sum rules," *Phys. Rev.*, vol. D71, p. 014015, 2005.

[62] Y.-Y. Charng, T. Kurimoto, and H.-n. Li, "Gluonic contribution to $B \to \eta^{(\prime)}$ form factors," *Phys. Rev.*, vol. D74, p. 074024, 2006. [Erratum: Phys. Rev.D78,059901(2008)].

[63] 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.

[64] 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. f_s/f_d value updated in LHCb-CONF-2013-011.

[65] M. R. Ahmady, C. S. Kim, S. Oh, and C. Yu, "Heavy baryonic decays of $\Lambda_b^0 \to \Lambda\eta^{(\prime)}$ decays," *Phys. Lett.*, vol. B598, pp. 203–210, 2004.

[66] C. Q. Geng, Y. K. Hsiao, Y.-H. Lin, and Y. Yu, "Study of $\Lambda_b \to \Lambda(\phi, \eta^{(\prime)})$ and $\Lambda_b \to \Lambda K^+ K^-$ decays," *Eur. Phys. J.*, vol. C76, no. 7, p. 399, 2016.

[67] R. Aaij *et al.*, "Measurement of matter-antimatter differences in beauty baryon decays," *Nature Physics*, vol. 13, p. 391, 2017.

[68] L. R. Evans and P. Bryant, "LHC Machine," *JINST*, vol. 3, p. S08001. 164 p, 2008. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[69] R. Aaij *et al.*, "LHCb detector performance," *Int. J. Mod. Phys.*, vol. A30, p. 1530022, 2015.

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

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

[72] H. Abramowicz *et al.*, "Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data," *Eur. Phys. J.*, vol. C75, no. 12, p. 580, 2015.

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

[74] "LHCb magnet: Technical Design Report," 2000. LHCb-TDR-001, `https://cds.cern.ch/record/424338`.

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

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

[77] "LHCb Trigger and Online Technical Design Report," 2014. LHCb-TDR-016, `https://cds.cern.ch/record/1701361`.

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

[79] J. Badier *et al.*, "Shashlik calorimetry a combined Shashlik + preshower detector for LHC: R & D proposal," 1993.

[80] E. Picatoste Olloqui, "LHCb preshower(PS) and scintillating pad detector (SPD): Commissioning, calibration, and monitoring," *J. Phys. Conf. Ser.*, vol. 160, p. 012046, 2009.

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

[82] M Alemi et al., "First operation of a hybrid photon detector prototype with electrostatic cross-focussing and integrated silicon pixel readout," *Nuclear Inst.ruments and Methods A*, vol. 449, no. 1, pp. 48 – 59, 2000.

[83] T. Bellunato et al., "Performance of aerogel as cherenkov radiator," *Nuclear Instruments and Methods A*, vol. 519, no. 3, pp. 493 – 507, 2004.

[84] "Measurement of time-dependent *CP* violating asymmetries in $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ decays at LHCb," Jan 2017. LHCb-CONF-2016-018, `https://cds.cern.ch/record/2243039`.

[85] R. Aaij *et al.*, "Measurement of the time-dependent *CP* asymmetries in $B_s^0 \to J/\psi\, K_S^0$," *JHEP*, vol. 06, p. 131, 2015.

[86] R. Aaij *et al.*, "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays," *JHEP*, vol. 08, p. 055, 2017.

[87] R. Aaij *et al.*, "Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays," *Phys. Rev. Lett.*, vol. 118, p. 191801, 2017.

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

[89] O. Callot and S. Hansmann-Menzemer, "The Forward Tracking," 2007. https://cds.cern.ch/record/1033584.

[90] 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.

[91] S. Agostinelli *et al.*, "Geant4: A simulation toolkit," *Nucl. Instrum. Meth.*, vol. A506, p. 250, 2003.

[92] M. Clemencic *et al.*, "The LHCb simulation application, Gauss: Design, evolution and experience," *J. Phys. Conf. Ser.*, vol. 331, p. 032023, 2011.

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

[94] G. A. Cowan, D. C. Craik, and M. D. Needham, "RapidSim: an application for the fast simulation of heavy-quark hadron decays," *Comput. Phys. Commun.*, vol. 214, pp. 239–246, 2017.

[95] V. N. Ivanchenko, O. Kadri, M. Maire, and L. Urban, "Geant4 models for simulation of multiple scattering," *Journal of Physics: Conference Series*, vol. 219, no. 3, p. 032045, 2010.

[96] J. Apostolakis *et al.*, "Validation and verification of Geant4 standard electromagnetic physics," *Journal of Physics: Conference Series*, vol. 219, no. 3, p. 032044, 2010.

[97] McCarthy, James, *Search for rare baryonic b decays with the LHCb experiment at the LHC*. PhD thesis. CERN-THESIS-2015-159.

[98] R. Aaij *et al.*, "Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays," *Phys. Rev. Lett.*, vol. 113, p. 151601, 2014.

[99] R. Eisberg and R. Resnick, *Quantum physics of atoms, molecules, solids, nuclei, and particles*. Wiley, 1974.

[100] H. A. Bethe, "Molière's theory of multiple scattering," *Phys. Rev.*, vol. 89, pp. 1256–1266, Mar 1953.

[101] B. Gottschalk, A. Koehler, R. Schneider, J. Sisterson, and M. Wagner, "Multiple coulomb scattering of 160 MeV protons," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 74, no. 4, pp. 467 – 490, 1993.

[102] H. W. Lewis, "Multiple scattering in an infinite medium," *Phys. Rev.*, vol. 78, pp. 526–529, Jun 1950.

[103] V. L. Highland, "Some practical remarks on multiple scattering," *Nuclear Instruments and Methods*, vol. 129, no. 2, pp. 497 – 499, 1975.

[104] GEANT4 Collaboration, "GEANT4 Physics Reference Manual v10.3," 2016. http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsReferenceManual/BackupVersions/V10.3/fo/PhysicsReferenceManual.pdf.

[105] M. Clemencic and B. Couturier, "LHCb build and deployment infrastructure for run 2," *Journal of Physics: Conference Series*, vol. 664, no. 6, p. 062008, 2015.

[106] R. Aaij *et al.*, "Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions," *Phys. Rev. Lett.*, vol. 118, p. 052002, 2017. Erratum: Phys. Rev. Lett.119,no.16,169901(2017).

[107] R. Aaij *et al.*, "Observation of the $B_s^0 \to \eta'\eta'$ decay," *Phys. Rev. Lett.*, vol. 115, p. 051801, 2015.

[108] R. Aaij *et al.*, "Search for the $B_s^0 \to \eta'\phi$ decay," *JHEP*, vol. 05, p. 158, 2017.

[109] 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.

[110] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *Ann. Math.Stat.*, vol. 9, pp. 60–62, 1938.

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

[112] M. Hoballah, P. Perret, and O. Deschamps, "Measurement of the photon polarization using $B_s^0 \to \phi\gamma$ at LHCb.," Apr 2015. CERN-THESIS-2015-073, https://cds.cern.ch/record/2026848.

[113] Y. Amhis *et al.*, "HFLAV Averages for the PDG 2018 review," 2018. https://www.slac.stanford.edu/xorg/hfag/osc/PDG_2018/#FRACT.

[114] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, p. 119, 1997.

[115] A. Hocker *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[116] 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.

[117] E. M. Aitala *et al.*, "Multidimensional resonance analysis of $\Lambda_c^+ \to pK^-\pi^+$," *Phys. Lett.*, vol. B471, pp. 449–459, 2000.

[118] R. Aaij *et al.*, "Measurements of the $\Lambda_b^0 \to J/\psi\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7\,\text{TeV}$," *Phys. Lett.*, vol. B724, p. 27, 2013.

[119] C. J. Clopper and E. S. Pearson, "The use of confidence or fiducial limits illustrated in the case of the binomial," *Biometrika*, vol. 26, no. 4, pp. 404–413, 1934.

[120] Anderlini, Lucio *et al.*, "The PIDCalib package," Tech. Rep. LHCb-PUB-2016-021. `https://cds.cern.ch/record/2202412`.

[121] E. Govorkova, "Study of π^0/γ efficiency using B meson decays in the LHCb experiment," *Phys.Atom.Nucl.*, vol. 79, pp. 1474–1476. 3 p, May 2015.

[122] R. Aaij *et al.*, "Measurement of the $B_s^0 \to \phi\phi$ branching fraction and search for the decay $B^0 \to \phi\phi$," *JHEP*, vol. 10, p. 053, 2015.

[123] R. Aaij *et al.*, "Measurement of b-hadron masses," *Phys. Lett.*, vol. B708, p. 241, 2012.

[124] S. Fang, A. Kupsc, and D. Wei, "An overview of η and η' decays at BESIII," *Chin. Phys.*, vol. C42, no. 4, p. 042002, 2018. `https://arxiv.org/pdf/1710.05173.pdf`.

[125] D. Martínez Santos and F. Dupertuis, "Mass distributions marginalized over per-event errors," *Nucl. Instrum. Meth.*, vol. A764, pp. 150–155, 2014.

[126] K. S. Cranmer, "Kernel estimation in high-energy physics," *Comput. Phys. Commun.*, vol. 136, pp. 198–207, 2001.

[127] R. Aaij *et al.*, "Observation of the decay $\Xi_b^- \to pK^-K^-$," *Phys. Rev. Lett.*, vol. 118, p. 071801, 2017.

[128] L. Šantelj *et al.*, "Measurement of Time-Dependent CP Violation in the decay $B^0 \to K_S^0\eta'$," *JHEP*, vol. 10, p. 165, 2014.

[129] B. Aubert *et al.*, "Measurement of branching fractions and charge asymmetries in B decays to an η meson and a K* meson," *Phys. Rev. Lett.*, vol. 97, p. 201802, 2006.

APPENDIX A

Comparisons of 2011 MC and Background Subtracted Control Channel Data

This appendix compares variables, used in both the pre-selection and BDT, between MC and control channel background subtracted data. The background subtraction is performed using the *sPlot* technique [111]; the mass fit used is shown in Figure 4.1.

