First observation and branching fraction measurement of the rare decay $\Lambda_b^0 \to pK\mu^+\mu^-$ at the LHCb experiment, CERN

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY AT THE UNIVERSITY OF BIRMINGHAM

BY

PETER NOEL GRIFFITH

The University of Birmingham
Birmingham
United Kindgom

2016

UNIVERSITYOF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

THE DEPARTMENT OF PHYSICS AND ASTRONOMY

Abstract

The analysis for the first observation and branching fraction measurement of the rare decay $\Lambda_b^0 \to pK\mu^+\mu^-$ with respect to the control channel $\Lambda_b^0 \to J/\psi\,pK$ was performed. The analysis was kept blind in $\Lambda_b^0 \to pK\mu^+\mu^-$ to avoid potential selection bias as this decay is unobserved. The data is from proton-proton collisions recorded by the LHCb experiment in 2011 and 2012, corresponding to an integrated luminosity of $1\,\text{fb}^{-1}$ at $7\,\text{TeV}$ and $2\,\text{fb}^{-1}$ at $8\,\text{TeV}$ respectively. A corrected yield of $(7.83 \pm 0.13 \pm 0.65) \times 10^6$ $\Lambda_b^0 \to J/\psi\,pK$ candidates was measured for the full $3\,\text{fb}^{-1}$ and the analysis is currently under review by the LHCb Collaboration, awaiting approval to un-blind for the branching fraction measurement of $\Lambda_b^0 \to pK\mu^+\mu^-$.

Declaration of the author's contribution

The technical work in Section 2.3.10 and the analysis in Chapter 3 was carried out by myself.

The technical work involved developing backend and frontend software for the extraction and analysis of data for the validation of electromagnetic physics simulation at LHCb. The tools are designed for implementation into a validation and regression testing framework currently under development at LHCb.

All of the analysis is my own work. It is stated in the text where LHCb tools and software are used. The concept behind the kinematic weighting procedure and the machinery for the five dimensional efficiency modelling was developed in collaboration with Michal Kreps. Nigel Watson, Michal Kreps and members of the LHCb rare decays working group provided feedback and suggestions throughout the development of the analysis.

Acknowledgements

Firstly I would like to thank my fianceé Valentina for her dedication and patience throughout my PhD. This would not have been possible without her unwavering support.

My supervisor, Nigel Watson and collaborator Michal Kreps have both been a tremendous help in all aspects of my studies. Their advice and assistance in the development of techniques and ideas has been invaluable.

I would also like to acknowledge members of the rare decays working group and the LHCb Collaboration as a whole, who have given feedback and assistance on the analysis as it has developed, particularly Albert Puig and Mark Oliver.

Gloria Corti has been a superbly helpful in all things Gauss-related and I am grateful for her assistance in the development of the simulation validation tools for LHCb.

The members of office W313 at Birmingham have all been a great help. I would like to thank them along with my PhD colleagues with which I shared an invaluable two year attachment to CERN, Luca Pescatore, Mark Levy and Didier Alexandre.

A special thanks goes to my parents for their years of guidance and encouragement.

for Valentina

Contents

List of Figures

8

List of Tables

Chapter 1

Introduction

To date, the Standard Model (SM) of particle physics is the most successful attempt to describe the nature and behaviour of our universe. The 20th century discoveries of the strong and weak nuclear interactions completed our knowledge of the existence of at least four fundamental forces in nature: Electromagnetism, the weak nuclear force, the strong nuclear force and gravity.

The SM is a quantum field theory which attempts to describe three of these fundamental forces of nature. This effort began in 1961 with Glashow's unification of the weak and electromagnetic interactions [1]. Subsequent developments in the 60's and 70's lead to the Standard Model's current form, further corroborated by discoveries such as the existence of the top quark [2] and Higgs boson [3].

Despite the success of the Standard Model in explaining the interactions of matter on a subnuclear level, there are certain astronomical observations and particle physics anomalies for which there is no adequate explanation. For instance, there is no mechanism stong enough to describe the magnitude of the matter-antimatter asymmetry observed in the universe. We lack a fundamental explanation for the existence of neutrino oscillations or their finite masses, neither does the SM offer candidate particles of which dark matter

may be composed. Furthermore it does not explain the homogeneity and isotropy of the universe, such as a field to drive the proposed explanation for this: *cosmic inflation*. The attempts to develop a quantum description of gravity have yet to be fruitful, making it the only known fundamental force that is not included in the SM at all.

This thesis presents an experimental analysis that can directly probe the SM through the previously unobserved decay, $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ using data from the LHCb experiment at the Large Hadron Collider (LHC), CERN. This decay proceeds through a process known as a flavour changing neutral current (FCNC). Forbidden at tree level, decays such as these can be very sensitive to new physics, as new particles may modify the quantum loop contributions to the amplitude and hence cause observables, such as the rate at which the process occurs to deviate away from SM predictions. It is possible that studying these processes may uncover indirect evidence for physics beyond the Standard Model (BSM).

Efforts are ongoing to constrain and challenge the Standard Model in the fields of both astro and particle physics. The precise measurement of observables associated with these rare FCNC decays represent an invaluable means by which the predictions of the SM can be tested.

The Standard Model

The fundamental particles of the SM can be catagorised into quarks, leptons and gauge bosons. Quarks and leptons are spin $\frac{1}{2}$ fermions, and these particles are the constituents of matter. Interactions between fermions, through the EM, weak and strong forces are mediated by the spin-1 gauge bosons. The SM is a *Quantum Field Theory* (QFT) which describes particles and their interactions in terms of fields, as shown in Table 1.1 below.

The electroweak boson and gluon fields describe the spin-1 gauge bosons of the electroweak and strong interactions respectively. These gauge interactions are associated

Ψ	Fermion field
W_1, W_2, W_3, B	Electroweak boson fields
G_α	Gluon field
ϕ	Higgs field

Table 1.1: Quantum fields of the Standard Model.

with the symmetry group,

$$G = U(1) \times SU(2) \times SU(3) \tag{1.1}$$

The unitarity groups, $U(N)$ consist of $N \times N$ matrices (U groups have determinant $|1|$). In the case of of the SU groups, the S denotes 'special' meaning the determinant of the matrices is $+1$. The manifestation of gauge invariance gives rise to deep symmetries and conservation laws such as the conservation of electromagnetic charge.

Each group has $N^2 - 1$ generators, as $U(1)$, $SU(2)$ and $SU(3)$ act on their respective quantum fields to generate the gauge bosons of the standard model. $U(1)$ gives rise to a single massless boson, the photon, while $SU(3)$ acts on the gluon field producing eight types of gluon.

$SU(2)$ gives rise to three massless gauge bosons to mediate the weak force, however the $W^\pm$ and Z bosons are not massless. To reconcile this within the SM picture, the process of spontaneous symmetry breaking describes how particles, such as the massless gauge bosons of the weak interaction can acquire mass through interaction with a scalar field, hence the need for the scalar field ϕ, and Higgs particle.
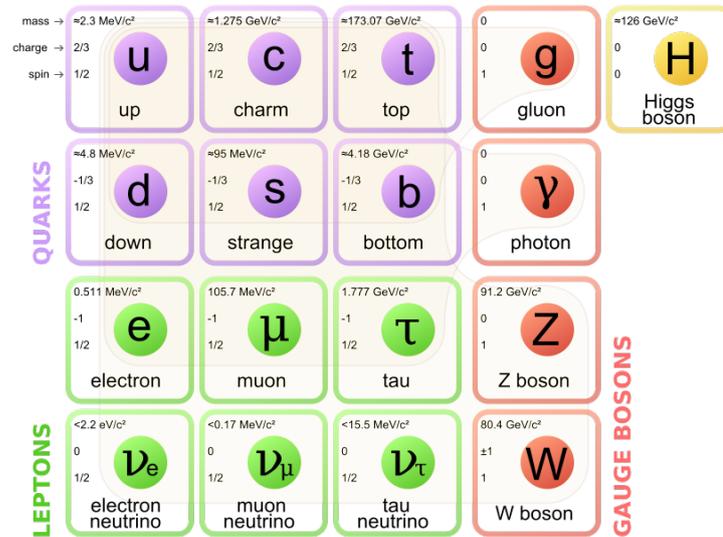
Figure 1.1: The fundamental particles of the Standard Model [4]

1.1 The electromagnetic interaction

The electromagnetic interaction, described in the SM by the theory of quantum electro-dynamics (QED) [5], concerns the interaction of particles with electromagnetic charge and is mediated by the photon. The elementary QED vertex shown in Figure 1.2 is the fundamental interaction from which all QED processes derive, for instance Møller scattering and electron-positron annihilation (Figure 1.3 and higher-order processes, for example those shown in Figure 1.4
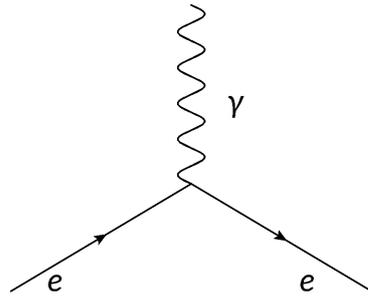
Figure 1.2: Elementary interaction vertex in QED where e is any charged particle in the SM.
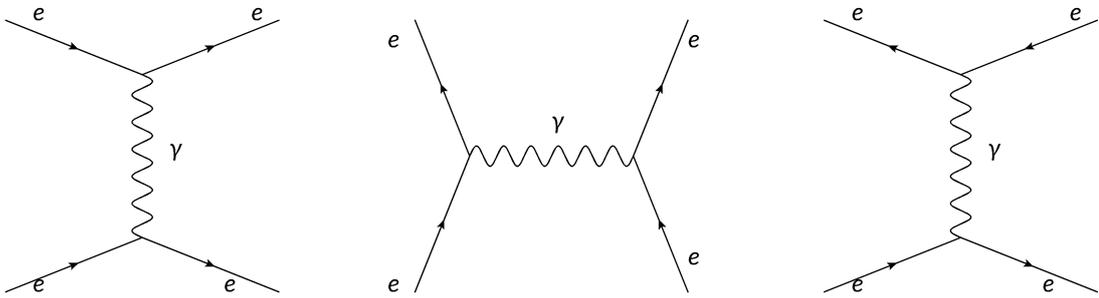
Figure 1.3: Tree level QED processes
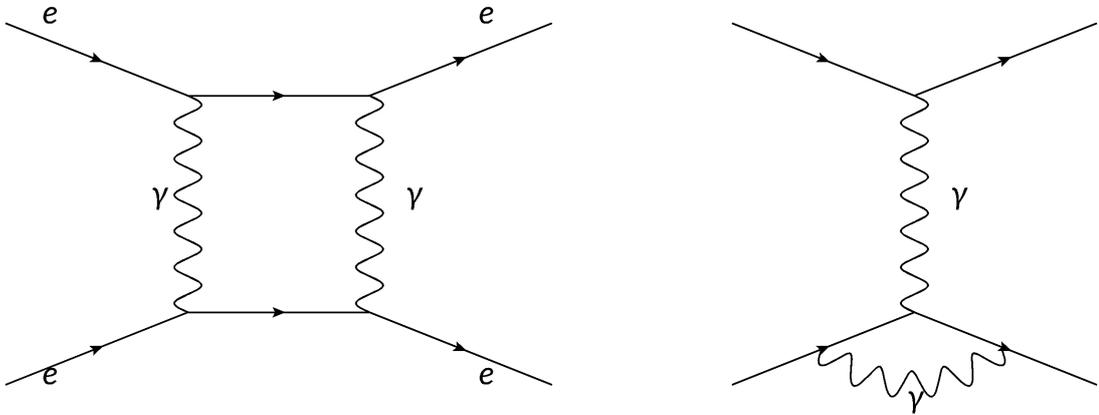
Figure 1.4: Higher order QED diagrams. (a) a box diagram. (b) a vertex correciton.

Feynman diagram calculations of QED processes are in principle straightforward with respect to QCD, because the inclusion of increasingly higher order diagrams or radiative photons perturb the calculation to a lesser degree since each QED vertex picks up an

additional factor of the fine structure constant, α, where[1],

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \tag{1.2}$$

meaning that QED calculations converge relatively quickly.

1.2   The strong interaction

The strong interaction, described in the SM by Quantum Chromodynamics (QCD) [5], is the interaction between particles that possess the property of *colour charge*, specifically gluons and quarks. The discovery of numerous hadronic states in the 1950's and 1960's preceded QCD and the notion of quarks themselves. In 1963 Gell-Mann and Zweig proposed the idea of hadrons being composed of three flavours of quarks, but the quantum number of colour charge was needed to explain how quarks could exist in a bound state with fully symmetric spin wavefunction and also a symmetric flavour wavefunctions, such as in baryons without violating the Pauli exclusion principle.

Unlike the electromagnetic charge in QED, in QCD there are three 'colours' (and their anti-colours) red, green and blue. This analogy to additive colours is used under the hypothesis that these bound states of quarks (mesons and baryons) must be colour singlets, i.e. 'colourless'[2]. For example, $q_r q_g q_b$ for a baryon and $q_r q_{\bar{r}}$ for a meson. Recent measurements from LHCb also show evidence of a pentaquark [6] state, where the color singlet state is achieved by quark combinations of the form $q\bar{q}q\bar{q}$ and $qqqq\bar{q}$ respectively.

The 'fundamental' QCD vertex shown in Figure 1.5 demonstrates how the gluon carries two colour degrees of freedom. Despite this interaction being quite similar to QED, the

---

[1]The fine structure constant is a running coupling, and varies with energy scale. The value of $\frac{1}{137}$ corresponds to the long distance (low energy) scale limit

[2]The notion of colour confinement is unambiguously observed but a formal proof that this must be the case has yet to be developed.
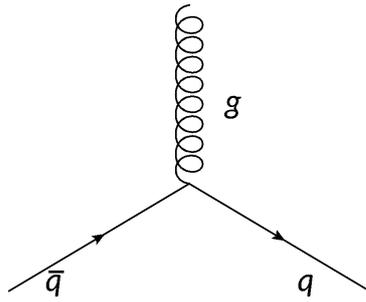
Figure 1.5: Fundamental interaction vertex in QCD, showing a specific colour exchange, however any combination of colour exchange can occur as long as it is conserved.

fact that gluons carry colour charge introduces a fundamental difference in that they will in general, self interact. This makes QCD calculations far more complex than QED. However, vacuum polarisation of virtual gluons has an anti-screening effect, (contrary to the screening effect that an electromagnetic charge experiences) causing qluons to interact weakly at higher energies (shorter distances). This means QCD is *asymptotically free* at higher energies.

1.3   The weak and electroweak interactions

The weak nuclear force is responsible for flavour change in the SM. It is in fact the only known mechanism in which flavour change can occur. There are two types of weak interaction: Charged currents, mediated by the $W^{\pm}$ bosons through which flavour change occurs (see Figure 1.6), and neutral currents via the Z boson.

Unlike the gluon and photon, the Z and W bosons are massive at $80.4 \, \mathrm{GeV}/c^2$ and $91.2 \, \mathrm{GeV}/c^2$ respectively [7]. Futhermore, the weak force does not conserve parity in constrast to the EM and strong interactions, due to the involvement of the axial vector current in weak interactions. EM and strong interactions involve the vector current V^{μ}, which under a parity inversion (a transformation of spatial coordinates from $\overline{x} \rightarrow -\overline{x}$ etc)
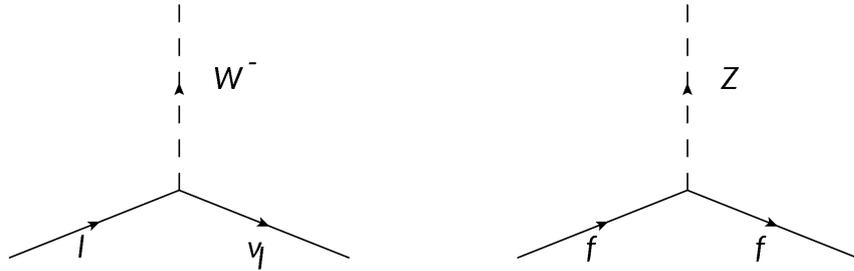
Figure 1.6: Charged and neutral currents in the weak interaction, showing quark and leption flavour change (left) and a neutral Z vertex (right) in which flavour remains unchanged.

becomes $-V^\mu$. Contrarily, the axial current A^μ is unchanged under parity inversion and the weak force is therefore a chiral gauge interaction in which only left handed fermions (and right handed anti-fermions) interact via the weak force. After violation of the parity symmetry was observed in 1957 by Wu [8], Landau proposed that the combined operation of both charge and parity symmetries (CP) as a symmetry group that would be respected by the weak interaction. However, it was only several years later, in 1964 that CP-violation was observed also in kaon decays [9]. Four years later the weak and electromagnetic forces were unified under the 'electroweak' force by Salam, Glashow and Weinberg [1]. This unification under a single theory of the electroweak interaction defines four massless bosons (W_1, W_2, W_3 and B, which through spontaneous symmetry breaking via the Higgs mechanism are transformed into the massive $W^\pm$ and Z bosons and the massless photon like so,

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \tag{1.3}$$

while,

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \pm iW_2) \tag{1.4}$$

therefore existing as two distinct forces, EM and weak, below the so-called *unification energy*, $\mathcal{O}(100\,\text{GeV})$. The angle θ_W is the weak mixing, or Weingberg angle and relates the coupling of the weak and EM forces as $e = g \sin\theta_W$, and the Z and $W^\pm$ boson masses

by $M_Z = \frac{M_W}{cos\theta_W}$ at lowest order [5].

1.4 Flavour physics

As shown in Figure 1.6, flavour change occurs through $W^{\pm}$ mediated interactions, but only flavour change within the same generation is shown in these Feynman diagrams, for instance $u \to d$ and $\mu \to \nu_\mu$. To date, no evidence has been seen for charged lepton flavour violation despite it being an active field for searches [10]. Experimental results demonstrating decays across generations of quarks are very well established however, such as the decay of the Λ baryon to a pion and proton. The Λ baryon constains a second generation valence quark (s), while the proton and pion and composed only of the first generation quarks u and d.

These transitions between quark generations, known as quark mixing, are explained via the Cabibbo-Kobayashi-Maskawa (CKM) matrix [11] which couples the weak eigenstates of the three quark generations to the mass eigenstates through a rotation, and describes the probability of transition bewtween two quark flavours.

The rotation of the mass eigenstates to the weak eigenstates via the CKM matrix is,

$$
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.5}
$$

where the left hand side primed components are the weak eigenstate doublets of down, strange and bottom type quarks, with the respective mass eigenstate doublets on the right. The CKM matrix in the middle (V_{CKM}) is a unitary complex matrix where the value of each $|V_{ij}|^2$ element represents the probability of the transistion $i \to j$ occurring. The

matrix are currently determined to be [7],

$$V_{CKM} = \begin{pmatrix} 0.97434^{+0.00011}_{-0.00012} & 0.22506 \pm 0.0005 & 0.00357 \pm 0.00015 \\ 0.22492 \pm 0.0005 & 0.97351 \pm 0.0001 & 0.0411 \pm 0.0013 \\ 0.00875^{+0.00032}_{-0.00033} & 0.0403 \pm 0.001 & 0.99915 \pm 0.00005 \end{pmatrix} \qquad (1.6)$$

where it is clear the elements of the leading diagonal, representing transistions within a single quark generation are close to unity, while off-diagonal elements are much smaller. Without the CKM matrix, the weak interaction would be 'universal' and the coupling would not depend on the quark generations involved, therefore 'universality' of weak interactions is broken.

As mentioned above, V_{CKM} is a complex unitary matrix and the presence of a complex phase allows for direct CP violation [12] as this gives a different between the rates for processes involving particles compared to anti-particles.

The representation above states only the magnitudes of the elements, and a widely used parametrisation of the CKM matrix is,

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \qquad (1.7)$$

where s and c denote sines and cosines of three Euler angles (representing rotations in 3D space) , $\theta_{12}, \theta_{13}, \theta_{23}$, and δ is the aforementioned complex phase. This leads to a very useful parametrisation introduced by Lincoln Wolfenstein in 1983 [13],
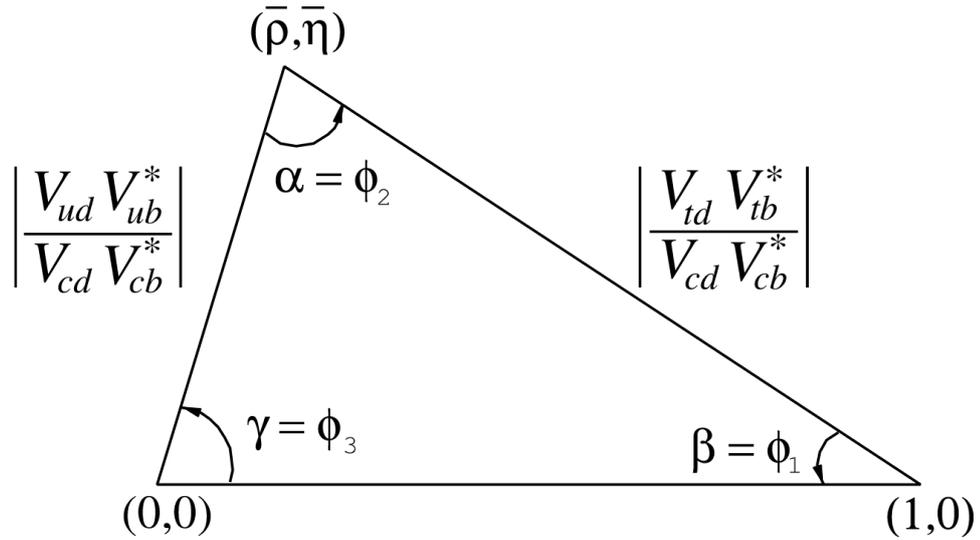
Figure 1.7: A CKM unitarity triangle [7] from Equation 1.9

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) + ..., \qquad (1.8)$$

where $\lambda = s_{12}$, $A\lambda^2 = s_{23}$ and $A\lambda^3(\rho - i\eta) = s_{13}e^{-i\delta}$, and therefore the complex phase, and all CP violation in the CKM is encoded in the parameter η. Untarity of V_{CKM} imposes that[3],

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \qquad (1.9)$$

which allows the construction of the CKM *unitarity triangle*, which re-scaled in the Wolfenstein parametrisation is as shown in Figure 1.7

The parameters of the unitarity triangle (the lengths of the sides, the angles and the area, which are related to the amount of CP violation in the SM) can be directly measured at particle physics experiments (see Figure 1.8). Examining whether the triangle closes by over contraining the values of the angles is an important test of the SM and is sensitive to

---

[3]Unitarity requires five other similar conditions that lead to representations as unitarity triangles but this is the most commonly used.
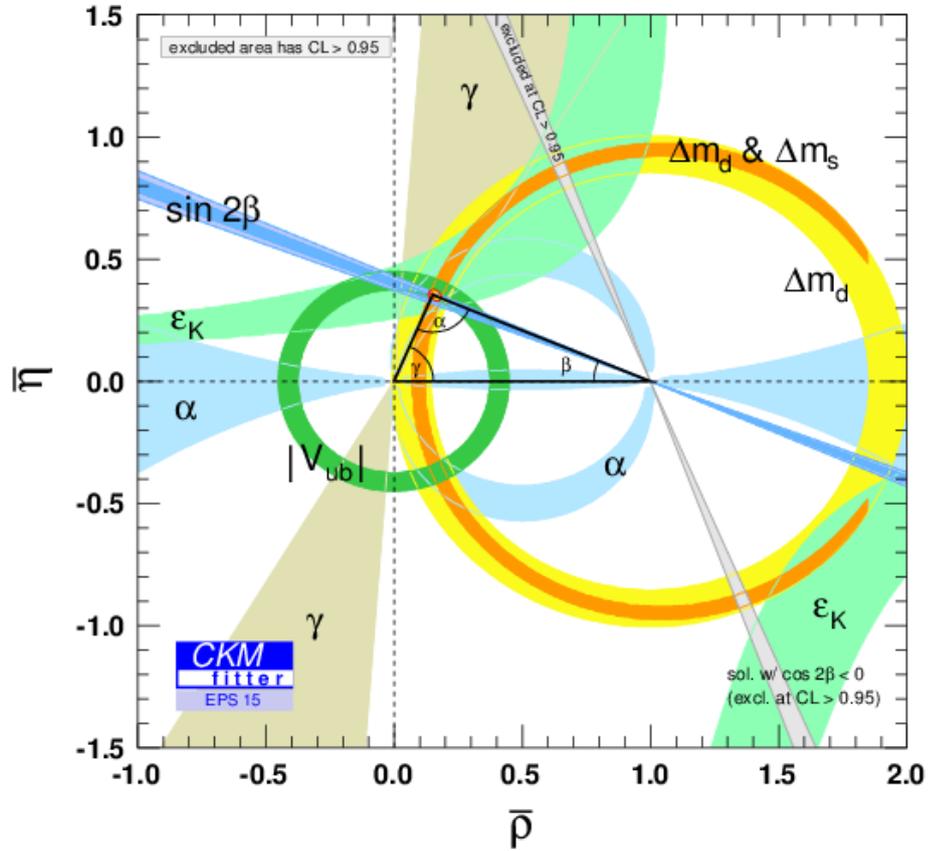
Figure 1.8: Latest constraints on the CKM unitarity triangle [16]

New Physics. Much of the experimental work carried out by the LHCb experiment relates directly to these parameters [14] [15].

As mentioned previously, one of the open puzzles in the SM is the matter-antimatter asymmetry. The universe initially had a baryon number $B = 0$ and within the first few picoseconds of the universe (before the quark epoch), through some mechanism (baryogenesis) a large asymmetry was generated. [17]. CP-violation provides a mechanism that would lead to asymmetry, and is in fact one of the 'Sakharov conditions' for baryogenesis [18]. However, if no BSM sources exist, currently the magnitude of CP violation within the SM seems to be insufficient in explaining the observed asymmetry [19], thus precision

measurements in the flavour sector are highly motivated.

1.5   Rare FCNC decays

At tree level, flavour change occurs through weak charged currents, however higher order interactions permit flavour change through loops, which would otherwise require a neutral current at tree level. These flavour changing neutral currents (FCNCs) are suppressed through the Glashow-Iliopoulos-Maiani (GIM) mechanism [20] and generally have a branching fraction of $< 10^{-6}$. The development of this theory led to the prediction of the existence of the charm quark [21], at a time when only u, d and s were known quarks.

Two types of FCNC diagrams are shown in Figure 1.9, the penguin and box diagrams. Here we see how the loops mediate the change in quark flavour from $b \rightarrow s$, and BSM particles may participate in these loops, either enhancing or suppressing through interference, the rate at which the process occurs. In fact, BSM particles such as extra vector-like quarks can permit FCNC decays to occur at tree level [22]. As the loop participants are off-shell, FCNC's are sensitive to NP contributions from particles with an on-shell mass far beyond that which could be produced directly at a given collider energy. Direct searches for BSM particles have not yet led to a discovery at the LHC, suggesting that the energy scale of NP may be beyond what the LHC can directly access [23], which further motivates indirect searches through rare decays.

To extract useful observables from the physics of $b \rightarrow s$ decays, a low energy effective theory can be constructed as m_Z and $m_{W^\pm}$ are much more massive than the b quark. The Operator Product Expansion (OPE) formalism is used to construct the effective Hamiltonian [24], in which the sum extends over all relevant operators and are weighted
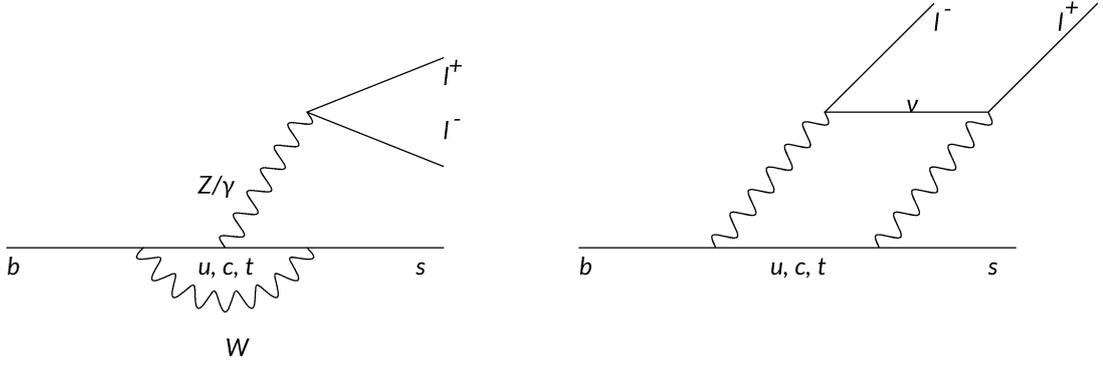
Figure 1.9: Penguin and W-box Feynman diagrams of the $b \to s$ flavour changing process.

by complex coefficients,

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}}V_{tb}V_{ts}^*\frac{\alpha_e}{4\pi}\sum_i C_i(\mu_s)\mathcal{O}(\mu_s), \qquad (1.10)$$

where G_F and α_e are the Fermi and fine structure constants respectively. The C_i are Wilson coefficients; complex numbers which describe the short distance physics, which are pertubative. The operators $\mathcal{O}_i$ contain the long distance contributions and μ_s is the renormalisation [25] scale for which C_i and $\mathcal{O}_i$ are evaluated. The relevant operators for radiative, leptonic and semi-leptonic decays are [26],

$$\mathcal{O}_7 = \frac{m_b}{e}\bar{s}\sigma^{\mu\nu}P_RbF_{\mu\nu}, \qquad\qquad \mathcal{O}_7' = \frac{m_b}{e}\bar{s}\sigma^{\mu\nu}P_LbF_{\mu\nu},$$

$$\mathcal{O}_9 = \bar{s}\gamma_\mu P_Lb\bar{l}\gamma^\mu l, \qquad\qquad \mathcal{O}_9' = \bar{s}\gamma_\mu P_Rb\bar{l}\gamma^\mu l,$$

$$\mathcal{O}_{10} = \bar{s}\gamma_\mu P_Lb\bar{l}\gamma^\mu\gamma_5 l, \qquad\qquad \mathcal{O}_{10}' = \bar{s}\gamma_\mu P_Rb\bar{l}\gamma^\mu\gamma_5 l,$$

where m_b is the mass of the b quark and $F_{\mu\nu}$ are the electromagnetic field strength tensors. $\mathcal{O}_i'$ represents the chirality flipped operators, obtained by replacing P_L with P_R, where $P_{R/L} = (1 \pm \gamma_5)/2$. Due to the left-handed nature of the weak interaction, the corresponding C_i' coefficients experience a suppression on the order of m_s/m_b. $\mathcal{O}_7$ describes

contributions to radiative and photon mediated decays, i.e. $b \to s\gamma$ and photon penguins, while $\mathcal{O}_{10}$ describes contributions to purely leptonic decays such as $B_s^0 \to \mu^+\mu^-$. As for semi-leptonic decays $(b \to sl^+l^-)$, they receive contributions[4] from $\mathcal{O}_7$, $\mathcal{O}_9$ and $\mathcal{O}_{10}$ and the particular contributions dominate depends on the kinematic regime of the decay (see Section 1.5.1).

In this effective theory formalism the Wilson coefficients can be calculated pertubatively, are model independent and are measurable observables in FCNC decays. In the SM case, the Wilson coefficients are determined by requiring the amplitudes from the full SM electroweak theory are consistent with those obtained by the effective theory, at the weak scale $(\mu_s = M_W)$. These solutions are then evolved to $\mu \approx m_b$, using the renormalisation group equations [27]. Any contributions from BSM particles will then be seen as a deviation from the SM prediction in the relevant Wilson coefficient(s) such that $C_i = C_i^{SM} + C_i^{NP}$ where NP denotes new physics. The coefficents can be tested through several experimental observables of FCNC decays, for instance the rate at which the process occurs (branching fraction) or the angular distributions of the decay.

1.5.1 $b \to sl^+l^-$ decays

As $b \to sl^+l^-$ decays receive contributions from several Wilson coefficients, they are powerful probes of the SM. The di-lepton mass squared (q^2) dependence of measurements such as the branching fraction are important to investigate because the relative importance of the contributions varies with the q^2 region under consideration. The q^2 distribution for a $b \to s\mu^+\mu^-$ decay can be separated into the following regimes, starting at lowest q^2:

- The photon pole: This region, at low q^2 is dominated by $\mathcal{O}_7$.

- $\mathcal{O}_7$-$\mathcal{O}_9$ interference: This region resides between the photon pole and the J/ψ

---

[4]The corresponding chirality flipped operators also contribute to these types of decays, albeit suppressed.
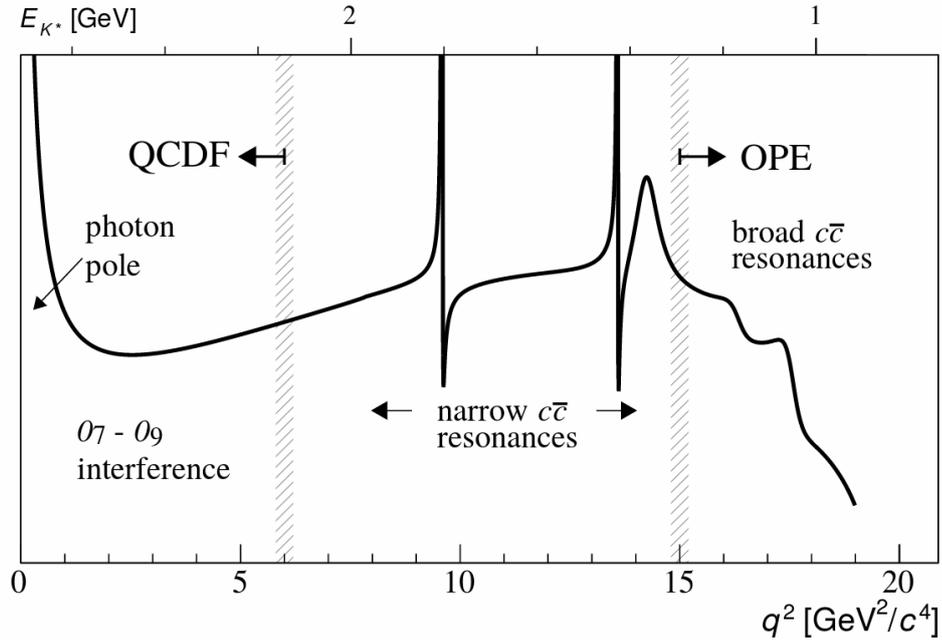
Figure 1.10: Representation of the q^2 differential decay rate for $B^0 \to K^{*0}\mu^+\mu^-$ [26].

charmonium resonance and is where $\mathcal{O}_7$-$\mathcal{O}_9$ interference dominates. This region is highly sensitive to new physics in the C_9 coefficient.