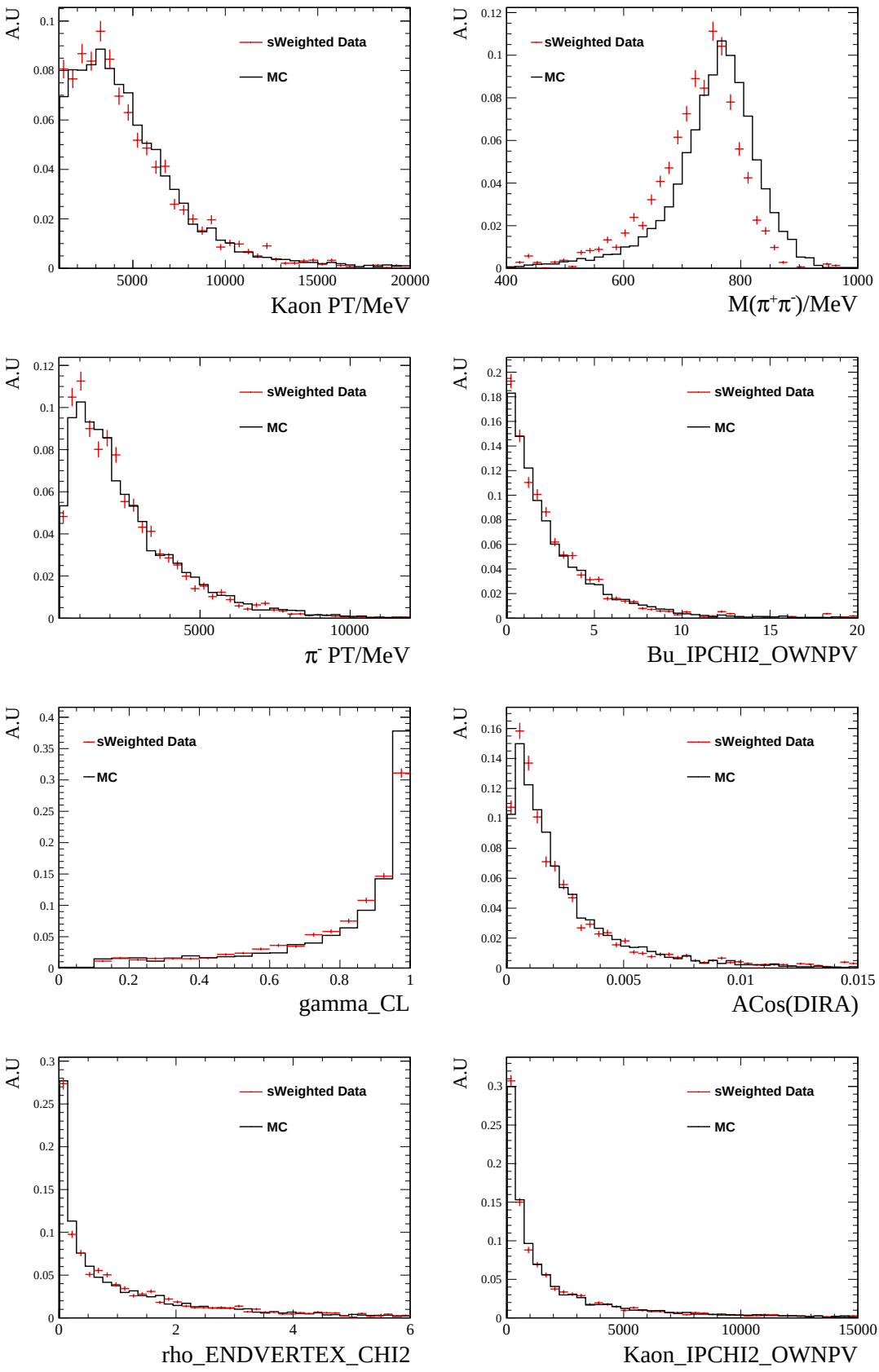


Figure A.1: Comparison between 2011 background subtracted control channel data and MC for variables used in the pre-selection.

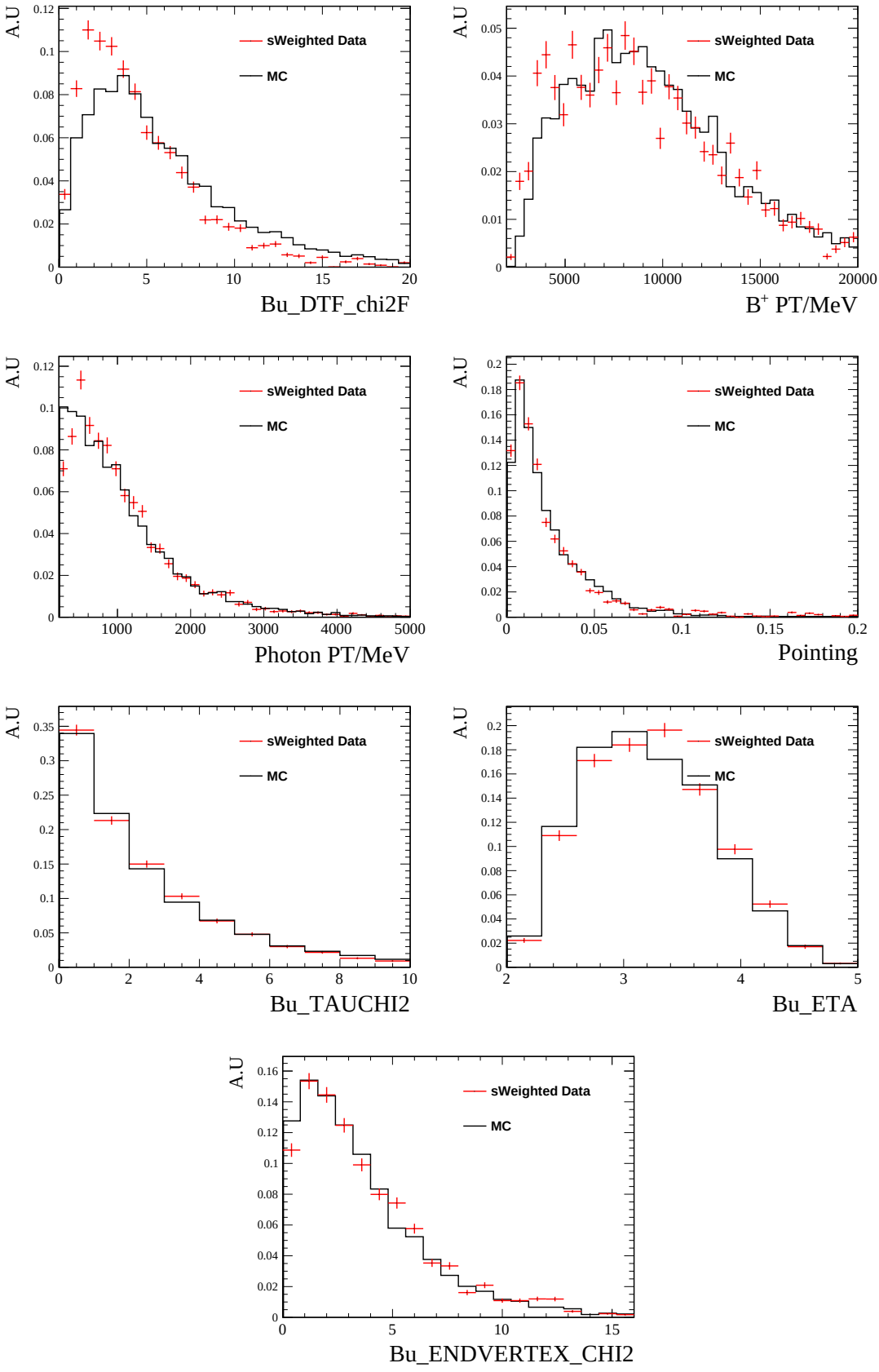


Figure A.2: Comparison between 2011 background subtracted control channel data and MC for variables used in the BDT.

APPENDIX B

Square Dalitz Plot Efficiencies

As mentioned in Section 4.3, the phase space of a three body decay can be binned in the Square Dalitz Plot (SDP) variables m' and θ'. However, these variables are not used in this analysis because they lead to low bin occupancies near the edges of the SDP. This appendix shows the rare channel efficiencies as a function of the SDP variables, which offers an alternative parameterisation of the phase space of the $\Lambda_b^0 \to pK\eta'$ decay. All of the efficiency maps shown are for 2011 and 2012 data combined.

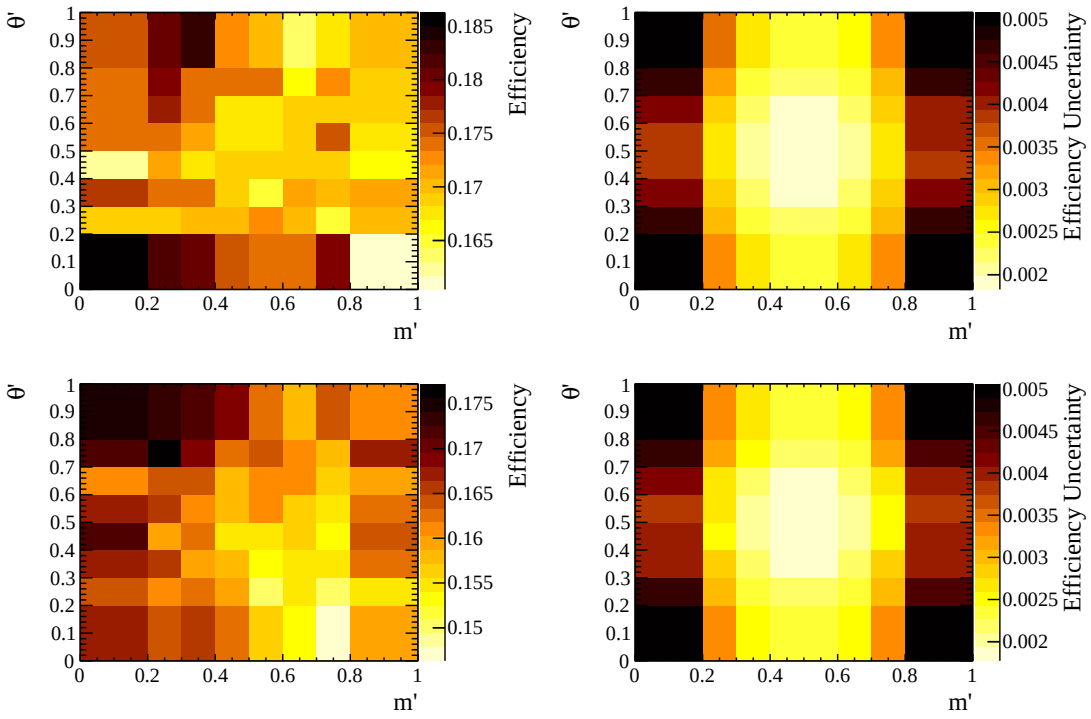


Figure B.1: Top(Bottom): Geometrical efficiencies and corresponding uncertainties of `DaughtersInLHCb` cut as a function of SDP position for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