- Narrow $c\bar{c}$ resonances: This region contains the J/ψ and $\psi(2S)$ charmonium resonances, where the muons come from the charmonium decay. As this proceeds through tree level diagrams they are much higher rate than the non-resonant regimes of the spectrum.

- Broad $c\bar{c}$ resonances: At high q^2, above the open charm threshold, the production of broad charm resonances dominates and a local OPE method can be used for calculations.

Figure 1.10 depicts the differential decay rate in q^2 for $B^0 \to K^{*0}\mu^+\mu^-$ in which these regions are annotated.

The decay amplitude of such processes can be calculated using,

$$\langle f | \mathcal{H}_{eff} | i \rangle = \frac{G_F}{\sqrt{2}} \sum_i V_i C_i(\mu) \langle i | \mathcal{O}_i | f \rangle , \tag{1.11}$$

for a process of $i \rightarrow f$. The $\langle i | \mathcal{O}_i | f \rangle$ are the hadronic form factors, and as they are generated from the $\mathcal{O}_i$ operators they cannot be calculated perturbatively. Techniques such as lattice QCD [28] and light cone sum rules [29] (often covering different kinematic regimes) are commonly employed. There is strong motivation to reduce the uncertainties on the form factors from these calculations as they remain one of the dominant sources of uncertainties.

The decay $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ involves a $b \rightarrow s$ transition, and is expected to have similar structure in the q^2 differential decay rate, as shown in Figure 1.10.

Chapter 2

The Large Hadron Collider and the LHCb experiment

The Large Hadron Collider and LHCb experiment are amongst the largest, most complex and technically challenging scientific instruments ever built. This chapter briefly describes the design and operation of these two machines, with emphasis on the aspects that are relevant to the analysis in Chapter 3. Section 2.3.10 describes technical work undertaken for the simulation software of LHCb.

2.1 The accelerator complex
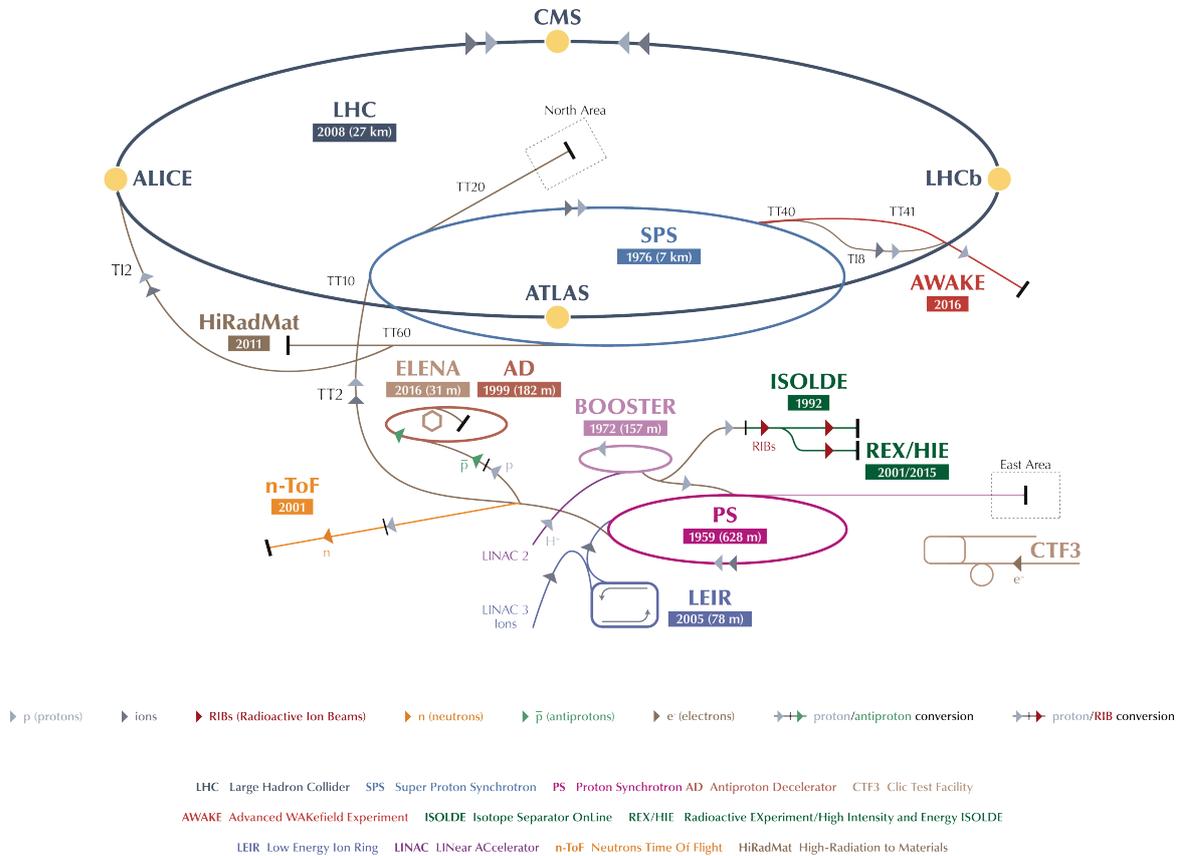
Figure 2.1: Diagram of the full accelerator complex at CERN. [30]

The Large Hadron Collider (LHC) [31] is located in Geneva, Switzerland at the CERN laboratory. Currently the LHC is the world's largest particle physics experiment, consisting of a ~27km storage ring, supported by the smaller accelerators; the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), the Superproton Synchrotron (SPS) and several linear accelerators.

The LHC collides beams of protons[1] at four interaction points along the 27km ring. Exploiting these interaction points are the four large experiments, ATLAS [32], CMS [33],

---

[1]Beams of lead ions as well as proton-lead ion collisions are also performed

Alice [34] and LHCb [35], along with smaller experiments TOTEM [36], LHCf [37] and MoEDAL [38].

A schematic view of CERN's full accelerator complex is shown in Figure 2.1.

The LHC is designed to collide protons at a centre of mass energy of 14 TeV. In 2011 the LHC operated at half the design energy, colliding protons at 7 TeV, while in 2012 this was increased to 8 TeV. The data collected by the LHCb experiment during this period of Run 1 are analysed in Section 3 for the first observation and branching fraction measurement of $\Lambda_b^0 \to pK\mu^+\mu^-$ decays.

2.1.1   Beam preparation and injection

The protons are initially acquired through the ionisation of hydrogen gas, after which they are accelerated to 50 MeV along the LINAC2 linear accelerator and injected into the PSB. The PSB ring consists of four vertically stacked beamlines of ~157m circumference. This simultaneous storage and acceleration of four proton beam injections from LINAC2 allows the 628m circumference PS to be filled from a single combined injection of the four PSB beams. The PS accepts the PSB beam at 1.6 GeV and accelerates them to 25 GeV before injection into the SPS. The ~7km long SPS ramps the beam to LHC injection energy of 450 GeV. The SPS nominally works on an injection cycle time on the order of a few tens of seconds, filling the LHC with proton bunch trains. Each bunch in the LHC has a population of $\sim 1.1 - 1.4 \times 10^{11}$ protons and the machine is filled with a maximum of 2808 proton bunches out of a possible 3564 configurations in the LHC. Design specifications give a bunch spacing of 25ns but during Run 1[2] the LHC ran with a nominal bunch spacing of 50ns and an increased bunch density [39].

---

[2]Run 1 of the LHC lasted between 2009 and 2013. For brevity, unless states otherwise, Run 1 will be used in this report to specifically mean the main proton-proton physics portion of Run 1 from 2011-2012

Maximum bunch population of the LHC is less than the total number of 'RF buckets' they can fill to account for the rise time in the beam dump's kicker magnet - leaving an "abort gap" of 3μs.

2.2 The Large Hadron Collider

The LHC consists of a ring of 1232 superconducting dipole magnets [40] that are designed to operate at a field strength of 8.33T to achieve the bending power required at the nominal beam energies of 7 TeV each. The design bunch crossing frequency, f_b is 40MHz, with a peak instantaneous luminosity of 1×10^{34}cm$^2 s^{-1}$. However, during Run 1, the LHC achieved peak stable luminosities of 1.26×10^{33}cm$^2 s^{-1}$ [41] and 7.7×10^{33}cm$^2 s^{-1}$ in 2011 and 2012 respectively[3]. These relatively high luminosities for the 7 and 8 TeV runs were achieved due to the higher proton bunch density.

2.3 The LHCb Spectrometer

The LHCb experiment, shown schematically in Figure 2.2 resides in France, at point 8 of the LHC. It is designed primarily to study CP-violation and rare decays in the beauty and charm sector. At the LHC, b and c quarks are produced primarily at low angles from the beam and in abundance. At 7 TeV, the charm production cross-section[4] $\sigma c\bar{c}_{p_{\mathrm{T}}<8\,\mathrm{GeV}/c,2.0<y<4.5}$ is measured to be $(1419 \pm 193)\,\mu$b [42], while the $\sigma(pp \to b\bar{b}X)$ cross section is $(284 \pm 69)\,\mu$b [43]

---

[3]As detailed in Section 2.3, these luminosity values are not the nominal luminosities at the interaction point of LHCb due to the lumi-levelling technique

[4]y denotes rapidity as defined in Equation 2.1
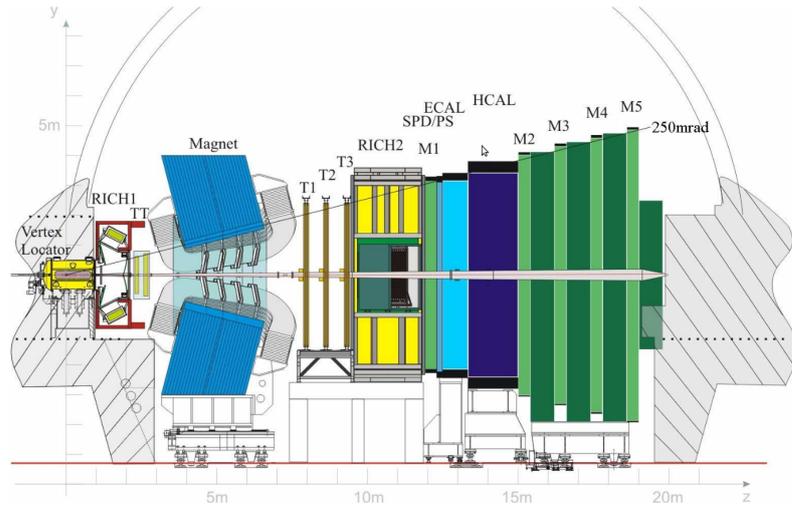
Figure 2.2: Schematic view of the LHCb detector [35]

LHCb is built as a single forward-arm spectrometer, covering a geometric acceptance of 10-400mrad from the beam, the reason for which is apparent in Figure 2.3, which shows Monte-Carlo simulation of $b\bar{b}$ production at LHCb. The red band in this figure represents the detector's geometric acceptance. This acceptance corresponds to a pseudorapidity of $2 < \eta < 5$.

In special relativity, rapidity, y, is the hyperbolic angle that relates two frames of reference in relative motion. In particle physics it is often preferable to express rapidity relative to the beam axis, as,

$$y = \frac{1}{2}ln\left(\frac{E + p_z}{E - p_z}\right),\tag{2.1}$$

where E is the particle's energy, and p_z is its momentum component parallel to the beam axis. In the high energy limit where $p \approx E$, the expression can be simplified to what is called pseudorapidity,

$$y \approx \eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right],\tag{2.2}$$

where θ is the angle between the particle and the beam axis.
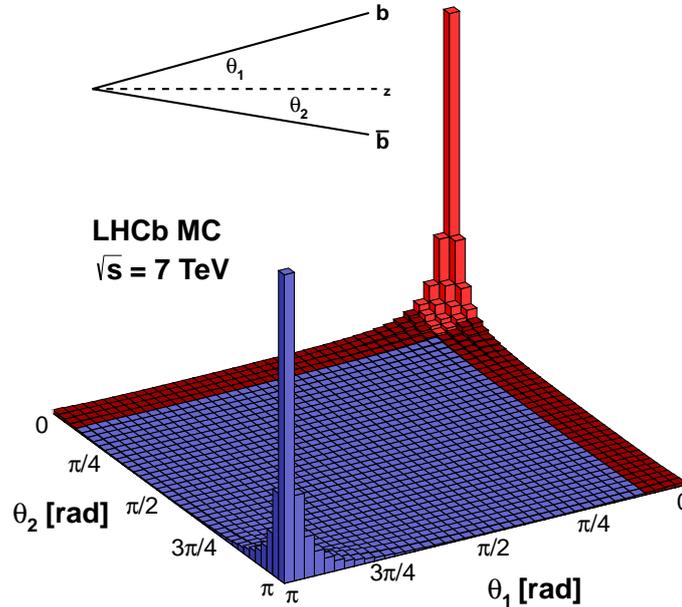
Figure 2.3: Monte-Carlo simulation of $b\bar{b}$ production at LHCb, with LHCb acceptance region in red. [44]

LHCb operates a luminosity-levelling technique using a controllable transverse beam offset [45]. This ensures a stable luminosity is achieved, rather than decaying throughout the physics run. At $4 \times 10^{32} cm^{-2} s^{-1}$, the luminosity is also significantly lower than the nominal luminosity used at ATLAS and CMS, and this keeps the number of interactions per bunch crossing ('pileup') close to one ($\sim$1.3). This relatively low beam pile-up increases the lifetime of the subdetectors, specifically the VELO which has sensitive silicon tracking sensors 8mm from the interaction point during stable physics data taking conditions. It is also optimal for several LHCb systems for which performance decreases under conditions of high particle and primary vertex multiplicity.

As LHCb is not a hermetic detector, covering only a few hundred mrad on a single

side of the interaction point it somewhat resembles the design and operation of a fixed target experiment. Consequently, similar terminology is also used. 'Downstream' refers to further away from the interaction point towards the rear of the detector, while 'upstream' refers to closer to the interaction point. They are used as relative terms, for instance the the magnet is downstream of the VELO but upstream of the calorimeters. A right handed coordinate system is used, where the z direction travels along the beam pipe from upstream to downstream in the positive direction. The y axis is vertical from the base to the top of the detector, and the x direction is from right to left along the horizontal plane when looking downstream.

The detector consists of a large dipole magnet and several tracking systems for track reconstruction and momentum measurements of charged particles, two ring imaging cherenkov (RICH) detectors for particle identification, electromagnetic and hadronic calorimeters and muon detection systems for the identification of muons. These all surround a suspended beryllium beam pipe which is conically tapered with the smallest end closest to the interaction point. The data acquired from the readout electronics of LHCb are passed through a three tier trigger system consisting of a pure hardware trigger, L0, and two software triggers Hlt1 and Hlt2 which select potentially interesting events in real-time to store to disk. This data is further processed offline into specific 'data lines' in a process known as 'stripping' before finally being made available for offline analysis. In addition to the online and offline data analysis software infrastructure, a complex system of simulation software which uses Monte-Carlo (MC) techniques is used to simulate collision events, particle decays and material interactions within the detector. The hardware and software triggers, as well as the stripping process, are also simulated. The simulation infrastructure is paramount for the determination of quantities such as detection efficiencies, but also in accurate reconstruction of particle tracks, which depends on good modelling of material interactions in the detector.

The main components of the hardware and software are described in detail below.

2.3.1   Vertex locator

The vertex locator (VELO) [46] is a retractable array of 25 silicon detector stations which resides in a vacuum and surrounds the interaction point of LHCb. The VELO is a crucial component in LHCb's 'b-tagging' technique [47]. Correctly identifying the flavour of reconstructed b-hadrons at production is required for measurements such as time dependant CP asymmetries and b-oscillations. Furthermore, the high resolution reconstruction of primary and secondary vertices offered by the VELO facilitates high precision lifetime measurements.

The VELO is also able to measure the Impact Parameter (IP) to a high precision. The IP is defined as the distance of closest approach of the measured track from the primary vertex (the proton collision point). This measurement is crucial for the suppression of background from prompt particles, i.e. particles produced at the PV. The vast majority of particles produced at LHCb are prompt, while many particles of interest such as B and D mesons have a measurable decay length in the VELO. Therefore prompt background can be suppressed by cuts to the IP. IP resolution along the x-axis as a function of transverse momentum is shown in Figure 2.4

Each VELO station consists of two semi-circular halves (see Figure 2.5) of radial (r) and angular (ϕ) f 300 μm silicon strip sensors, with pitch that varies linearly between 38 μm and 102 μm from the inner radius of 8.2 mm to the outer radius of 42mm. Combining the information from these types of sensor allow high resolution tracking of charged particles from the primary vertex. Each semi-circular element is mounted to a carbon fibre paddle and during beam injection and setup, the VELO is withdrawn by 29mm in the horizontal plane. During collisions for physics data taking, the VELO modules move inwards to form circular detector planes surrounding the interaction point.

(a)   (b)
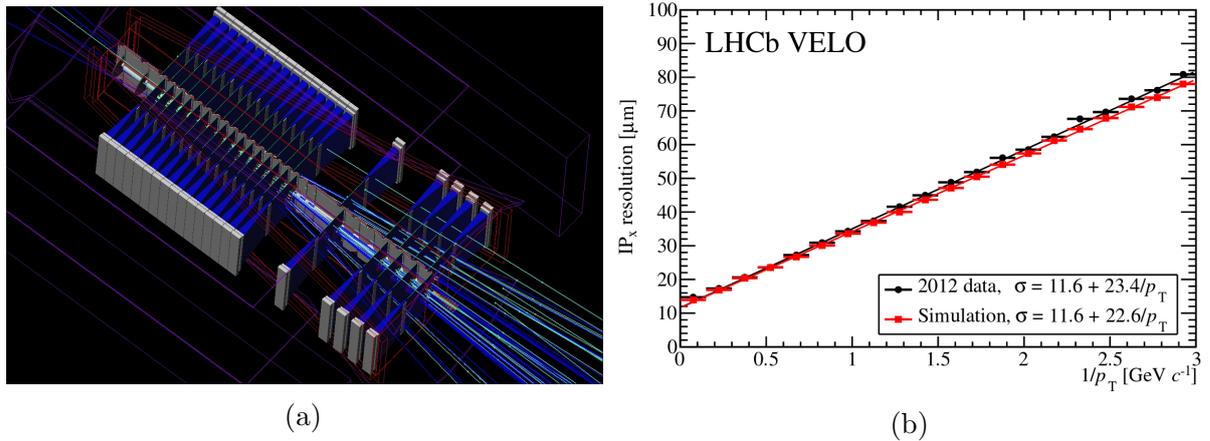
Figure 2.4: (a): Visualisation of tracks from a simulated event in the VELO. [48]. (b) IP_x as a function of p_T. [49]

The stations extend from 17.5cm upstream of the interaction point to 75cm downstream. Having information on 'backwards' tracks is useful for the reconstruction of primary vertices, hence the upstream extension. The centres of the stations leave an 0.8cm hole for the passage of the beam, and this proximity of the sensors to the beam is limited in part due the irradiation the modules receive at such close distances, as well as RF pickup from the beams.

RF interference between the beams and the VELO can significantly perturb the beams themselves as well as interfere with the performance of the VELO. For this reason, sheets of corrugated aluminium $\sim$0.25 mm thick separate the VELO sensors from the beam. This RF foil, along with wake field suppressors, minimises RF interference as well as separates the primary LHC vacuum from that of the VELO's vacuum vessel. This vacuum isolation protects the primary vacuum of the LHC from out-gassing of the detector's components.

(a) Single VELO module.   (b) VELO stations and enclosure
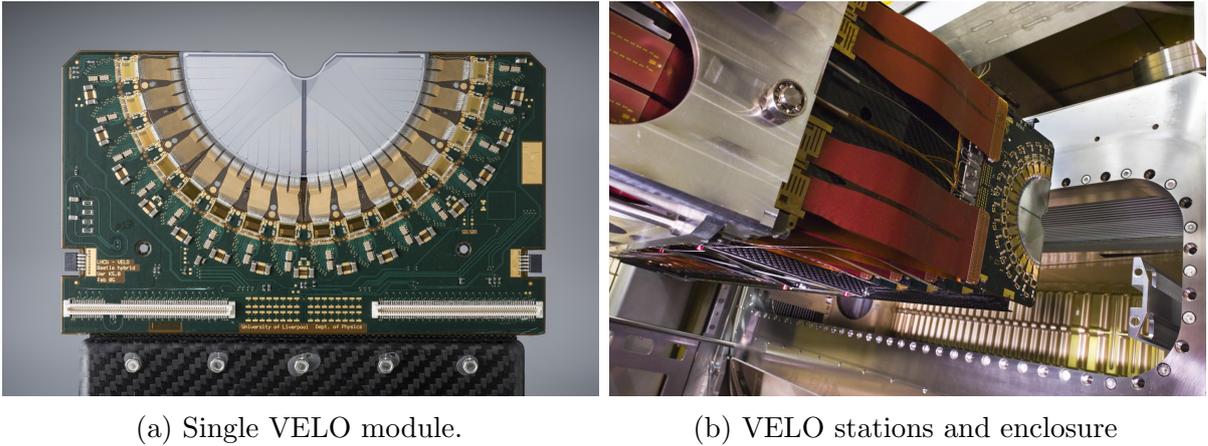
Figure 2.5: Photographs of VELO subdetector. The RF foil is visible at the bottom of the enclosure [50]

2.3.2 RICH 1 and RICH 2

There are two RICH detectors (RICH 1 and RICH 2) at LHCb [51], which are designed to provide particle identification over a momentum range of 1-100 GeV/c [52]. The first RICH detector is directly downstream of the VELO, while the larger RICH 2 is after the bending magnet and tracking stations.

Ring imaging Cherenkov detectors exploit the phenomenon of Cherenkov radiation to identify charged particles that traverse the transparent refractive medium, or 'radiator' inside the detector. When a charged particle traverses the radiator faster than the phase velocity of light in the medium, photons are emitted as a coherent wavefront as shown in Figure 2.6. The relationship between velocity and angle is $\cos\theta = \frac{1}{n\beta}$ where $\beta = \frac{v}{c}$.

In a RICH detector, these shock waves are focused and projected as a ring onto the photomultiplier tubes. Pattern recognition algorithms are used to infer the Cherenkov angle, and this process involves several different approaches [53].

In RICH 1, there are two radiators. The first, a 5cm thick block of aerogel provides K^+ identification for particle momenta above 2 GeV/c and π-K separation up to 10 GeV/c, while an 85cm long region of C_4F_{10} gas provides further π-K separation up to 50 GeV/c.
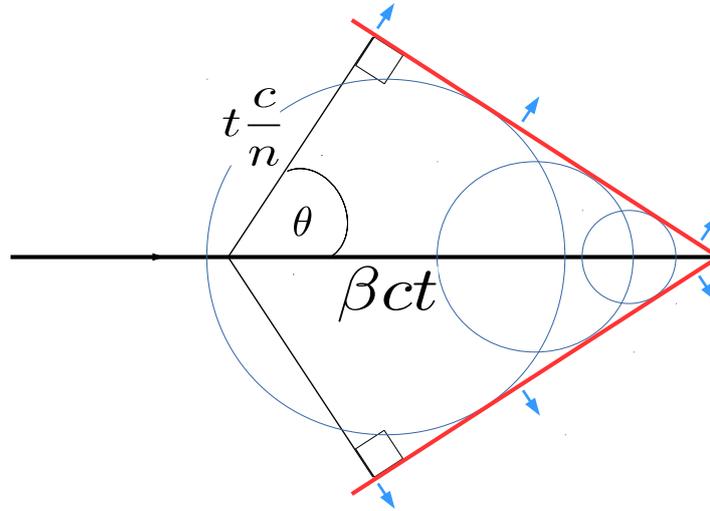
Figure 2.6: Geometry of Cherenkov radiation showing the shock front in red, and the angle with respect to the particle's velocity vector.

(a) Without RICH information.

(b) With RICH information.
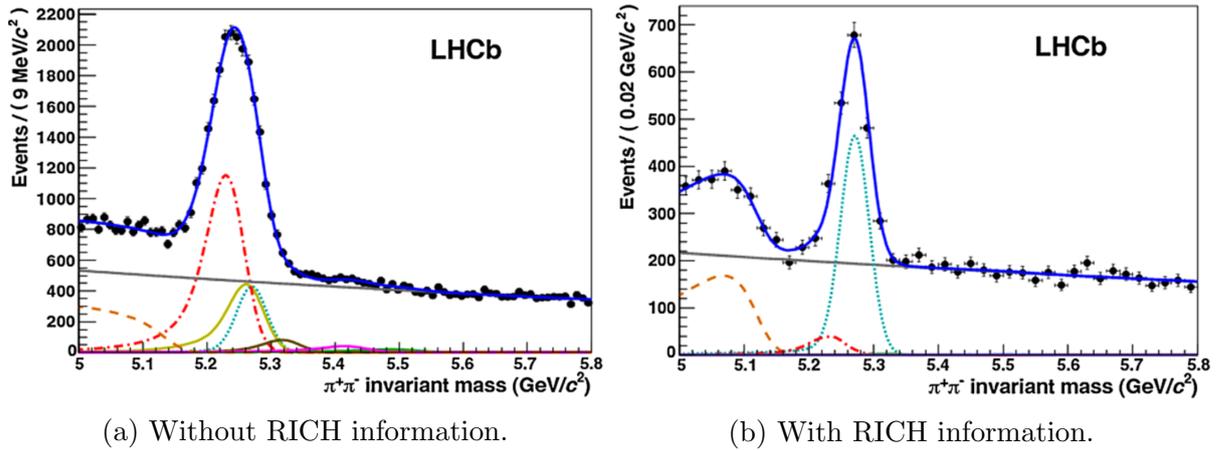
Figure 2.7: Invariant mass for $B \rightarrow h^+h^-$ decays with signal channel of $B^0 \rightarrow \pi^+\pi^-$ in turqoise, and various background components: $B^0 \rightarrow K\pi$ *red dash-dotted line*, $B^0 \rightarrow$ 3-body *orange dash-dotted line*, $B^0_s \rightarrow K\pi$ *brown line*, $\Lambda^0_b \rightarrow pK$ *purple line* and $\Lambda^0_b \rightarrow p\pi$ *green line*). (b) shows the powerful background suppression achieved when information from the RICH system is used for particle identification [54].

RICH 2 on the other hand has a single 167cm thick radiator of CF_4 gas and provides π-K separation from 50–100 GeV/c and beyond. The benefit of the RICH information is illustrated in Figure 2.7.

2.3.3   Bending Magnet

LHCb has a single, iron yoke dipole magnet with a bending power, $\int Bdl$, of 4 Tm over a 10 m track length between the VELO and the final tracking station. This provides the tracking system with the ability to measure the momentum of charged particles through their deflection with a resolution of 0.3–0.6%. The magnetic field is orientated vertically, giving a horizontal bending plane and as the magnet is a conventional magnet, it facilitates fast ramping and periodic inversion of the field polarity. This can be extremely useful in the cancellation of systematic effects and detection asymmetries, especially in measurements related to CP-violation.

As shown in Figure 2.8, a 1450 ton iron yoke surrounds two trapezoidal coils. The shape of the magnet gap between the coils follows the angular acceptance of the downstream detector systems, ensuring all particles within the acceptance of LHCb are not obstructed by the magnet yoke or coils. The warm magnet is watercooled and operates at a current of $\sim$50 kA, dissipating ~ 4.2 MW.
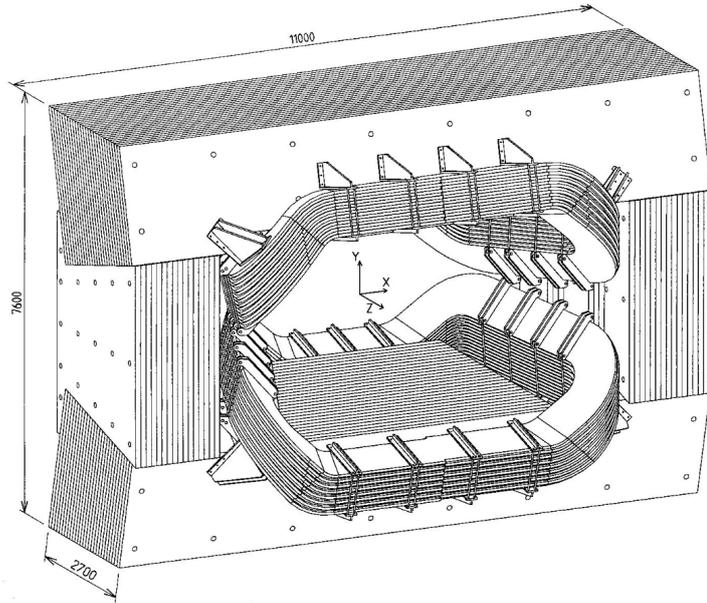
Figure 2.8: Schematic of the LHCb magnet, from the wider aperture end and looking upstream. This displays the wedge-shaped geometry of the coils and the inside of the yoke to match the angular acceptance of LHCb. [55]

2.3.4 Tracking systems

The main tracking system (Figure 2.9) is composed of four rectangular stations, the Tracker Turicensis (TT) between RICH 1 and the magnet, and three other stations (T1-
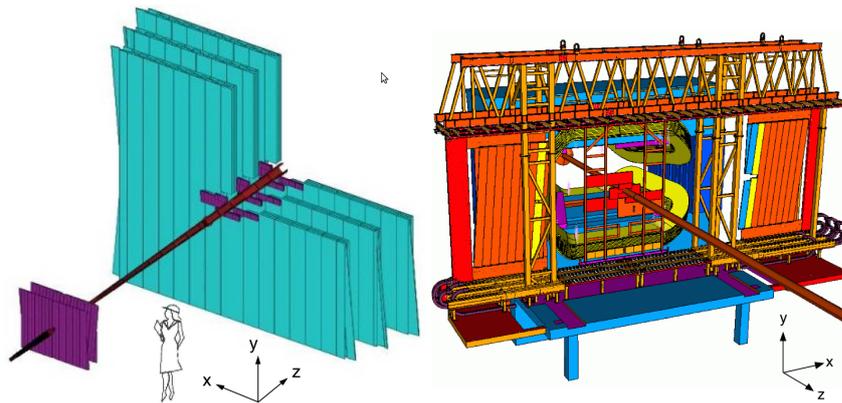
Figure 2.9: Left: Schematic of the tracking stations cut-away quadrant. Silicon modules od the TT and IT are shown in purple and straw tube modules in turquoise. Right: Tracker stations T1 shown retracted and in relation to the dipole magnet and beam pipe. [56]

T3) between the magnet and RICH 2. The TT is a silicon microstrip detector which provides momentum information at trigger level and also provides tracking for particles originating from decays outside of the VELO sensitive volume (downstream tracks) or low momentum tracks that are bent out of the detector's acceptance before reaching the downstream trackers (upstream tracks). The downstream stations, T1-T3 each consist of two parts, the Inner Track (IT) and the Outer Tracker (OT). The IT form a 'cross shape' close to the beam pipe where particle multiplicity and track density is high, and is composed of silicon microstrip sensors. The rest of each station is composed of straw tube detectors and is collectively called the OT. Further away from the beam pipe, particle fluxes are lower while tracks are bent more due to their lower longitudinal momentum. Therefore, the resolution offered by straw tubes over silicon microstrips is sufficient, especially since silicon microstrip technology is far more expensive per m^2 of sensor. Performance results from data taken in 2012 show that the TT and IT deliver, respectively, 99.8% and 99.9% hit efficiencies and hit resolutions of 53.4 µm and 53.9 µm. The OT offers 205 µm hit resolution and a 99.2% single cell hit efficiency. [56]

As mentioned in the description of the TT, at LHCb tracks that traverse the TT and stations T1-T3 are defined as either 'long' or 'downstream' (abbreviated to LL and DD when in decay pairs) depending on where the tracks originate in the detector. Relatively long-lived particles, for instance the Λ baryon, can have a decay length on the order of a metre, giving rise to charged tracks originating downstream of the VELO. Since this can give rise to differences in the measurements between long and downstream tracks, for instance momentum resolution, the long and downstream categories of tracks are often analysed separately to ensure optimal use of the data. The diagram in Figure 2.10 illustrates long and downstream tracks in relation to the tracking system and magnetic field of the dipole.
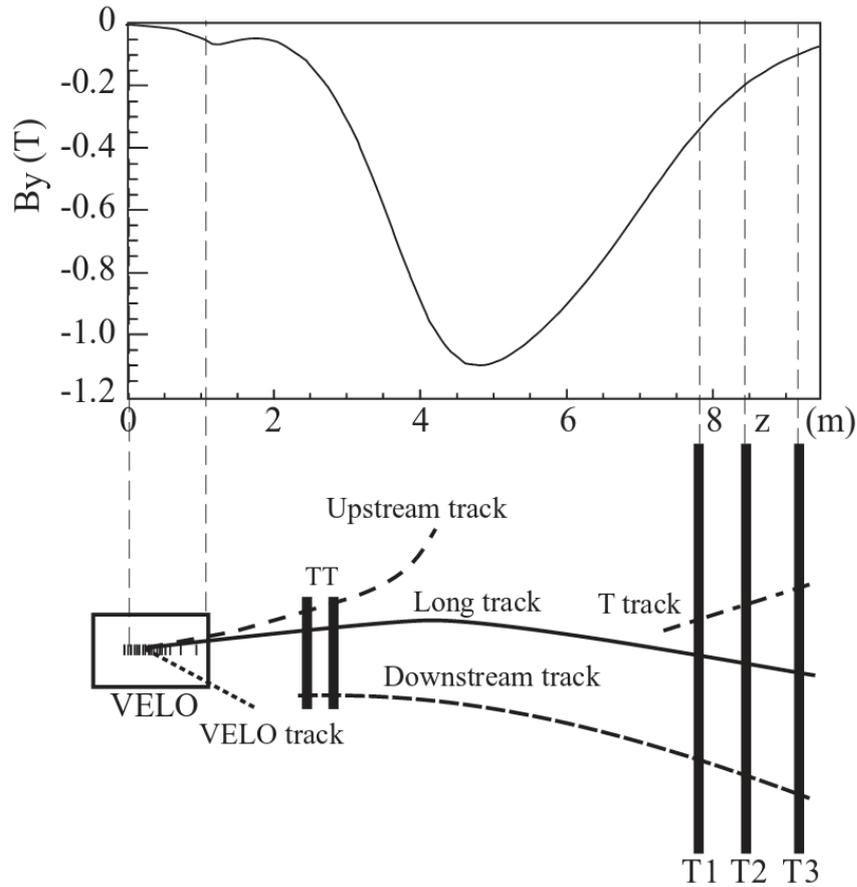
Figure 2.10: Diagram detailing the definition of upstream, downstream and long tracks in relation to the tracking system and magnetic field [56] . Also depicted are T and VELO tracks. T tracks are only present in the T1-T3 stations. These usually originate from secondary interactions but must be considered in the RICH pattern recognition algorithms since they can enter RICH 2. VELO track information is used mainly for the reconstruction of primary vertices. [35].