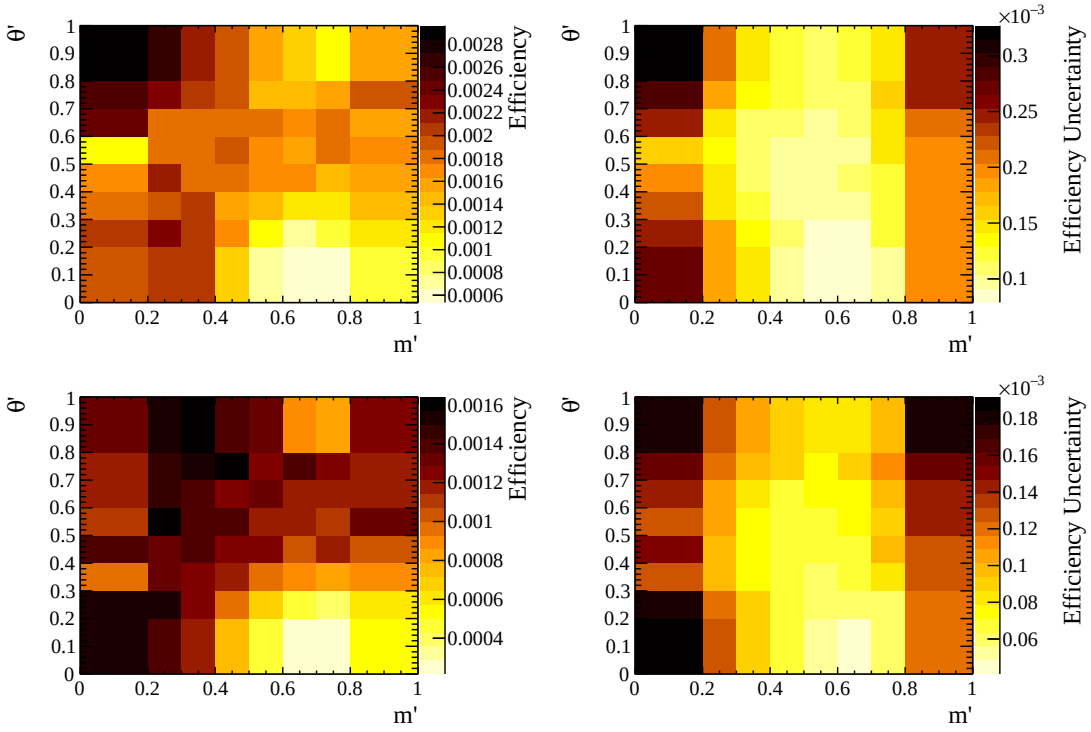


Figure B.2: Top(Bottom): Selection efficiencies and corresponding uncertainties as a function of SDP position for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

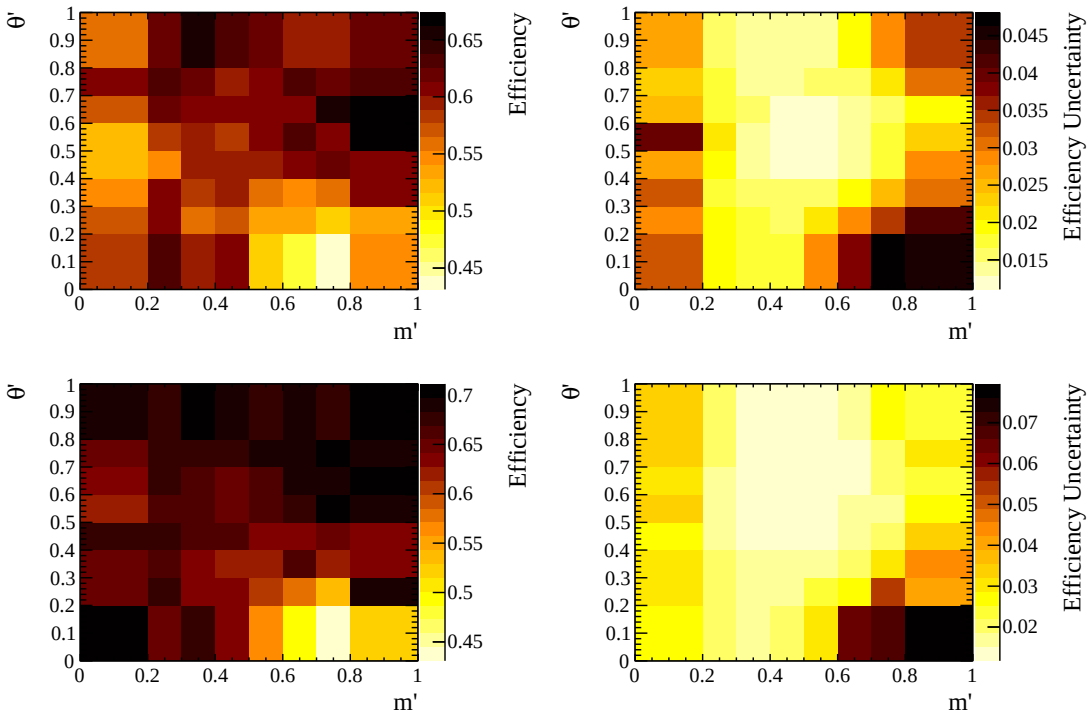


Figure B.3: Top (Bottom): PID Efficiencies and their uncertainties as a function of SDP position for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel.

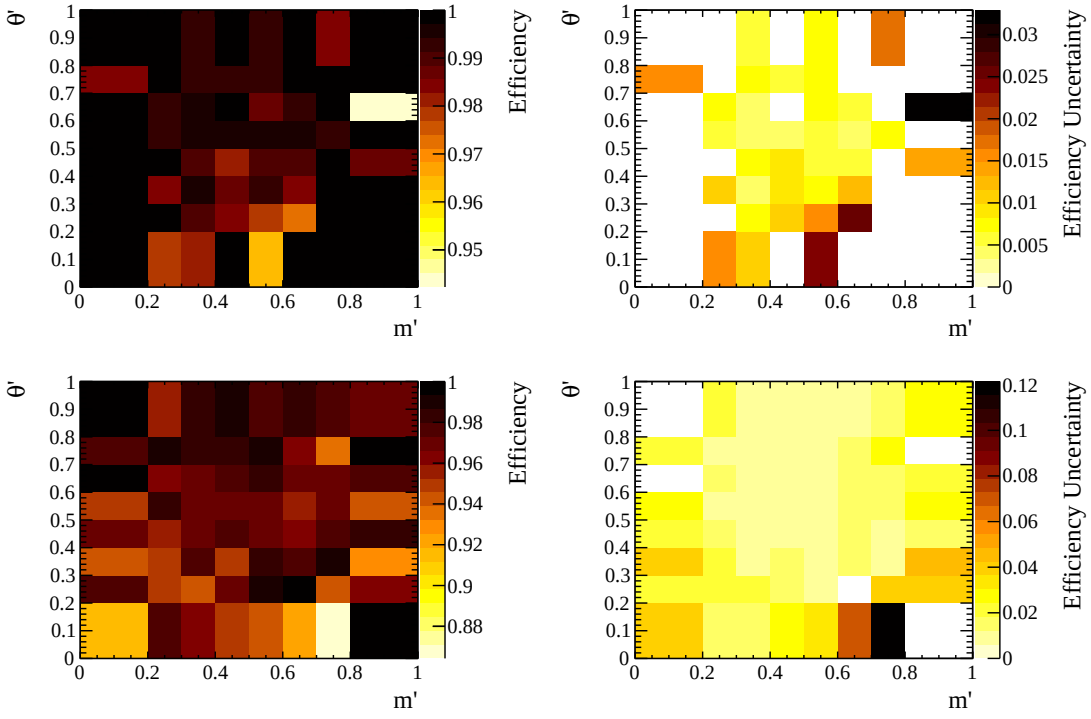


Figure B.4: Top (Bottom): Efficiency of retaining one candidate per event at random as a function of SDP position for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel and uncertainties.

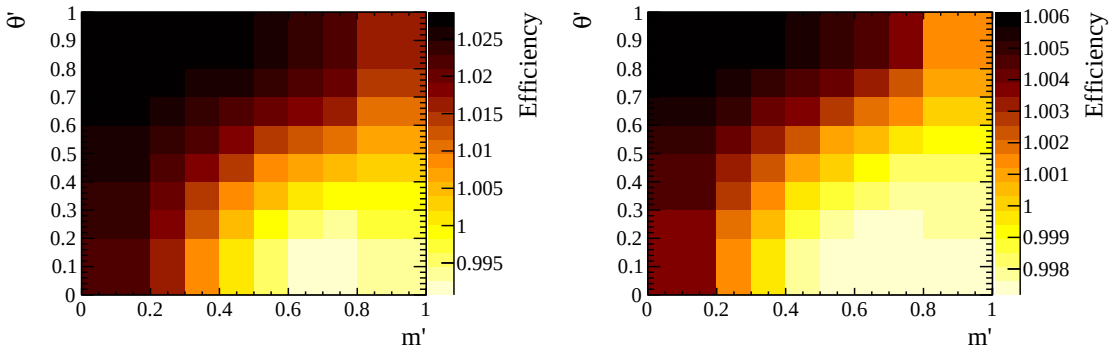


Figure B.5: Left (Right): The photon efficiency correction factors for the $\eta' \to \pi^+\pi^-\eta$ ($\eta' \to \pi^+\pi^-\gamma$) rare channel. The calculation of these correction factors is discussed in Section 4.3.5.

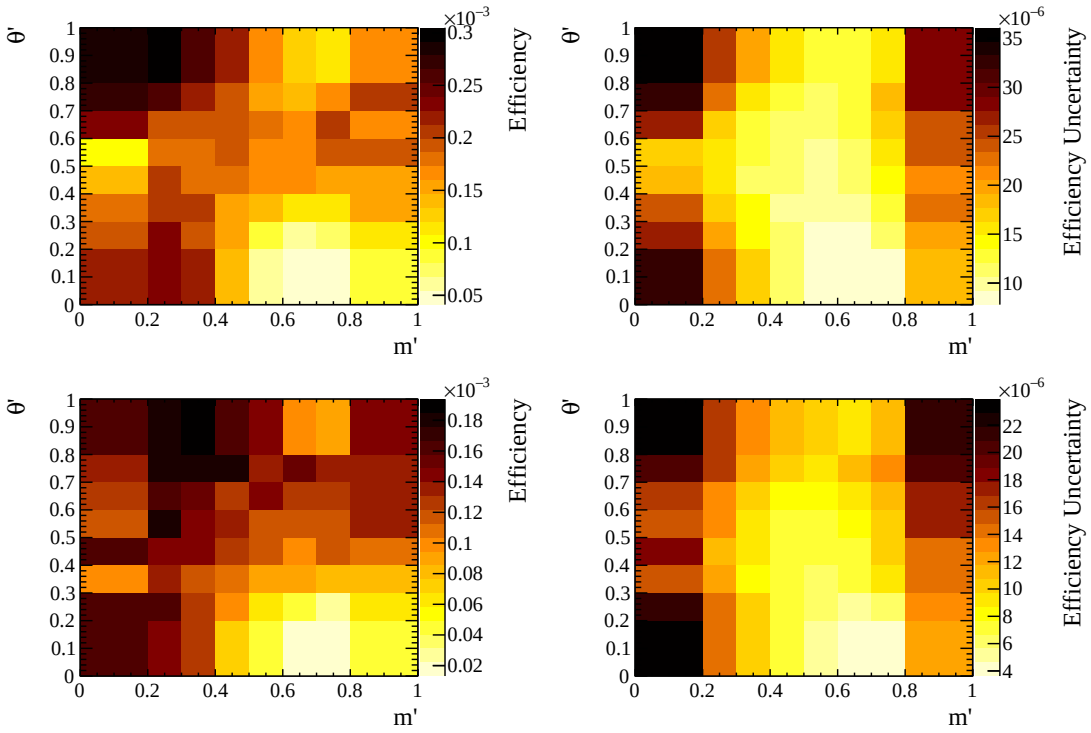


Figure B.6: Top (Bottom): Total efficiency for all stages of the selection combined for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel along with the total uncertainties.