2.3.5   Calorimetry

The LHCb calorimeter is composed of four subdetectors; the Preshower detector (PS), Scintillating Pad detector (SPD), Electromagnetic calorimeter (ECAL) and Hadronic calorimeter (HCAL). In addition to performing energy measurements, the calorimeter system has to reject the large background from inelastic pp collisions (at a rejection level of $\sim$99%) by triggering on electrons of high transverse energy, E_{T}. There is also an large neutral and charged pion background which is suppressed by the PS and SPD systems.
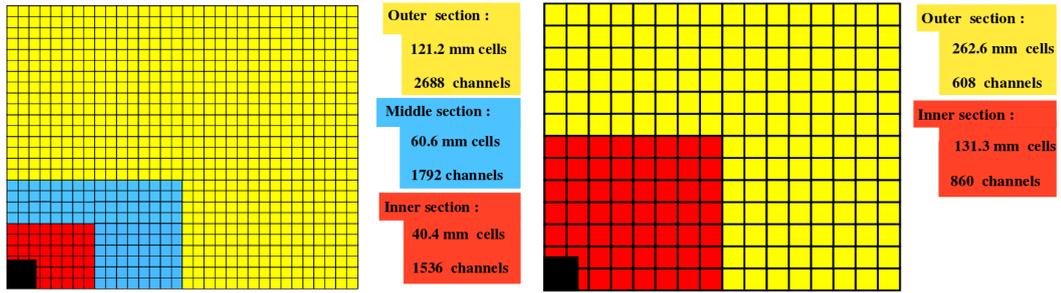
Figure 2.11: Projection of the SPD, PS and ECAL segmentations (left) and HCAL segmentation (right). Only the top right quadrants are shown, with the beam pipe in the bottom left. [35]

The SPD determines whether the traversing particle is charged or neutral, while the PS differentiates between electrons and photons, information that is primarily useful for the trigger in suppressing these pion backgrounds.

Both of these detectors use scintillating pads separated by a thin lead converter sheet. The ECAL and HCAL both use 'shashlik' technology in which each module consists of alternating layers of scintillator and lead (ECAL) or iron (HCAL) plates. All four detectors use multi-anode photomultipliers to detect the scintillation photons. In a similar design to the tracking stations, the calorimeter systems are segmented to offer varying degrees of granularity. The hit density varies by two orders of magnitude across the surface of the calorimeters, with the highest densities closest to the beam pipe, therefore all four calorimeter systems are segmented in the fashion displayed in Figure 2.11.

The energy resolution of the ECAL was measured using test beams [57]. The performance depends on the modules are tested, but a resolution of around

$$\frac{\sigma(E)}{E} \approx \frac{9\%}{\sqrt{E}} \oplus 8\% \oplus \frac{0.003\,\text{GeV}}{E}, \tag{2.3}$$

was measured. The three percentages correspond to the stochastic, constant and noise

terms. While for the HCAL this was measured to be,

$$\frac{\sigma(E)}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%, \qquad (2.4)$$

where in both cases, E is in GeV.

2.3.6 Muon systems

Muon triggering and identification is crucial for extracting interesting physics from the collisions at the LHC. The last main section of LHCb is composed of five muon detector stations (Figure 2.12). Four of the muon stations (M2–M5) are downstream of the HCAL, and separated by 80cm thick iron filter plates while one muon station (M1) occupies the space between RICH 2 and the calorimeters. Information from M1 is used primarily for the L0 trigger, in particular for transverse momentum measurements at the L0 level, when combined with information from the downstream stations. The M2-M5 stations are used for both triggering and reconstruction. The muon system selects muons with a high efficiency of ∼95%.

M1–M5 are composed of modules of Multi-Wire Proportional Chambers (MWPC) apart from the inner section of M1 which uses Triple-GEM (Gas Electron Multiplier) detectors due to the high particle flux. The MWPC sections are divided into projective regions around the beam pipe like the tracker stations and calorimeters that scale with particle flux. All these regions are composed of MWPC chambers but the number of 'logical pads' per chamber scales in the ratio 1:2:4:8 to increase granularity close to the beam pipe. This division of chambers into logical pads is shown in Figure 2.13

MWPC's consist of an array of wires held at high voltage between two conductive plates. The space between the two plates is filled with a gas mixture which is ionised by the passage of charged particles such as the muons. The placement of the iron filters between
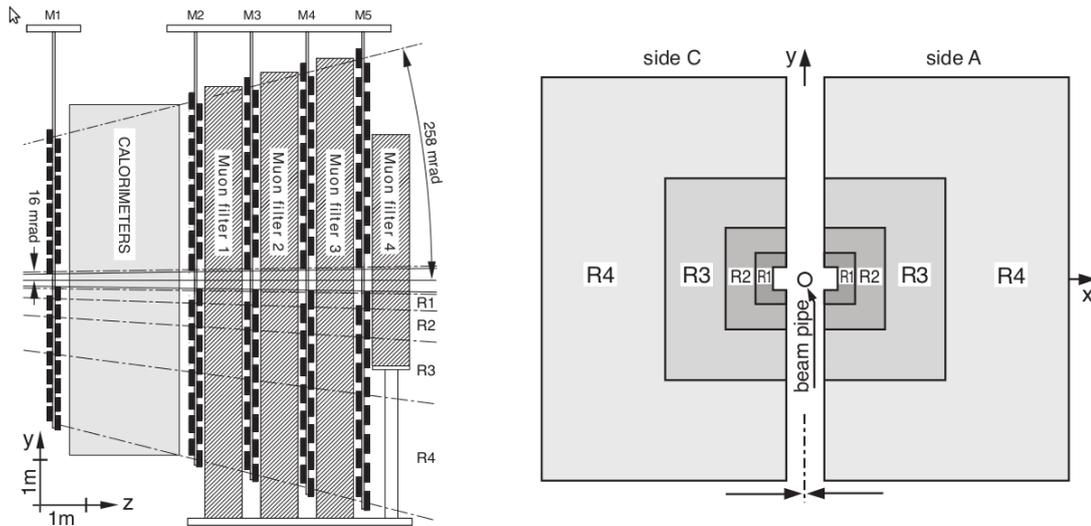
Figure 2.12: Left: Side view of the five muon stations showing the iron filters and acceptances of the different station regions. Each consecutive station and its regions are scaled in size to follow the acceptance of LHCb. Right: front view of a station showing the different regions. Side A and C close to be hermetic along the central axis but can be opened for maintenence. [58]

the muon stations is to remove background charged particles since the boosted muons that LHCb is interested in are highly penetrative; if they are inside the detector's acceptance, nominally they will traverse the filters and stations, decaying far downstream of the entire detector. In the muon stations of LHCb, the 30 µm thick gold-plated tunsgten wires of the MWPC's are spaced 2 mm apart with a gas gap of 5 mm and operated at a voltage of ~2.6 kV. The gas is a mixture of argon, carbon dioxide and CF_4 with a ratio of 40/55/5. the physical principle behind detection is same as with the straw tubes of the OT. The main difference in design is having an array of wires inside each gas chamber rather than individual wires inside conductive gas filled straws. The electric field between the wires and plates causes charge carriers liberated by the traversing muon to drift to the wires and plates, producing an electrical signal in the wires.

GEM detectors, such as the ones used in the M1 station use a similar principle, except the field is produced by small conductive rings around holes in polymer foils [59]. These holes can be thought as 'pixels' on the polymer. Electrons from ionisation in the gas
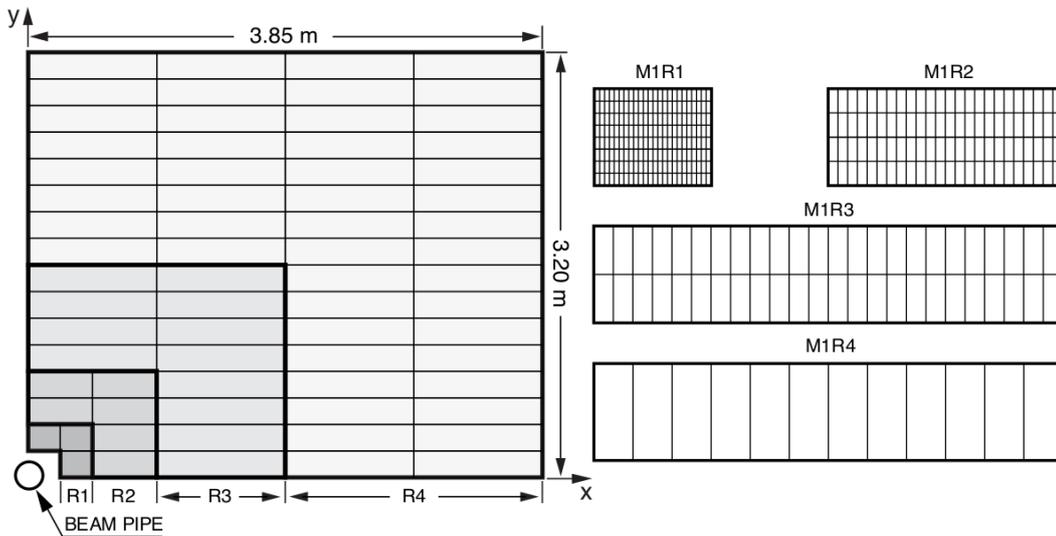
Figure 2.13: Left: Top right quadrant of station M1 showing the division into chambers (the horizontal rectangles) and regions (R1–R4.) Right: Division of chambers into logical pads for each type of region. The vertical alignment of the rectangular pads gives higher granularity in the bending plane. [58]

gap are drawn into the holes where they initiate an electron avalanche and therefore a detectable signal. The ability to fabricate these holes with spaces on the order of μm allows a precise hit resolution to be achieved. In the centre of the M1 station there are 12 GEM detectors which use the same gas mixture as the MWPC's but with a ratio of 45/15/40.

The x–y spatial resolution varies from $\sigma_x \times \sigma_y = 4 \times 10 (mm^2)$ for the innermost region of M1, to $\sigma_x \times \sigma_y = 150 \times 180 (mm^2)$ for the outer most region of M5 [58]. Note the higher spatial resolution in the bending plane.

2.3.7 Trigger

The trigger system [60] [61] is required to massively throttle the raw data rate coming from the LHCb readout systems in a way that is heavily biased towards potentially 'interesting' events. Expressly, it must perform online selection of events to reject background and

select these possibly interesting events with a high efficiency in real-time. At nominal operation, there are 40 million bunch crossings a second (15 million with Run 1's 50 ns bunch spacing) inside the VELO, with an average of 1–2 pp collisions per crossing. To store all of this data for offline processing is not feasible with current technology, so a hardware first level trigger (L0) is used to decrease this rate to ∼1 MHz, while the two tiers of software triggers, High Level Trigger 1 (Hlt1) and High Level Trigger 2 (Hlt2), decrease the write-to-disk rate to a manageable 3 kHz.

The algorithms used, especially in the higher level triggers are many and varied therefore only a general description of the trigger system is given below.

L0 trigger

The L0 trigger is composed of three independent triggers: The L0-Muon trigger, L0-Calorimeter trigger and the L0-PileUp trigger. The latter is used only for measurement of the luminosity, while the other two select potentially interesting events to pass to the higher level triggers.

The L0-Muon processors select events that contain muons by searching for hits in the five muon stations corresponding to particles with the highest p_T and the second highest p_T in an event. The hits must form a straight line pointing towards the interaction point. If this largest p_T passes a set threshold in L0 then the muon candidate is selected by the `L0Muon` line. If $p_{T1} \times p_{T2}$, (where p_{T1} and p_{T2} are the highest and second highest muon transverse momenta respectively) passes this threshold, the candidate is selected by `L0DiMuon`.

The L0-Calorimeter processors analyse the E_T of clusters in the calorimeter systems for the selection of hadron and photon candidates. For the selection of hadron candidates, the trigger looks at the highest E_T cluster in the HCAL; any ECAL cluster directly upstream may be matched to this and they are selected as a hadron candidate by the `L0Hadron`

line if the sum of these E_T's passes the threshold. If the highest E_T cluster in the ECAL has hits in the PS directly in front of this cluster, but no hits in the SPD, L0 fires as an `L0Photon` candidate. Alternatively, if there are hits in the SPD as well, L0 fires on the `L0Electron` line.

High level trigger

The high level triggers run on PC farms, reducing background and overall event rate to disk to a few kHz. For Hlt1 trigger lines which do not require muons to be present, tracks in the VELO are used to find the smallest IP's. Quality criteria are imposed, based on the number of hits associated with the track relative to number of expected hits. For the muon lines, namely events that trigger `L0Muon` or `L0DiMuon`, a search window is defined in muon station M3. The size of this window is determined by the expected size of the multiple scattering effect in the vertical, non-bending plane and by the expected deflection of a 6 GeV muon in the bending-plane, as 6 GeV is the minimum cut off for selected muons. If the extrapolated VELO track falls within this window, corresponding hits are searched for in M2, M4 and M5. A linear fit is then performed on this extrapolation and the requirement of $\chi^2/nDOF < 20$ is imposed to select the track as a muon candidate.

Hlt2 employs one of the two offline track reconstruction algorithms to fully reconstruct tracks while the inclusive b-decay topological lines perform partial reconstruction of b hadrons, covering all b-hadrons with at least two charged particles in the final state and a displaced decay vertex. There are also several exclusive lines dedicated to perform a full reconstruction of b and c decays.

2.3.8   Particle Identification

The VELO, RICH, tracking, calorimeter and muon systems all play a roll in identifying particles[5]. The combined information from the RICH detectors, calorimeters and muon stations is used to identify charged particles; electrons, muons, pions, kaons and protons. The ECAL is used to identify neutral pions and photons.

The RICH detectors identify particles using the techniques described in Section 2.3.2.

Muons are identified by extrapolating the tracks with a momentum of $3\,\mathrm{GeV}/c$ or more into the muon stations; muons with momentum less than $3\,\mathrm{GeV}/c$ would not arrive at the muon stations. Hits are then searched for in the muon stations in the regions in which they are expected. A cut is applied to the number of stations with which there are hits and this cut varies with track momentum. A log-likelihood between muon and pion hypotheses is also generated for the track.

Electron identification is performed in the ECAL by comparing the ECAL cluster centre with the expected centre from track extrapolation. Energy deposition in the PS and HCAL along the extrapolated track is also used to improve electron identification. Bremsstrahlung photons are also considered. If the electron emits bremsstrahlung radiation before the magnet, the electron will be deflected sideways by the magnetic field while the photons will not, and produce a separate trackless hit cluster in the ECAL. These cluster energies are recombined with electron clusters by extrapolating back the photon cluster to see if it matches the electron track before the magnet.

Photons themselves are identified by clusters in the ECAL and their separation from tracks, as well as hit information in the SPD. Clusters in the ECAL separated by a given distance from any tracks identified as photons, while photons from secondary interactions with the detector material after the magnet are identified by corresponding hits in the

---

[5]Mainly the RICH, calorimeter and muon systems play a roll in particle identification, but parameters from the VELO and tracking systems are also used as training variables in the generation of global particle ID's using multivariate techniques.

SPD.

Neutral pions are reconstructed and identified by pairs of photon hits in the ECAL, from $\pi^0 \to \gamma\gamma$. However, pions with sufficiently high p_{T} will be seen as two merged photon clusters. Algorithms are used to identify these merged clusters [56].

Global particle ID

The information from these particle identification techniques described above is merged into a single log likelihood variable by adding the log likelihood from each subdetector linearly. Each likelihood is relative to the pion hypothesis. A particle ID and therefore mass hypothesis can then be assigned by performing cuts on combinations of these likelihoods. The global identification is capable of distinguishing protons, pions, kaons, electrons and muons with high efficiency and relatively low contamination. For example, during Run 1, a requirement on the log likelihoods of $\Delta log\mathcal{L}(K - \pi) > 0$ gives a $\sim$95% kaon identification efficiency with 10% mis-id rate. Alternatively $\Delta log\mathcal{L}(K - \pi) > 5$ gives a efficiency and mis-id rate of 85% and 3% respectively [62][6].

An additional technique to assign a particle ID based on multivariate techniques in which the above information is used. This also includes the correlations between subdetectors and further information such as track quality χ^2 etc from the tracking system as training variables in a multi-layer perceptron neural network (trained using simulated events). This is used to assign a Bayesian probability for each particle type, to each track. These are called `probNN` variables and are used extensively in the analysis described in chapter 3.

---

[6]For tracks between 2 and 100 GeV/c

2.3.9   Simulation

LHCb employs several software packages to describe the behaviour of particles which are all contained in the in-house simulation framework, Gauss [63]. The *pp* collisions are generated with Pythia [64]. The decays of the particles that are produced are described the EvtGen software [65]. Final state radiation of the decay is generated with Photos [66]. When simulating data at 'reconstruction (reco) level, where the simulated data replicates observed data as closely as possible, the material interactions with the detector have to be modelled along with the online and offline processing such as trigger, stripping and reconstruction software. A highly detailed geometric model of LHCb is constructed in the simulation software and describes the materials and their densities as well as the geometry. Interaction with this material is simulated with the Geant4 package [67] [68].

Accurate and precise modelling of the detector materials and these interactions is crucial for good track reconstruction and accurately simulated data. For this reason, Geant4 and its usage within Gauss regularly undergo validation testing, especially after patches, changes or version upgrades have been implemented. Section 2.3.10 describes my contribution to the simulation validation development at LHCb.

2.3.10   Electromagnetic physics simulation at LHCb

The Geant4 package is widely used across many research diciplines for the purpose of modelling the interaction of particles with materials. Geometric models representing objects such as detector systems, detector infrastructure, solid and gaseous volumes are described, with the corresponding material properties assigned. Due to its wide range of applications, Geant4 supplies numerous model plug-ins and options known as *physics lists*. There is no 'one-size-fits-all' modelling of the interaction behaviour when dealing with a plethora of electromagnetic and hadronic interactions across energy ranges that span several orders of

magnitude, therefore several specialised physics lists are supplied in which different physics models and configurations are used. There are some complementary models that deal with the same types of interactions but cover different energetic regimes, while in some cases there also exists competing models to describe the same type of interactions but where one model may have advantages over the other in certain applications or depending on whether the priority is speed or accuracy. It is a crucial aspect of Gauss development to perform regression testing, compare available physics lists and options and validate production versions of Gauss whenever the implmentation of GEANT4 is updated or changed. It is also worthwhile to perform these validation tests on a periodic basis to ensure any unexpected changes that affect material interactions are likely to be detected.

Several components of the testing and validation of EM physics interactions are performed in a simulation of the VELO detector. In this environment, there exists only the VELO silicon, beam pipe and beam pipe support strictures in a configuration identical to LHCb during data taking. Instead of proton-proton collisions being simulated at the interaction point, particle emitters (particle guns) fire mono-energetic charged particles downstream through the VELO. In these tests the particles are fired from the emitter within 0 and 0.39 mrad of the z-axis. The simulation is performed within the full Gauss framework to test both GEANT4 and its implementation in Gauss. The interactions of the particles with the VELO silicon are probed by extracting the energy deposition dE/dx where the particle has a path length through the silicon of $300 \pm 10\,\mu$m, as the VELO silicon is $300\,\mu$m thick.

Additionally, photon information is captured, such as the number of photons emitted for each track, and each photon's energy. This is to look primarily at the bremsstrahlung radiation emitted by the particles as they traverse the material.

The packages `EmGaussMoni` (derived from the pre-existing `VeloGaussMoni` tool) and `BremVeloCheck` were developed for Gauss in order to extract the energy loss and

bremsstrahlung information from the simulation.

2.3.11   VELO Energy Deposits

The dE/dx behaviour follows a skewed Landau distribution characterised by the most probable value (MPV) at which the distribution peaks, as well as the peak's full width at half maximum (FWHM). These two values are estimated for several types of charged particles across a range of energies. The MPV and FWHM values and are plotted as a function of $\beta\gamma$ where γ is the Lorentz factor and β is the ratio of the particle velocity to the speed of light, allowing all particles and energies to be compared on a single graph. The chosen energy values and resulting $\beta\gamma$ values are displayed in Table 2.1.

Energy(GeV)	$\beta\gamma_{e^\pm}$	$\beta\gamma_{\mu^\pm}$	$\beta\gamma_{\pi^\pm}$
0.1	195.695	0.943396	0.714286
0.2	391.389	1.88679	1.42857
0.4	782.779	3.77358	2.85714
1	1956.95	9.43396	7.14286
5	9784.74	47.1698	35.7143
10	19569.5	94.3396	71.4286
17	32876.7	158.491	120
50	97847.4	471.698	357.143
100	195695	943.396	714.286
120	234834	1132.08	857.143
168	328767	1584.91	1200

Table 2.1: Considered particle gun energies and the corresponding $\beta\gamma$ values for each particle species.

Eleven particle gun energies, from 0.1 to 168 GeV are used for the simulation of electons, muons and charged pions. For each case the respective anti-particles are also simulated. These energies and particles cover a large range of $\beta\gamma$ in which the distributions of each particle overlap, with pions and muons covering the low $\beta\gamma$ regime, and electrons accessing high $\beta\gamma$. Due to the smearing of the energy distributions, a simple Landau fit is inappropriate. The MPV of the Landau distributions from each simulation is estimated using a parabolic fit to the peak while the FWHM is estimated using linear interpolation. An example of one of the dE/dx distributions is shown in Figure 2.14, alongside the simulated hits in the x-y plane of the VELO.

(a)

(b)
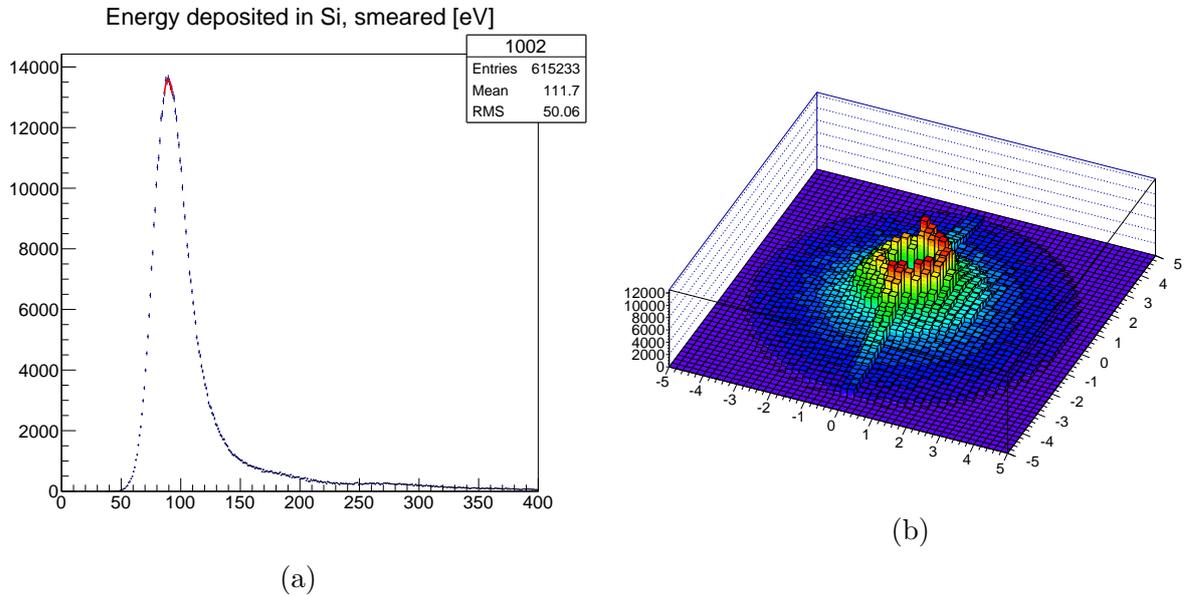
Figure 2.14: (a) Energy deposited in 300 μm of silicon for 120 GeV electrons. (b) Hit map of the electrons in the Velo silicon.

Figure 2.15 shows the dE/dx behaviour across a large $\beta\gamma$ for anti-muons in copper. For thin material layers, the MPV and FMHWM can be estimated using the Bethe-Bloch relation [7], with the consideration of density effects. The MPV is described by,

$$\Delta_p = \xi \left[ln\frac{2mc^2\beta^2\gamma^2}{I} + ln\frac{4\pi N_A r_e^2 m_e c^2 Z}{2IA} \right] + j - \beta^2 - \delta(\beta\gamma), \qquad (2.5)$$

and the FWHM is,

$$\xi = \frac{4\pi N_A r_e^2 m_e c^2 Z}{2A},$$ (2.6)

where the parameters are defined in Table 2.2.

Parameter	Description
m_e	Electron mass
N_A	Avogadro's number
r_e	Vlassical electron radius
Z	Atomic number of material
A	Atomic mass of material
I	Mean excitation energy
j	2.00 [7]
$\delta(\beta\gamma)$	density effect correction

Table 2.2: Bethe-Bloch equation parameters.

In the situation where the validation tests are required, but there is no change to which physics lists are used in LHCb, the same physics lists from each version are compared, along with theoretical estimates using Equations 2.5 and 2.6. In such cases where, for example, the viability of a new physics list available in Geant4 is to be tested, the results from each available list are compared with each other, alongside the theoretical estimate. The estimates calculated from Equations 2.5 and 2.6 are not expected to have perfect agreement with the results from the physics lists as the simulation contains much more detailed modelling than the analytical ξ and Δ_p values, but it is included as a 'sanity check.'

These tools are designed to be implemented into the LHCb performance and regression testing (LHCbPR) [69] software. LHCbPR allows for systematic extraction and comparison of results, with planned features that allow automatic running of jobs and validation of
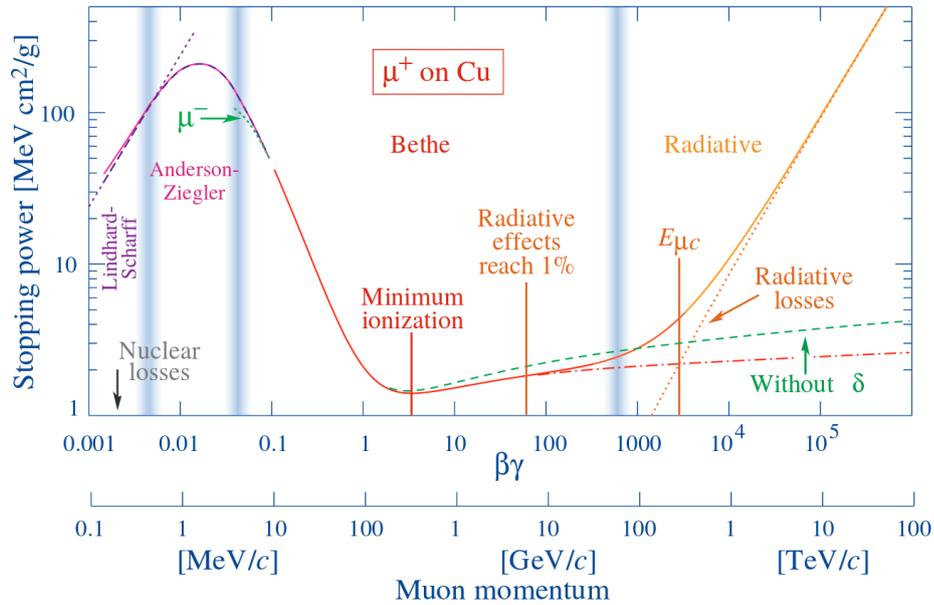
Figure 2.15: dE/dx MPV as a function of $\beta\gamma$ for μ^+ in copper [7]. In this case, the Bethe approximation is valid for $\beta\gamma$ values of ~ 0.1–1000 (indicated by the vertical bands), within which lies the minimum ionisation energy. The minimum ionisation energy is used as a validation parameter in these EM physics simulation tests.

the output. The full simulation validation procedures of physics simulation at LHCb are extensive and consists of numerous stages. The dE/dx and bremsstrahlung tests described here are the second stage, (the first being stand-alone testing in GEANT4) and they are designed to be relatively fast, repeatable and systematic so that they can be run frequently and automatically. For this reason, choices have been made such as to parametrise the dE/dx distributions with parabolic fits to the peak and linear interpolation for the FWHM so that the tests are quick and require minimal oversight. They provide results designed for visual inspection in which any discrepencies can be investigated further with a more sensitive and diverse set of tests.

The physics lists relevant to LHCb are all based on the 'EM Standard' list (`EmStd`) and consist of EM Standard Option1 (`EmOpt1`), EM Standard Option3 (`EmOpt3`) as well as two other lists, `EmLHCb` and `EmNoCuts`. `EmStd`, `EmOpt1` and `EmOpt3` are 'off the shelf'

GEANT4 physics lists, the properties of which can be found for GEANT4 v9.4, v9.5 and v9.6 in [70]. `EmLHCb` and `EmNoCuts` however are private LHCb lists, based off the `EmOpt1` list but with modifications to production cuts. The details of how these lists differ are shown in Table 2.3.

By default, in `EmOpt1`, production cuts are applied to the bremsstrahlung, ionisation and e^+e^- pair production processes. However, production cuts can also be applied to the photoelectric effect, Compton scattering and gamma conversion (set internally with the `SetApplyCuts()` method). In Table 2.3, `EmNoCuts` is split into `EmNoCutsOld` and `EmNoCutsNew`. This is due to the GEANT4 Collaboration recommending that with the upgrade from GEANT4 v9.5 to v9.6, LHCb updates the `EmNoCuts` list to the latest configuration of `EmOpt1`, in which there are changes to the multiple scattering models. In LHCb nomenclature, 'Sim06b' refers to the simulation version in which GEANT4 v9.4 is used, while 'Sim07' uses v9.5 and 'Sim09' uses v9.6[8].

Physics List	e+/e- Multiple Scattering Model	Other Properties
EmOpt1	`UrbanMsc93`<100MeV `WentzelVI`>100MeV `fMinimal` Step Limit Also Uses `G4CoulombScattering`	`opt.SetPolarAngleLimit(CLHEP::pi)` `opt.SetApplyCuts(true)`

EmOpt3	`UrbanMsc95`
	`SetStepLimitType(fUseDistanceToBoundary)`
	`SetRangeFactor(0.01)`
		`opt.SetPolarAngleLimit(CLHEP::pi)`

EmStd	`UrbanMsc95` <100MeV `WentzelVI`>100MeV Also uses `G4CoulombScattering`	`opt.SetPolarAngleLimit(CLHEP::pi)`
EmNoCutsOld	`UrbanMsc95` `fMinimal` Step Limit	`opt.SetPolarAngleLimit(0.2)`
EmNoCutsNew	`UrbanMsc93`<100MeV `WentzelVI`>100MeV `fMinimal` Step Limit Also Uses `G4CoulombScattering`	`opt.SetPolarAngleLimit(CLHEP::pi)`
EmLHCb	`UrbanMsc93`<100MeV `WentzelVI`>100MeV `fMinimal` Step Limit Also Uses `G4CoulombScattering`	`opt.SetPolarAngleLimit(CLHEP::pi)` `SetApplyCuts(true)`

Table 2.3: A table summarising relavant properties of the electromagnetic physics lists for LHCb. These properties are all for GEANT4 v9.6 except for `EmNoCutsOld` which is from v9.5 and is used as a reference when comparing to v9.6.

---

[7]Sim08 also used GEANT4 v9.5, so no validation tests were performed with these tools.

[8]Full version numbers are 4.9.5.p02 and 4.9.6.p04. v9.4 and v9.5 are used here for brevity.

For the upgrade between GEANT4 v9.4 and v9.5, the dE/dx tests were performed on pre-existing centralised productions of particle gun event simulations and the data was extracted from these using the aforementioned `EmGaussMoni` tool. These results are presented in Section 2.3.11. For the subsequent upgrade from v9.5 to v9.6, a new set of analysis tools was developed, facilitating the production of specialised MC samples by the end user for these validation tests. Furthermore, between the upgrade from v9.5 and v9.6, the `BremVeloCheck` tool was developed for the additional test on bremsstrahlung radiation. Therefore, these tests are only performed for the validation of v9.6. The results of the GEANT4 v9.6 validation studies are presented in Section 2.3.11

Geant4 v9.4 vs v9.5 results

Overall there is good agreement between the lists within v9.5, and consistency between v9.4 and v9.5 with `EmNoCuts`, however there are two exceptions. For `EmOpt3` there are not enough hit statistics between 290 and 310 μm to evaluate the distributions. This is unsurprising due to the difference in step size used in `EmOpt3` with respect to the other lists. Therefore, the consistency of `EmOpt3` with others is evaluated for path lengths greater than 300 μm, resulting in significantly more hits, and shows consistency with theory and other models. The results for $x > 300$ microns is shown in figures 2.22 and 2.23.