APPENDIX C

Efficiency maps for 2011 data

This appendix shows the rare channel efficiencies as a function of m'' and $\cos(\theta_{\eta' p})$ for 2011 data. These are used in the branching fraction measurement. The equivalent plots for 2012 data can be found in Section 4.3.

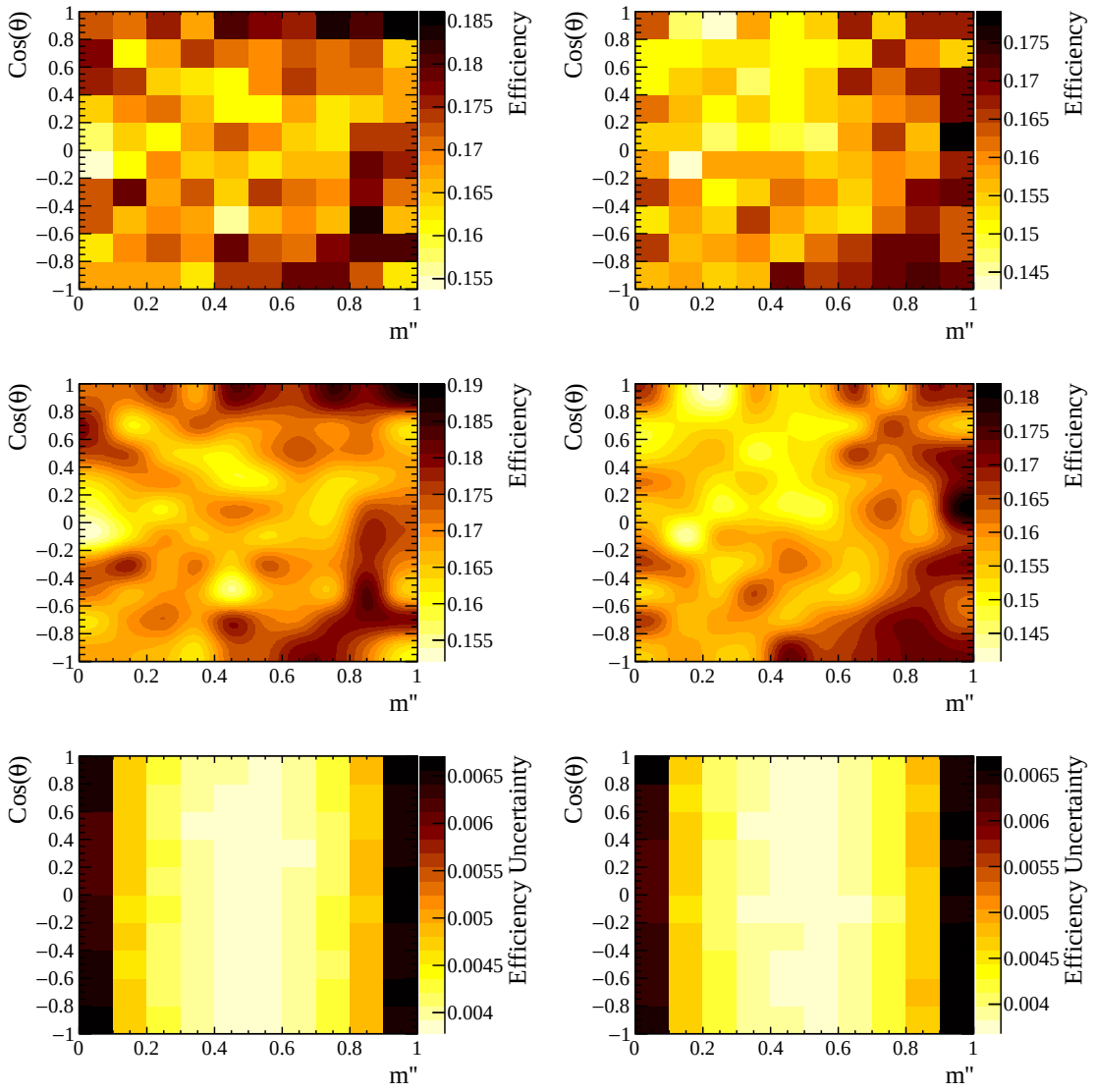


Figure C.1: Left(Right): Geometric efficiencies (top), with the results of the cubic spline interpolation(middle) and statistical uncertainties(bottom) for the `DaughtersInLHCb` cut as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2011 data. The uncertainties are calculated using the Clopper-Pearson method [119].

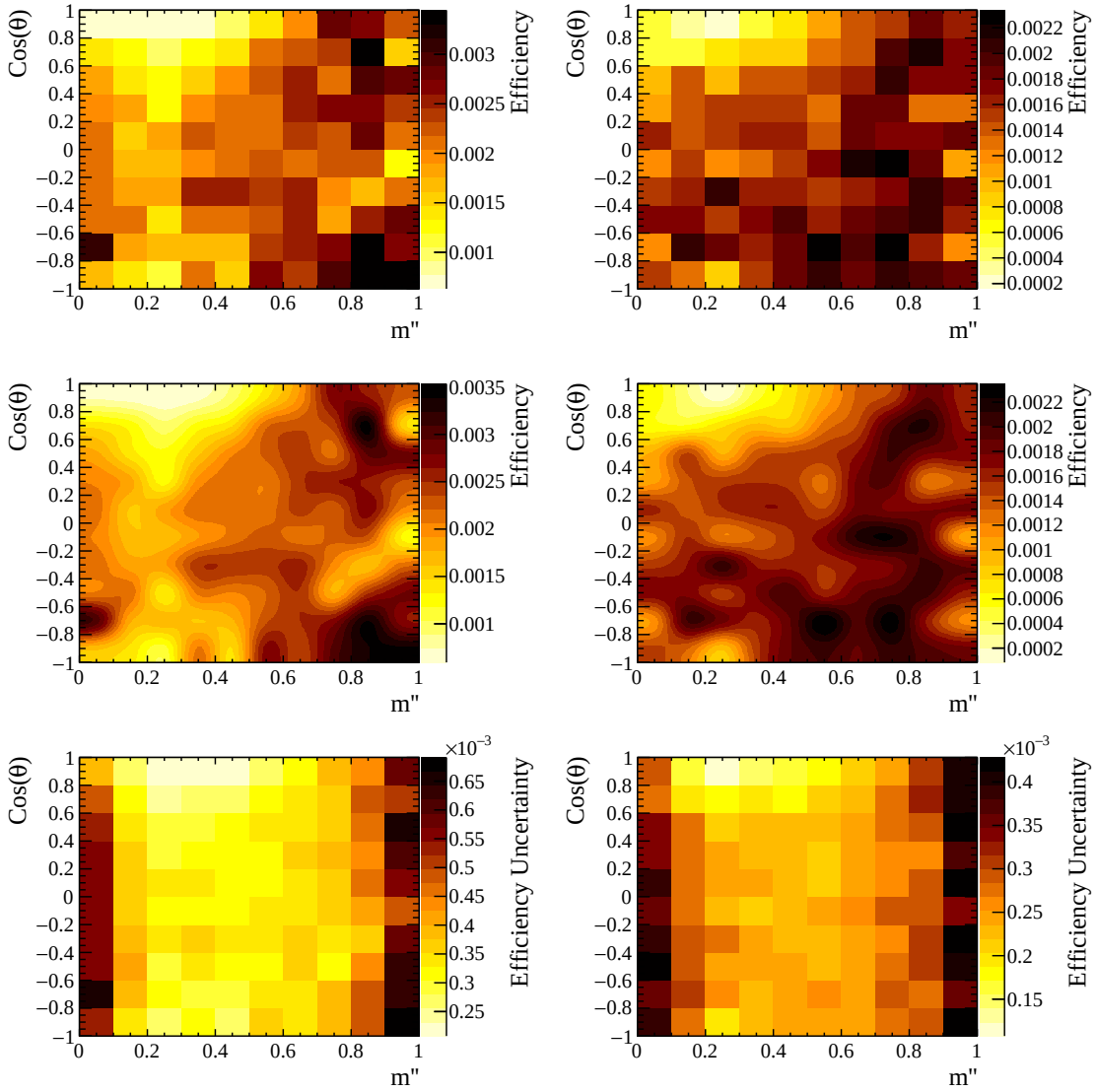


Figure C.2: Left(Right): Selection Efficiencies (Top) along with the results of the cubic spline interpolation (middle) and uncertainties(bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2011.