`EmNoCuts` shows some fluctuations in the FWHM at high $\beta\gamma$ from the expectation for electrons. The MPV however, is consistent across the full $\beta\gamma$ range, and the behaviour of the more critical muons and pions is consistent. These results can be seen in Figures 2.18 and 2.19. One possible reason for the deviations in `EmNoCuts` electrons for v9.5 is due to the accuracy limitations of the linear interpolation method used to estimate the FWHM. Figures 2.16 and 2.17 present a comparison between the energy loss distributions of 1 GeV electrons in v9.4 and v9.5 `EmNoCuts`, which corresponds to v9.5 $\beta\gamma$ point with the largest deviation, and it is clear they are actually highly consistent.

Figures 2.20 and 2.21 show the comparison of the lists in GEANT4 v9.5 that were considered. `EmOpt3` is omitted due to insufficient statistics. `EmOpt3` is included in Figures 2.22 and 2.23 for a path length cut greater than $300\,\mu$m only.
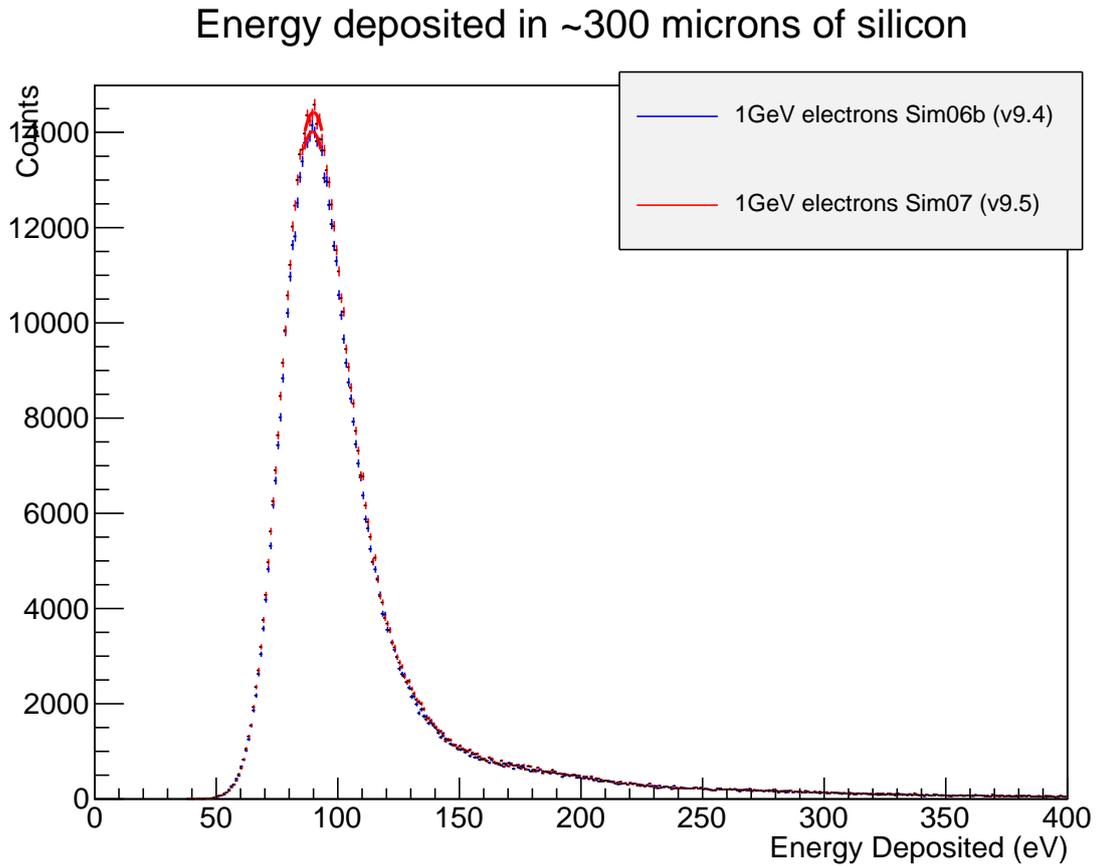
Figure 2.16: Deposited energy for 1GeV electrons, comparing Sim07 and Sim06b EmNoCuts.
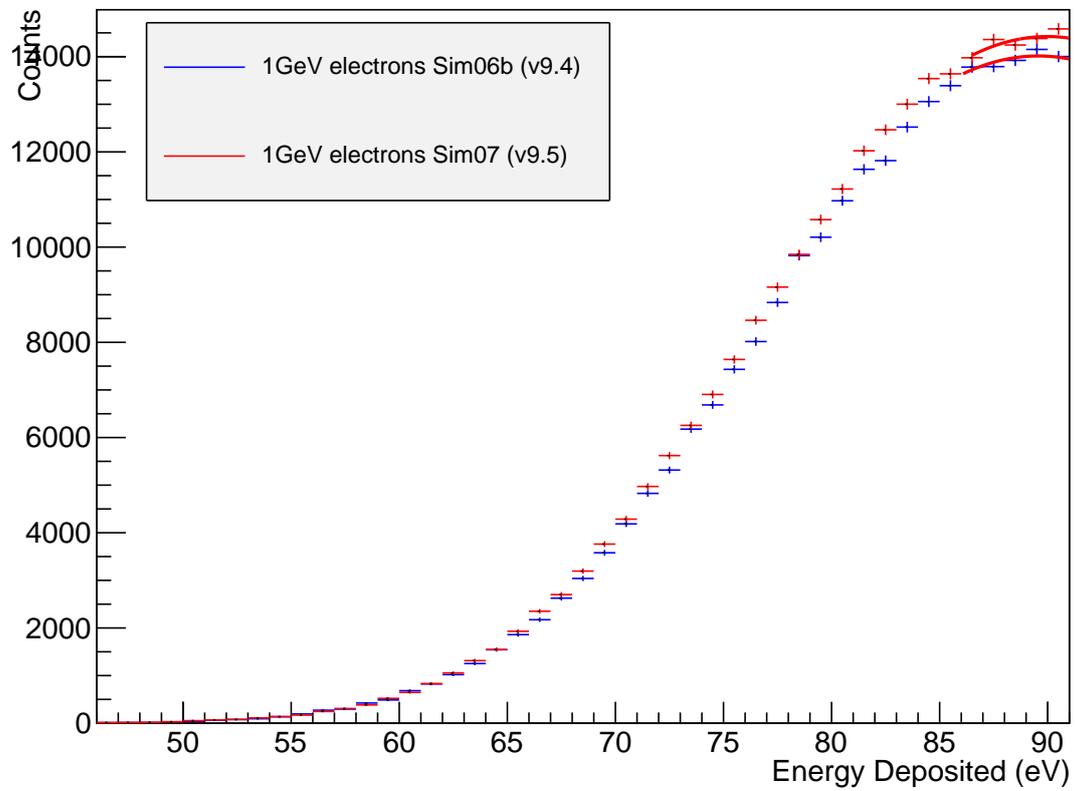
Figure 2.17: Magnification of distributions in figure 2.16.
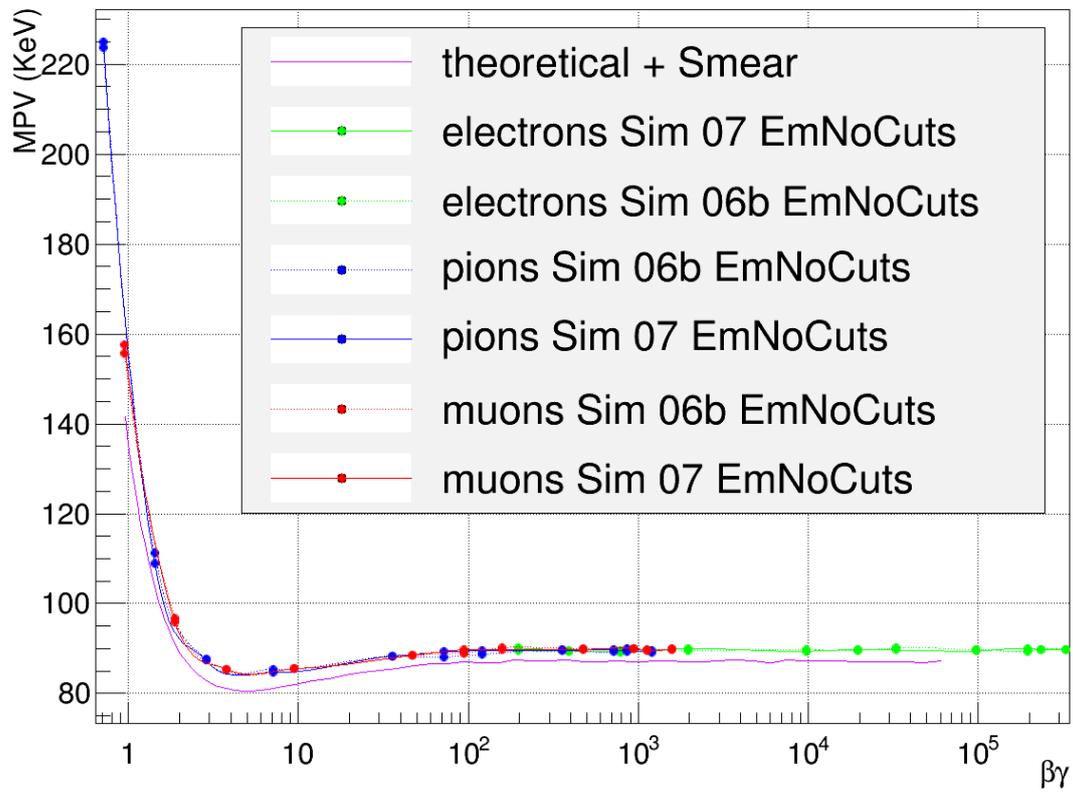
Figure 2.18: MPV comparison for v9.4 vs v9.5 EmNoCuts, $x \approx 300 \mu m$.
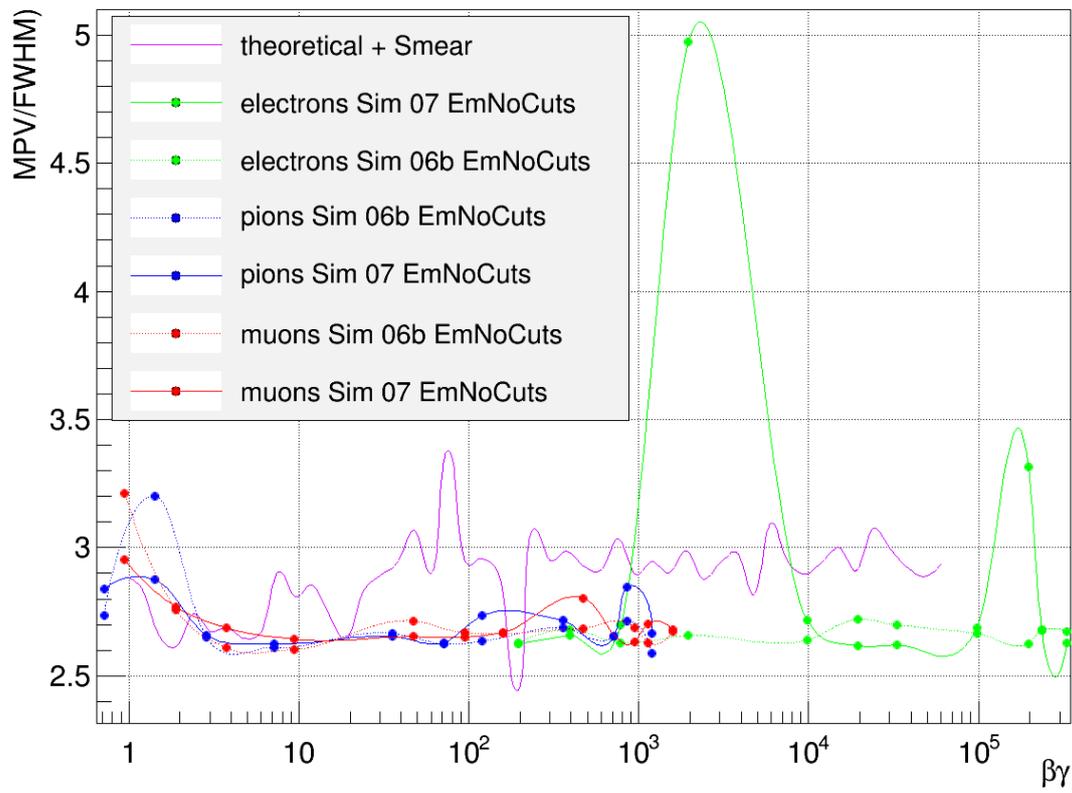
Figure 2.19: MPV/FWHM comparison for v9.4 vs v9.5 EmNoCuts, $x \approx 300\mu m$.
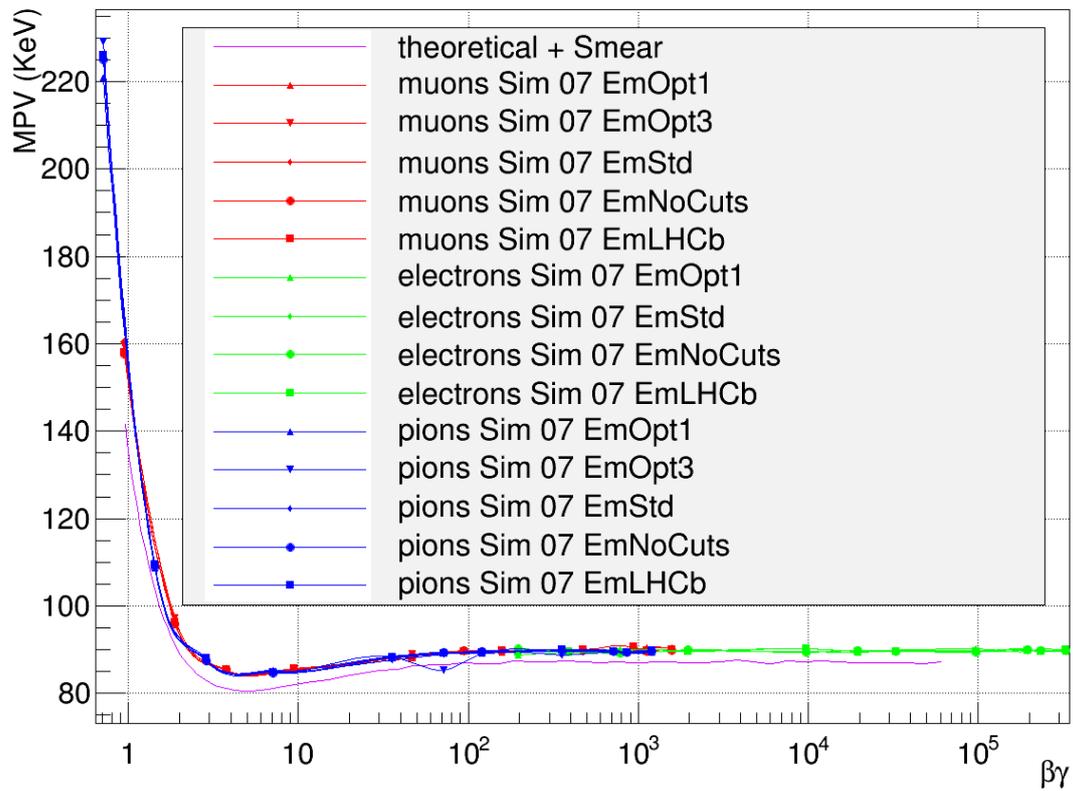
Figure 2.20: MPV comparison for v9.5 physics lists, $x \approx 300 \mu m$.
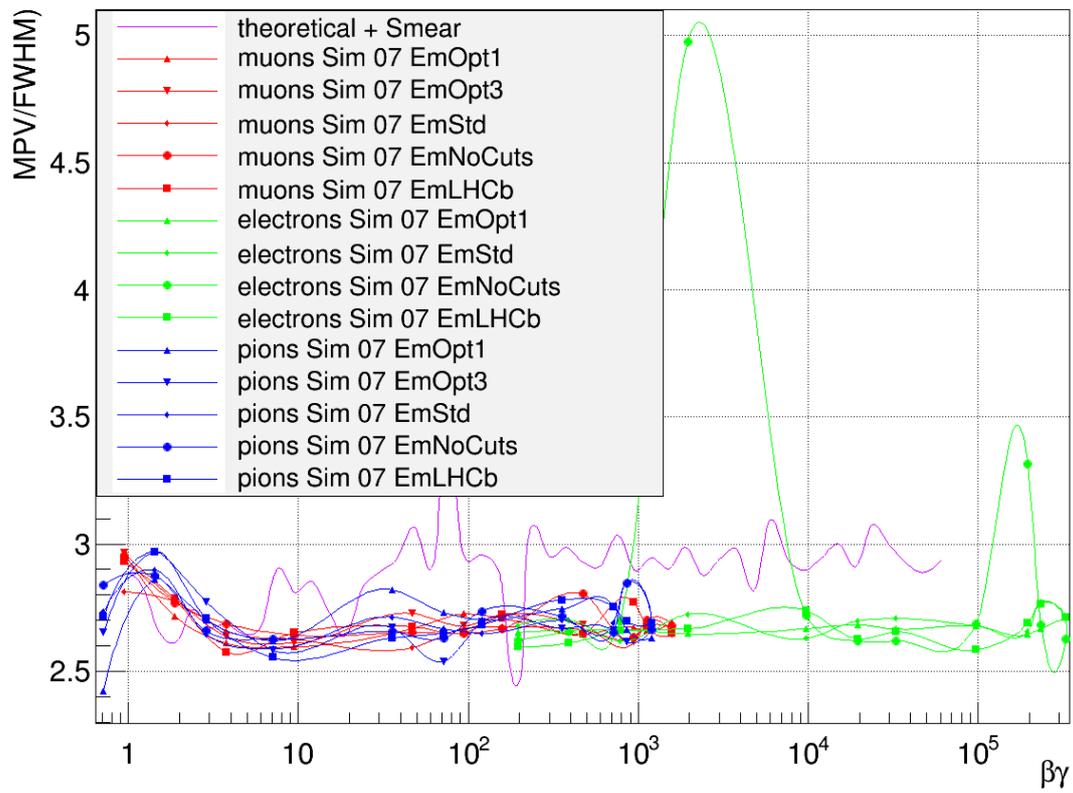
Figure 2.21: MPV/FWHM comparison for v9.5 physics lists, $x \approx 300\mu m$.
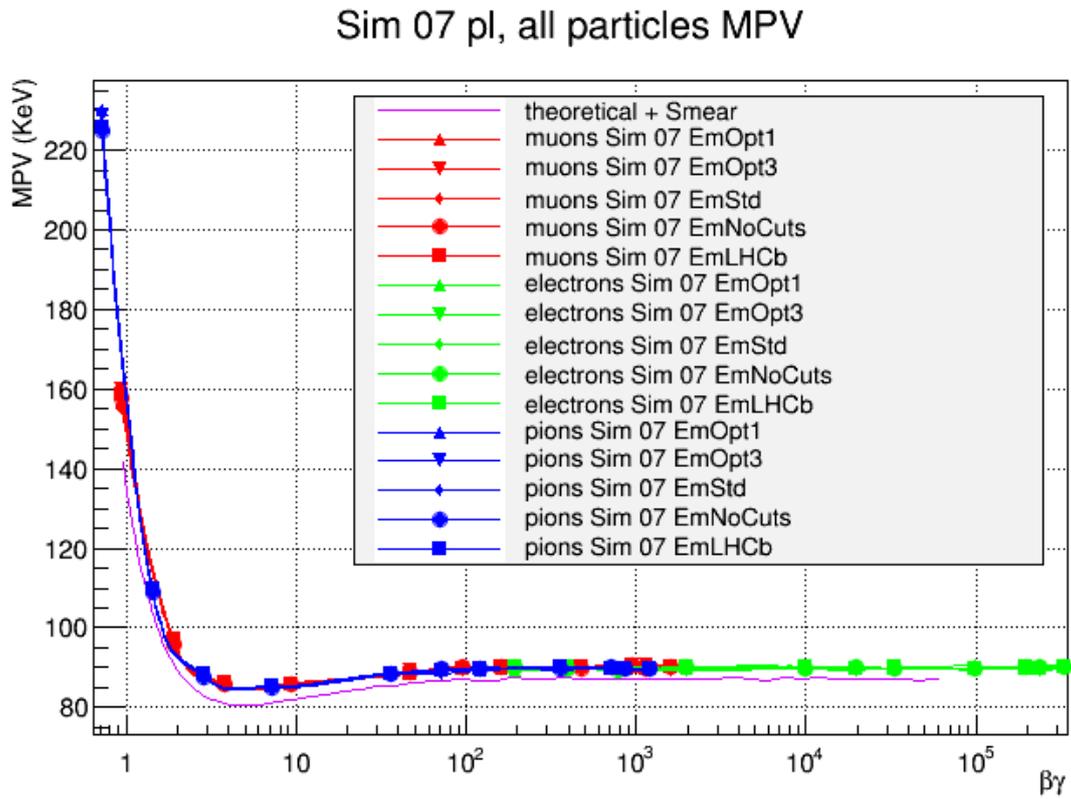
Figure 2.22: MPV comparison for v9.5 physics lists, $x > 300\mu m$.
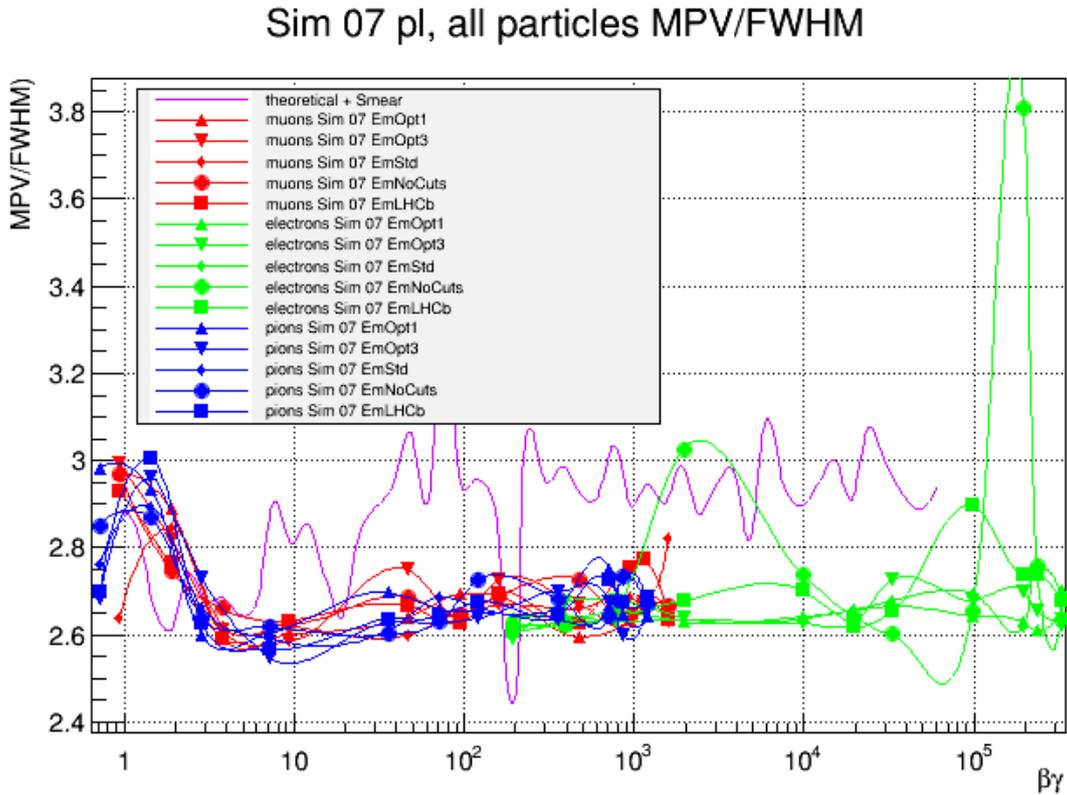
Figure 2.23: MPV/FWHM comparison for v9.5 physics lists, $x > 300\mu m$.