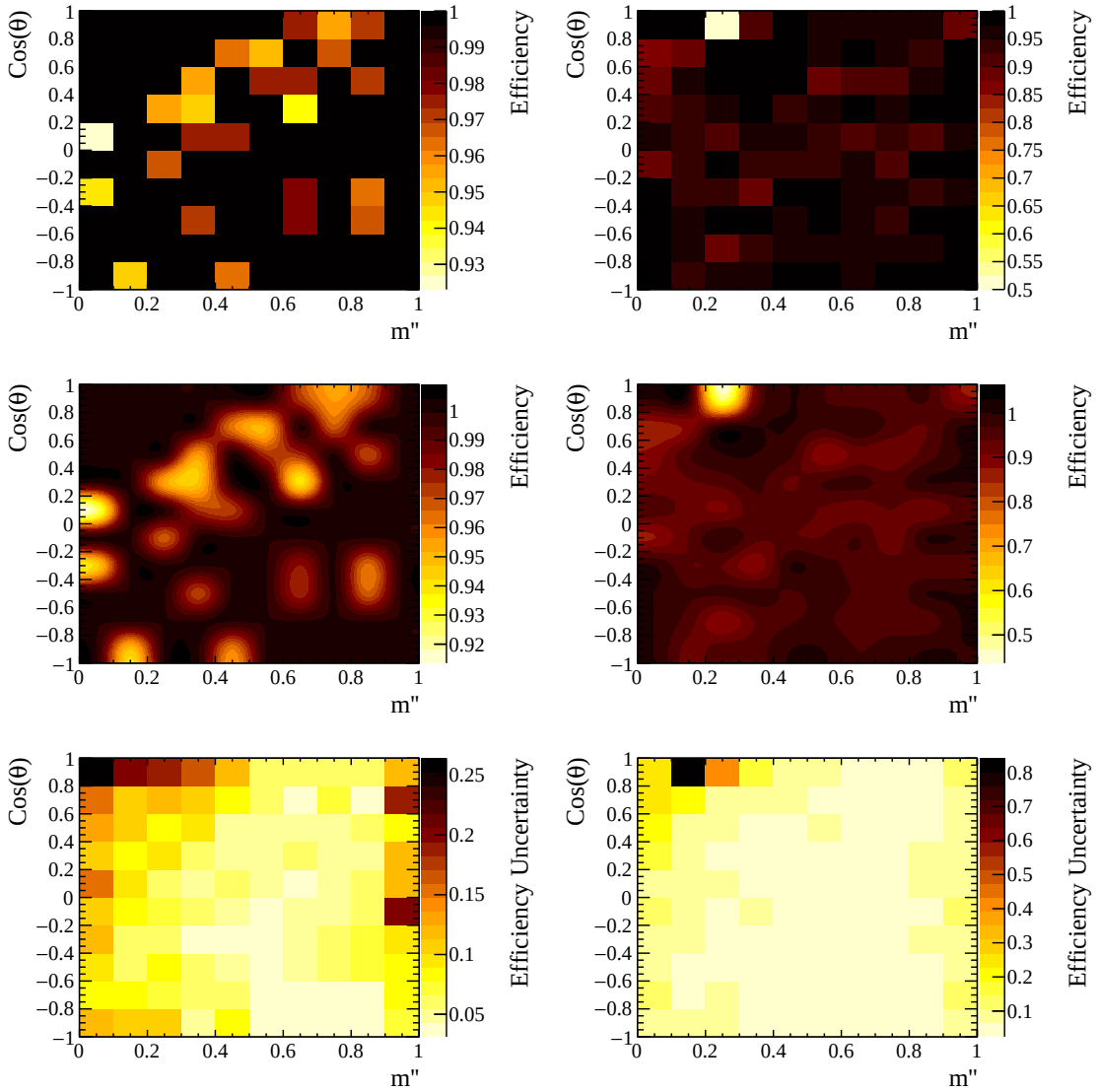


Figure C.3: Left(Right): Multiple candidate efficiencies (Top) along with the results of the cubic spline interpolation (middle) and total uncertainties(bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2011 data.

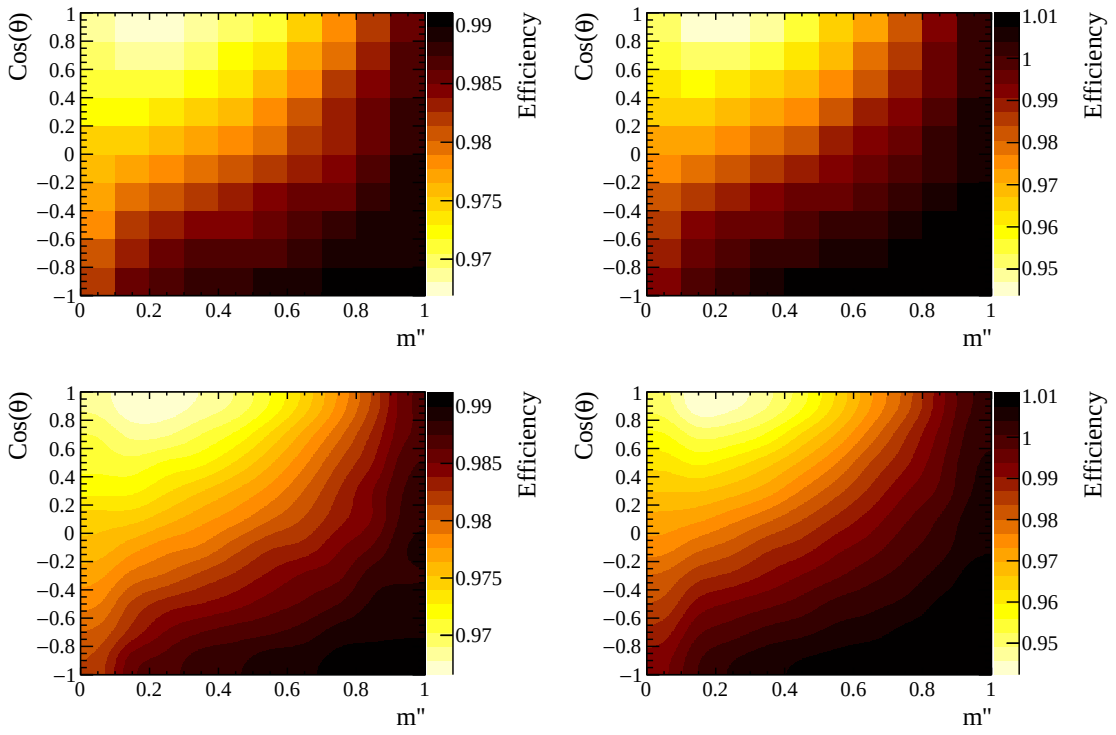


Figure C.4: Left(Right): Photon reconstruction efficiency correction (Top) along with the results of the cubic spline interpolation (bottom) as a function of m'' and $\cos(\theta_{\eta' p})$ for the $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) channel in 2011.

APPENDIX D

Branching Fraction Results By Decay Channel

As a cross check of the final branching fraction result, the branching fraction of $\Lambda_b^0 \to pK\eta'$ is calculated separately for each decay channel. The individual branching fractions are given by:

$$\mathcal{B}(\Lambda_b^0 \to pK\eta'(\eta' \to \pi^+\pi^-\gamma)) = \frac{N_\gamma}{N_c}\frac{\varepsilon_c}{\varepsilon_\gamma}\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_\gamma \times \mathcal{B}(B^+ \to K^+\eta'), \tag{D.1}$$

$$\mathcal{B}(\Lambda_b^0 \to pK\eta'(\eta' \to \pi^+\pi^-\eta)) = \frac{N_\eta}{N_c}\frac{\varepsilon_c}{\varepsilon_\eta}\left(\frac{f_u}{f_{\Lambda_b^0}}\right)_\eta \frac{\mathcal{B}_\gamma}{\mathcal{B}_\eta} \times \mathcal{B}(B^+ \to K^+\eta'), \tag{D.2}$$

where N is the signal yield, ε is the efficiency for a given channel and the subscripts c,γ and η refer to the control, $\eta' \to \pi^+\pi^-\gamma$ and $\eta' \to \pi^+\pi^-\eta$ channels respectively. $\left(\frac{f_u}{f_{\Lambda_b^0}}\right)$ are the fragmentation fractions and $\mathcal{B}_\gamma(\mathcal{B}_\eta)$ is the branching fraction of $\eta' \to \pi^+\pi^-\gamma$ ($\eta' \to \pi^+\pi^-\eta$) taken from the PDG.

By using the values of signal yields given in Table 5.6, efficiencies given in Table 5.10, and fragmentation fractions given in Table 4.1 the measurements of the branching fraction separated by channel are:

$$\mathcal{B}(\Lambda_b^0 \to pK\eta'(\eta' \to \pi^+\pi^-\gamma)) = (9.0 \pm 1.1(\text{stat.}) \pm 1.2(\text{sys.})) \times 10^{-6}, \tag{D.3}$$

$$\mathcal{B}(\Lambda_b^0 \to pK\eta'(\eta' \to \pi^+\pi^-\eta)) = (7.5 \pm 1.4(\text{stat.}) \pm 1.0(\text{sys.})) \times 10^{-6}. \tag{D.4}$$

These two measurements are consistent, which gives confidence in the final results. The systematic uncertainties are calculated by propagating the uncertainties on the

yields given in Table 5.7, on the efficiency ratios given in Table 5.13 and on the fragmentation fractions given in Table 3.1.

It is also possible to calculate the ratio of $\Lambda_b^0 \to pK\eta'$ BF results directly as

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\eta'(\eta' \to \pi^+\pi^-\gamma))}{\mathcal{B}(\Lambda_b^0 \to pK\eta'(\eta' \to \pi^+\pi^-\eta))} = \frac{N_\gamma}{N_\eta}\frac{\varepsilon_\eta}{\varepsilon_\gamma}\frac{\mathcal{B}_\eta}{\mathcal{B}_\gamma} = 1.19 \pm 0.26(\text{stat.}), \tag{D.5}$$

which gives further confidence that the results from the separate decay channels are consistent.

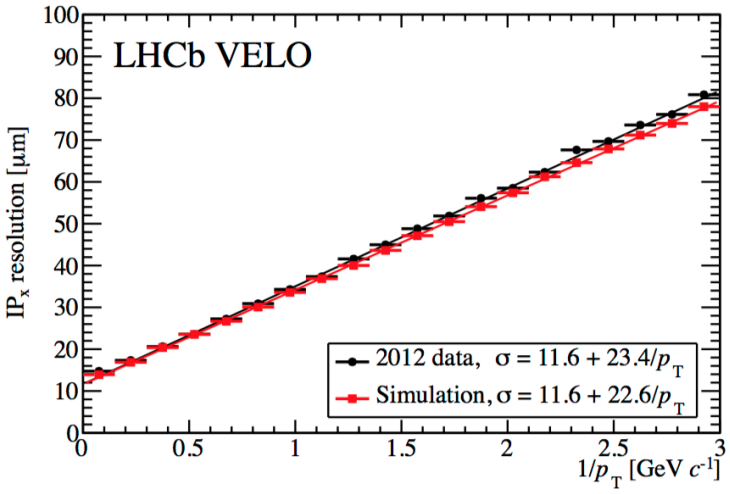


Figure 2.8: Impact parameter resolution of the VELO, as measured using prompt tracks in 2012 data and simulation [78].

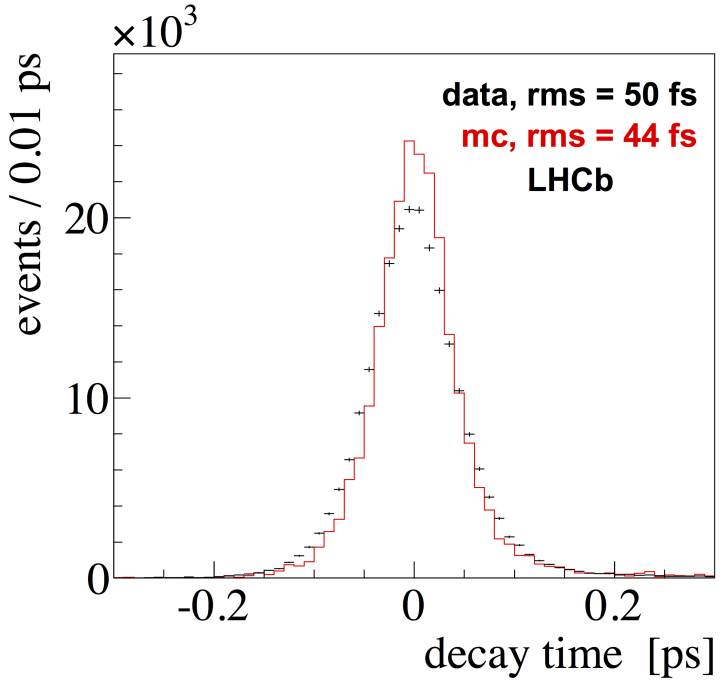


Figure 2.9: Decay time resolution of the VELO. This is measured using prompt events which mimic $B_s^0 \to J/\psi\,\phi \to \mu^+\mu^- K^+ K^-$ decays in 2011 data and simulation [78].

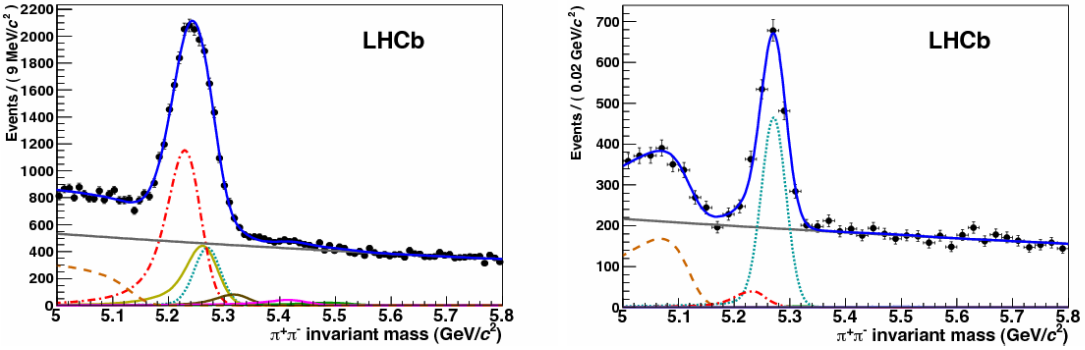


Figure 2.15: A comparison of the $B \to h^+h^-$ invariant mass distribution without (left) and with (right) the use of PID information. The signal decay under study, $B^0 \to \pi^+\pi^-$, is represented by the turquoise dotted line. Background contributions from: $B^0 \to K^+\pi^-$ *red dashed-dotted line*; $B^0 \to 3-$body *orange dashed-dotted line*; $B_s^0 \to K^+K^-$ *yellow/gold solid line*; $B_s^0 \to K^+\pi^-$ *brown line*; $\Lambda_b^0 \to pK^-$ *purple line* and $\Lambda_b^0 \to p\pi^-$ *green line* are also shown. The full fit function is represented by the *solid blue line* [81].

decays such as $B \to \mu^+\mu^-$ and lepton universality tests such as $\mathrm{R}(K^{*0})$ [85–87]. The LHCb muon system is composed of five detector stations (M1-M5). As can be seen in Figure 2.6, the first muon station (M1) is positioned upstream of the calorimeters whilst stations M2-M5 are positioned downstream of the calorimeters. The muon station positioned upstream of the calorimeter improves the resolution of the p_{T} measurement used in the trigger by $\approx 25\% - 35\%$[4] [76].

With the exception of the inner part of M1, each muon station is equipped with 276 Multi-Wire Proportional Chambers (MWPC). The inner region of M1 uses 12 triple Gas Electron Multiplier (GEM) gaseous detectors because this region receives the highest radiation dose, and it could not be guaranteed that MWPCs in this region would survive the required 10 years of operation. The muon stations are separated into logical regions in order to provide spatial information. Each muon station is separated by 80cm of iron, meaning that only muons with $p > 6\,\mathrm{GeV}$ will penetrate all five muon stations.

[4] M1 is not used in the offline reconstruction of muons because tracks from M2-M5 can be extrapolated back to the T1-T3 tracking stations which provides a more precise momentum measurement

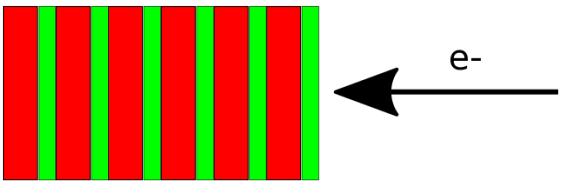


Figure 3.1: The geometry of the model calorimeter simulated by the calorimeter test. The red represents active plastic scintillator layers and the green represents passive lead layers.

The scenario simulated by the calorimeter test is electrons fired into the front face of a model calorimeter at normal incidence. The model calorimeter is a sampling calorimeter consisting of 66 alternating layers of both lead and plastic scintillator; the lead layers are 2 mm thick and the scintillator layers are 4 mm thick. A diagram showing this geometry is given in Figure 3.1. This geometry is deliberately chosen to model the design of the LHCb ECAL as closely as possible [73].

The main aim of this test is to use calorimeter resolution as a metric for comparison. The fractional resolution of an electromagnetic sampling calorimeter can be parameterised, in the absence of electronic noise, as

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} \oplus C, \tag{3.2}$$

where $\frac{\sigma}{E}$ is the fractional resolution of the calorimeter, A and C are free parameters of the model and $\oplus$ represents addition in quadrature. By determining the values of A and C for this model calorimeter, quantitative comparisons of different GEANT4 versions can be made.

The A term in Equation (3.2) arises from statistical fluctuations in the electromagnetic shower induced by the sampling calorimeter. In an electromagnetic shower, many particles are produced and the energy measured by the calorimeter is the sum of the energies deposited by each particle. However in a sampling calorimeter only a fraction of the shower takes place in active regions, therefore only a fraction of the shower particles are actually measured. Consequently the number of shower

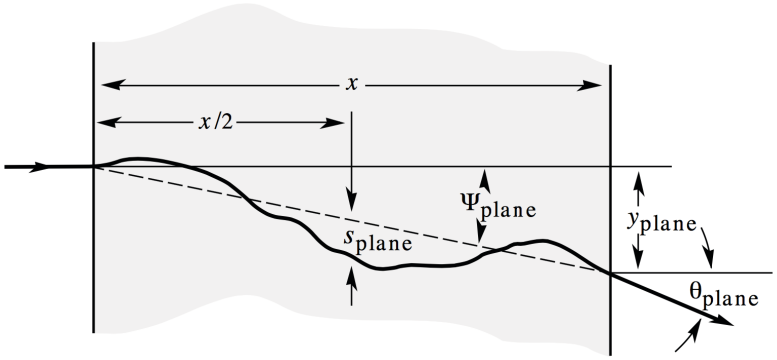


Figure 3.4: The lateral displacement and angular dispersion when a charged particle traverses a medium [1].

of the charged particle and θ is the angle through which the charged particle is scattered [99].

Except for cases where the scattering material is a very thin film, the charged particle will scatter multiple times before exiting the material. Hence, multiple coulomb scattering, which is more commonly known as just multiple scattering, occurs. The net effect is a lateral displacement as well as a scattering angle, as depicted in Figure 3.4. In this case a statistical treatment has to be used to obtain a distribution for the scattering angle, which is defined as θ in Figure 3.4. One such statistical treatment is Molière theory, which has been shown to give very good agreement with data over a wide range of particles, materials and energies [100, 101]. Several other theories have been shown to produce consistent results; Lewis theory also provides moments for the spatial displacement distribution [102]. Both the Moliere and Lewis theories give a scattering angle distribution that is Gaussian for the central 98% of scattering angle values, but the tails of the distribution fall off more slowly than a Gaussian function due to the $1/\sin^4(\theta/2)$ term in Equation 3.3. The width of the central Gaussian is defined as θ_0 which can be approximated by the Highland

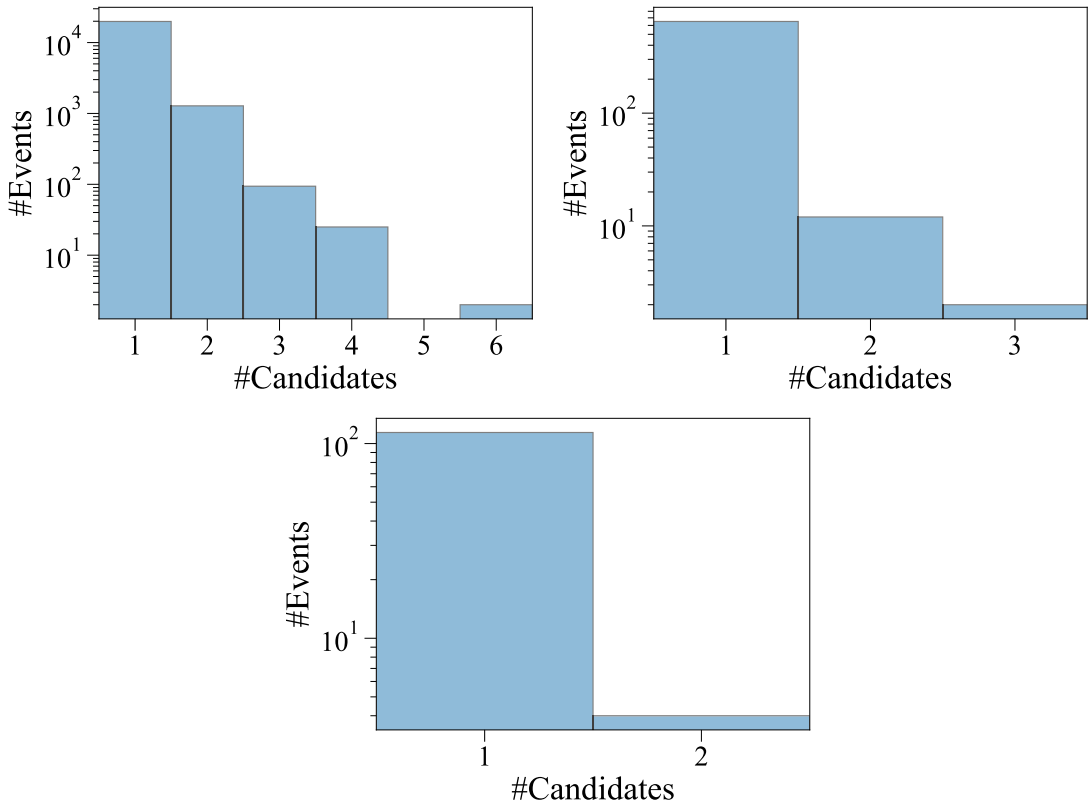


Figure 4.18: Number of candidates per event, for events passing the full selection in the Control channel (Top Left), $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$ channel (Top Right) and $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ channel (Bottom).

candidate. In the $\eta' \to \pi^+\pi^-\eta$ channel the low number of multiple candidates makes it difficult to conclude whether any particular area of the phase space is preferentially populated.

Table 4.13: The fraction of events that contain >1 candidate after all previous stages of the selection in each channel.

Channel	Data	MC
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$(6.61 \pm 0.18)\%$	$(5.33 \pm 0.21)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$(2.02 \pm 0.56)\%$	$(1.15 \pm 0.12)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$	$(4.31 \pm 1.94)\%$	$(5.40 \pm 0.22)\%$

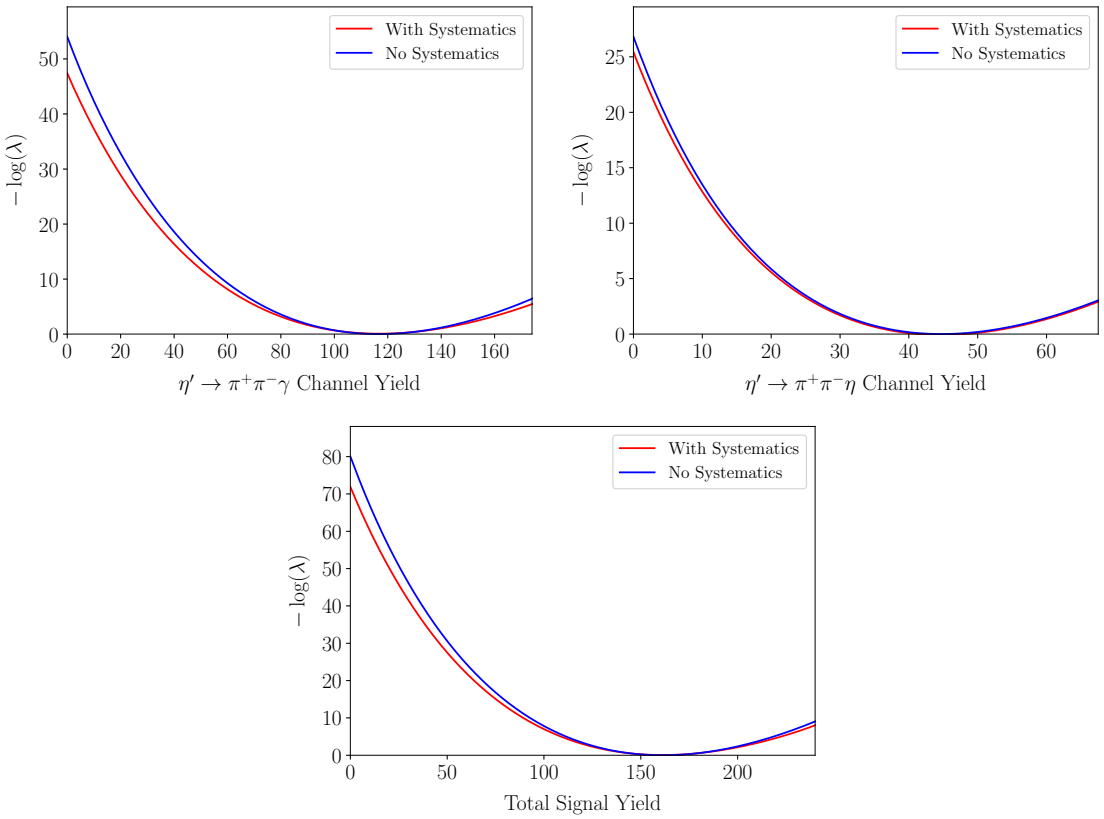


Figure 5.15: Scans of $-\ln\lambda$ for the two rare channel yields and the total signal yield both with and without systematic uncertainties included. The point at which the likelihood function crosses the yield=0 axis is equal to $\sigma^2/2$

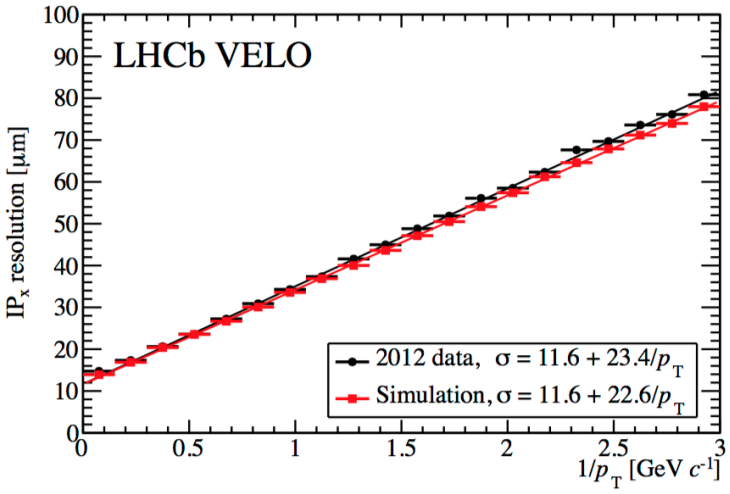


Figure 2.8: Impact parameter resolution of the VELO, as measured using prompt tracks in 2012 data and simulation [78].

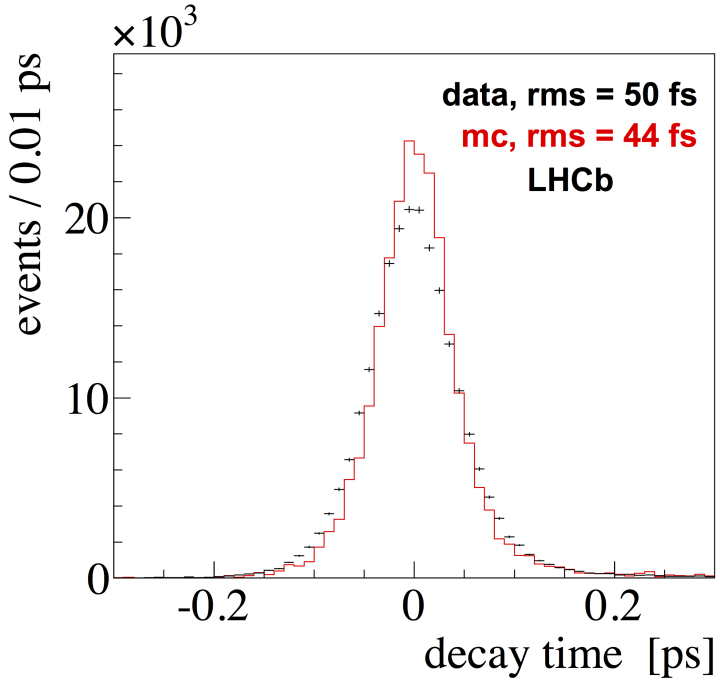


Figure 2.9: Decay time resolution of the VELO. This is measured using prompt events which mimic $B_s^0 \to J/\psi\,\phi \to \mu^+\mu^- K^+ K^-$ decays in 2011 data and simulation [78].

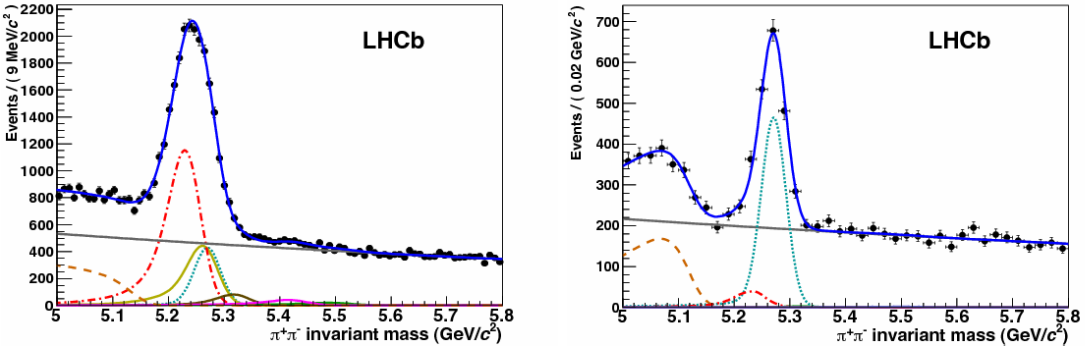


Figure 2.15: A comparison of the $B \to h^+h^-$ invariant mass distribution without (left) and with (right) the use of PID information. The signal decay under study, $B^0 \to \pi^+\pi^-$, is represented by the turquoise dotted line. Background contributions from: $B^0 \to K^+\pi^-$ *red dashed-dotted line*; $B^0 \to 3-$body *orange dashed-dotted line*; $B_s^0 \to K^+K^-$ *yellow/gold solid line*; $B_s^0 \to K^+\pi^-$ *brown line*; $\Lambda_b^0 \to pK^-$ *purple line* and $\Lambda_b^0 \to p\pi^-$ *green line* are also shown. The full fit function is represented by the *solid blue line* [81].

decays such as $B \to \mu^+\mu^-$ and lepton universality tests such as $\mathrm{R}(K^{*0})$ [85–87]. The LHCb muon system is composed of five detector stations (M1-M5). As can be seen in Figure 2.6, the first muon station (M1) is positioned upstream of the calorimeters whilst stations M2-M5 are positioned downstream of the calorimeters. The muon station positioned upstream of the calorimeter improves the resolution of the p_{T} measurement used in the trigger by $\approx 25\% - 35\%$[4] [76].

With the exception of the inner part of M1, each muon station is equipped with 276 Multi-Wire Proportional Chambers (MWPC). The inner region of M1 uses 12 triple Gas Electron Multiplier (GEM) gaseous detectors because this region receives the highest radiation dose, and it could not be guaranteed that MWPCs in this region would survive the required 10 years of operation. The muon stations are separated into logical regions in order to provide spatial information. Each muon station is separated by 80cm of iron, meaning that only muons with $p > 6\,\mathrm{GeV}$ will penetrate all five muon stations.

[4] M1 is not used in the offline reconstruction of muons because tracks from M2-M5 can be extrapolated back to the T1-T3 tracking stations which provides a more precise momentum measurement

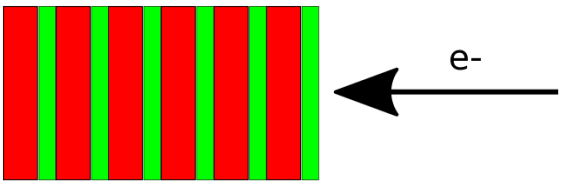


Figure 3.1: The geometry of the model calorimeter simulated by the calorimeter test. The red represents active plastic scintillator layers and the green represents passive lead layers.

The scenario simulated by the calorimeter test is electrons fired into the front face of a model calorimeter at normal incidence. The model calorimeter is a sampling calorimeter consisting of 66 alternating layers of both lead and plastic scintillator; the lead layers are 2 mm thick and the scintillator layers are 4 mm thick. A diagram showing this geometry is given in Figure 3.1. This geometry is deliberately chosen to model the design of the LHCb ECAL as closely as possible [73].

The main aim of this test is to use calorimeter resolution as a metric for comparison. The fractional resolution of an electromagnetic sampling calorimeter can be parameterised, in the absence of electronic noise, as

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} \oplus C, \tag{3.2}$$

where $\frac{\sigma}{E}$ is the fractional resolution of the calorimeter, A and C are free parameters of the model and $\oplus$ represents addition in quadrature. By determining the values of A and C for this model calorimeter, quantitative comparisons of different GEANT4 versions can be made.

The A term in Equation (3.2) arises from statistical fluctuations in the electromagnetic shower induced by the sampling calorimeter. In an electromagnetic shower, many particles are produced and the energy measured by the calorimeter is the sum of the energies deposited by each particle. However in a sampling calorimeter only a fraction of the shower takes place in active regions, therefore only a fraction of the shower particles are actually measured. Consequently the number of shower

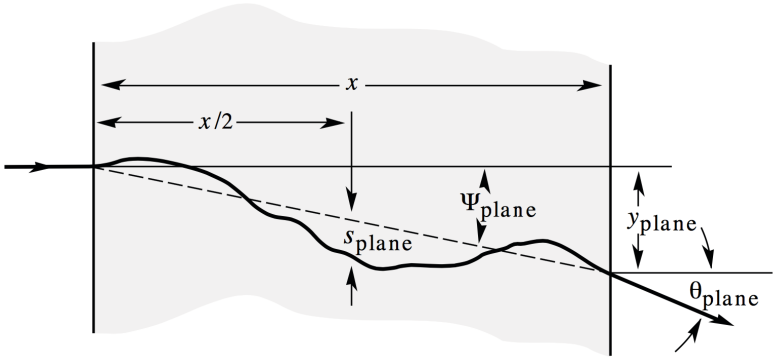


Figure 3.4: The lateral displacement and angular dispersion when a charged particle traverses a medium [1].

of the charged particle and θ is the angle through which the charged particle is scattered [99].

Except for cases where the scattering material is a very thin film, the charged particle will scatter multiple times before exiting the material. Hence, multiple coulomb scattering, which is more commonly known as just multiple scattering, occurs. The net effect is a lateral displacement as well as a scattering angle, as depicted in Figure 3.4. In this case a statistical treatment has to be used to obtain a distribution for the scattering angle, which is defined as θ in Figure 3.4. One such statistical treatment is Molière theory, which has been shown to give very good agreement with data over a wide range of particles, materials and energies [100, 101]. Several other theories have been shown to produce consistent results; Lewis theory also provides moments for the spatial displacement distribution [102]. Both the Moliere and Lewis theories give a scattering angle distribution that is Gaussian for the central 98% of scattering angle values, but the tails of the distribution fall off more slowly than a Gaussian function due to the $1/\sin^4(\theta/2)$ term in Equation 3.3. The width of the central Gaussian is defined as θ_0 which can be approximated by the Highland

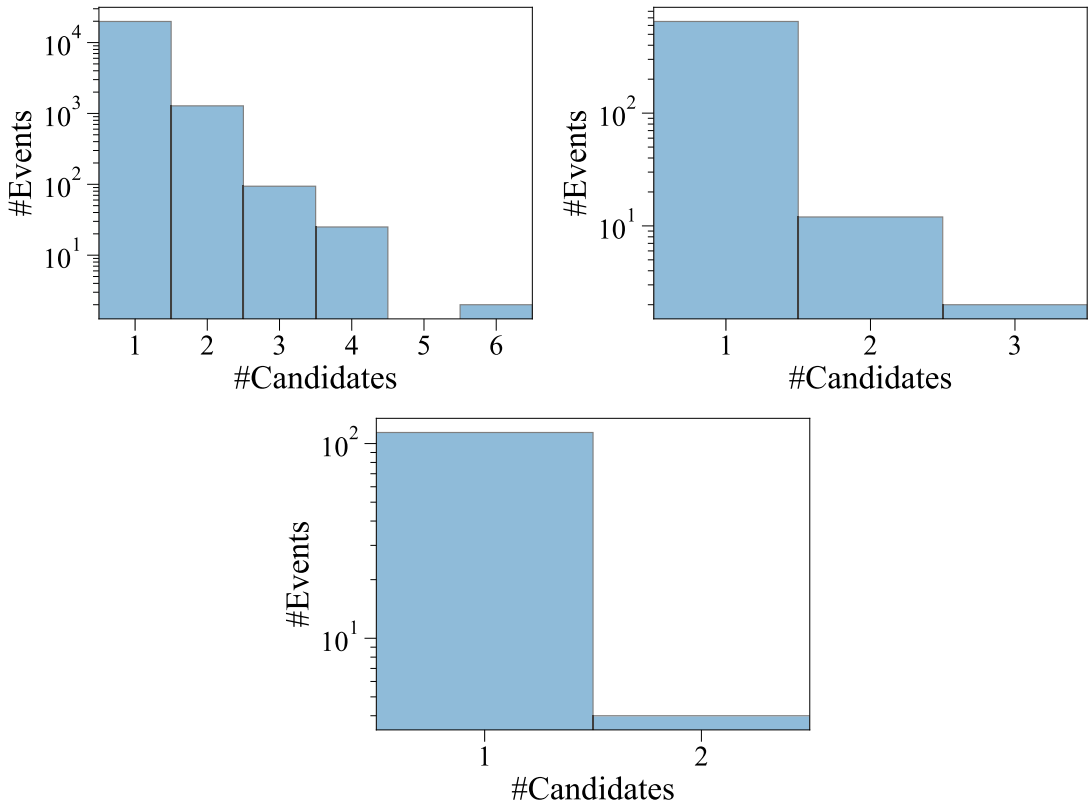


Figure 4.18: Number of candidates per event, for events passing the full selection in the Control channel (Top Left), $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$ channel (Top Right) and $\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$ channel (Bottom).

candidate. In the $\eta' \to \pi^+\pi^-\eta$ channel the low number of multiple candidates makes it difficult to conclude whether any particular area of the phase space is preferentially populated.

Table 4.13: The fraction of events that contain >1 candidate after all previous stages of the selection in each channel.

Channel	Data	MC
$B^+ \to K^+\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$(6.61 \pm 0.18)\%$	$(5.33 \pm 0.21)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\gamma\ )$	$(2.02 \pm 0.56)\%$	$(1.15 \pm 0.12)\%$
$\Lambda_b^0 \to pK\eta'$ $(\eta' \to \pi^+\pi^-\eta\ )$	$(4.31 \pm 1.94)\%$	$(5.40 \pm 0.22)\%$

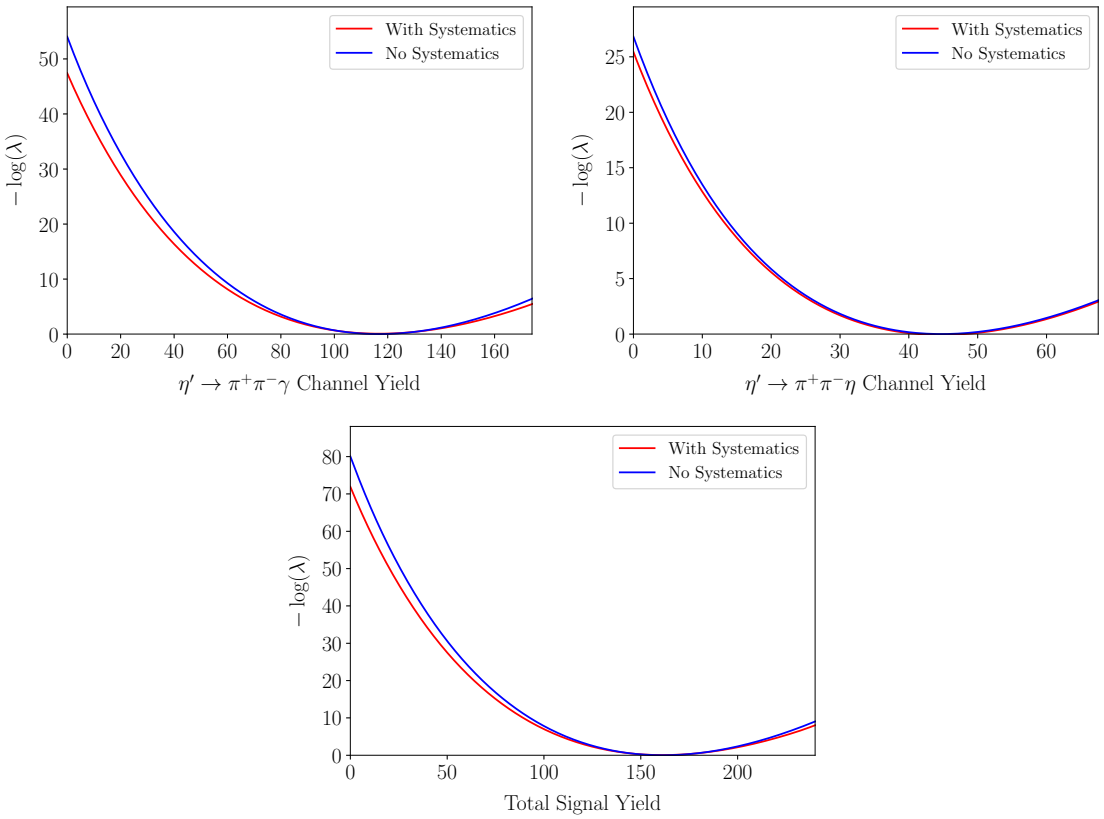


Figure 5.15: Scans of $-\ln\lambda$ for the two rare channel yields and the total signal yield both with and without systematic uncertainties included. The point at which the likelihood function crosses the yield=0 axis is equal to $\sigma^2/2$