Geant4 v9.5 vs v9.6 results

The same series of tests was performed with the validation of GEANT4 v9.6, in which the results from v9.5 `EmNoCuts` were used as a reference, and these are shown in Figures 2.24-2.27. The dE/dx tests showed good consistency between v9.5 and v9.6. The limiting factor in many of the comparisons is the small size of the centrally produced v9.5 sample, causing fluctuations in the muon comparison. The `BremVeloCheck` tool was prototyped in the v9.6 validation studies and the results are included in Table 2.4 for completeness.
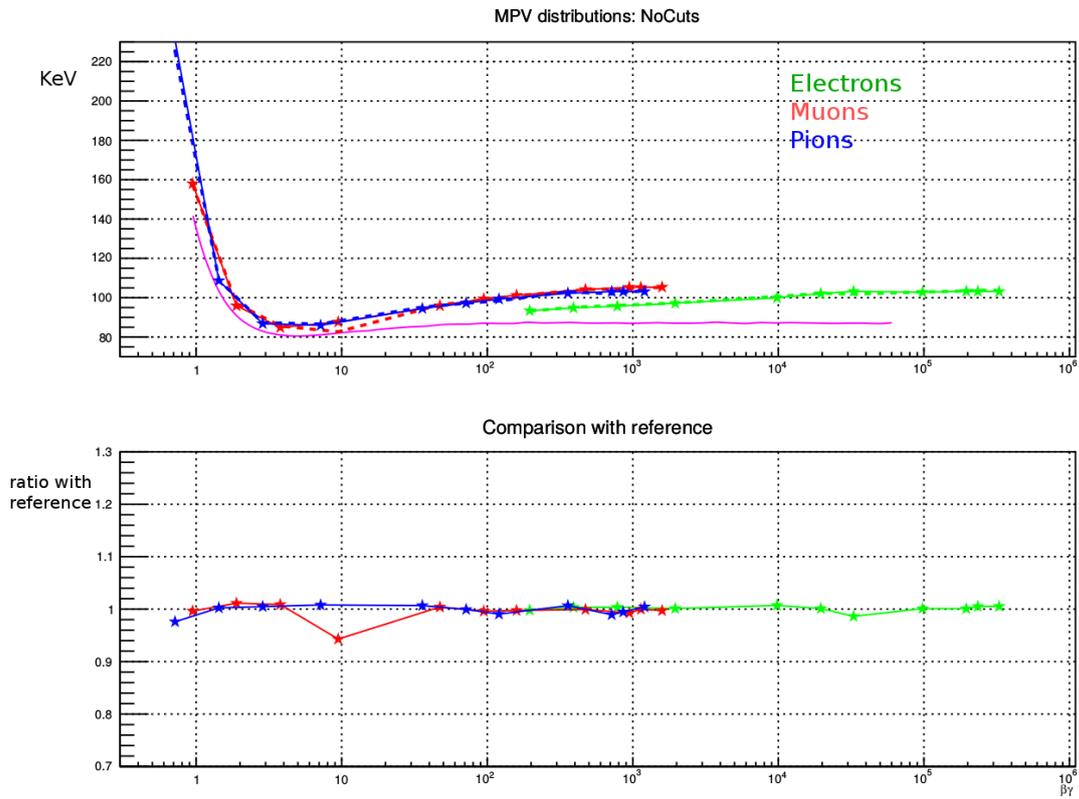
Figure 2.24: MPV comparison for v9.6 `EmNoCuts` and ratio with v9.5 `EmNoCuts` (bottom).
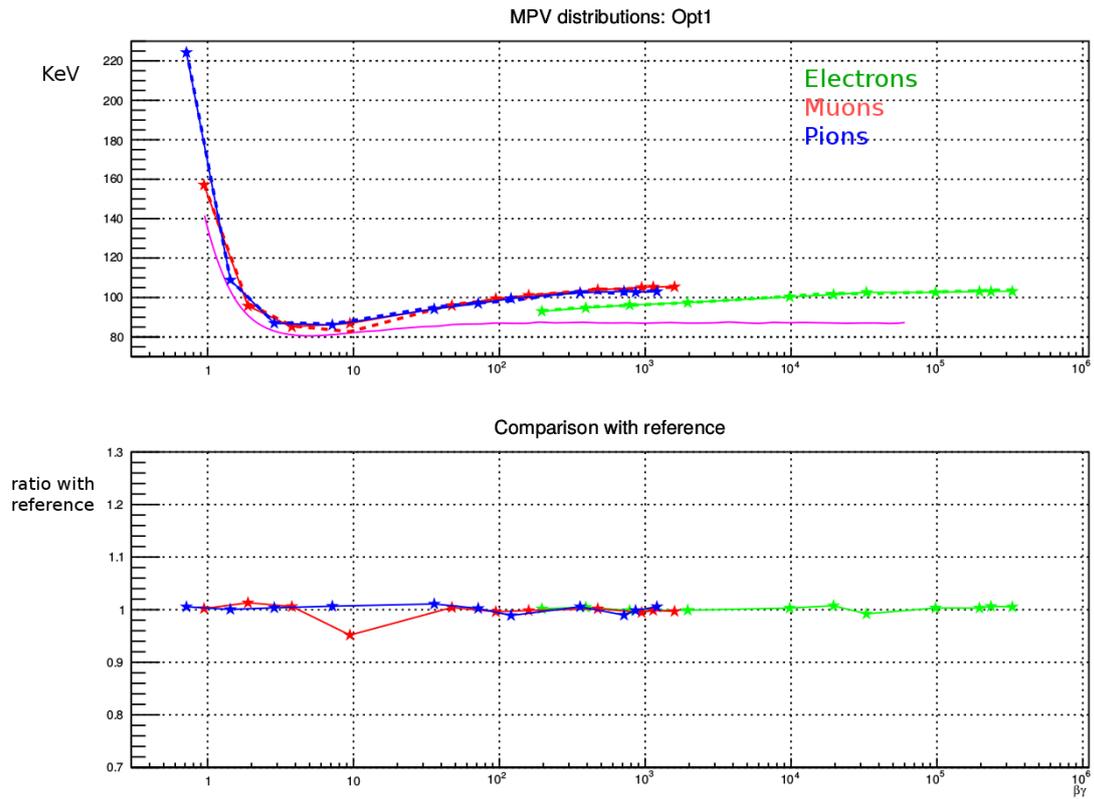
Figure 2.25: MPV comparison for v9.6 `EmOpt1` and ratio with v9.5 `EmNoCuts` (bottom).
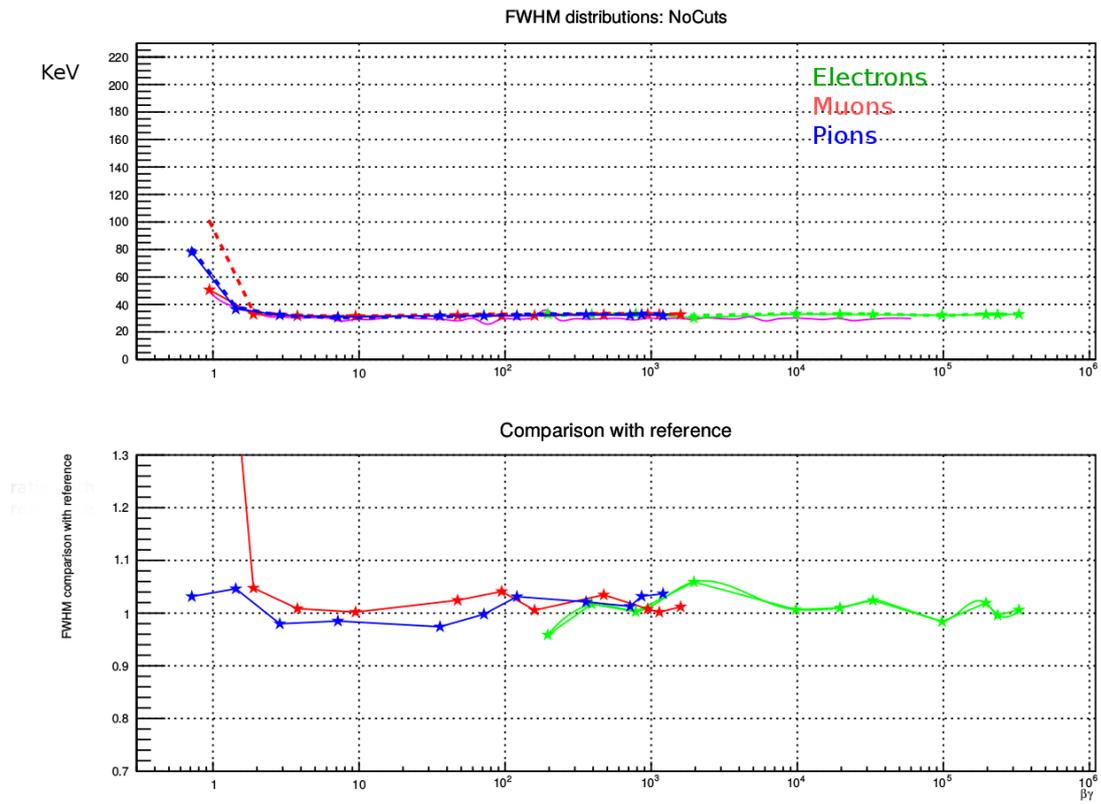
Figure 2.26: FWHM comparison for v9.6 `EmNoCuts` and ratio with v9.5 `EmNoCuts` (bottom).

List	Mean Energy	mean no. per track
EmNoCutsOld	1.431	5.00
EmOpt1	1.538	5.26
EmOpt2	1.555	5.28
EmLHCb	1.548	5.29
EmNoCutsNew	1.437	5.07

Table 2.4: Photon multiplicity and mean photon energy per track from 1 GeV electrons. Physics lists are v9.6
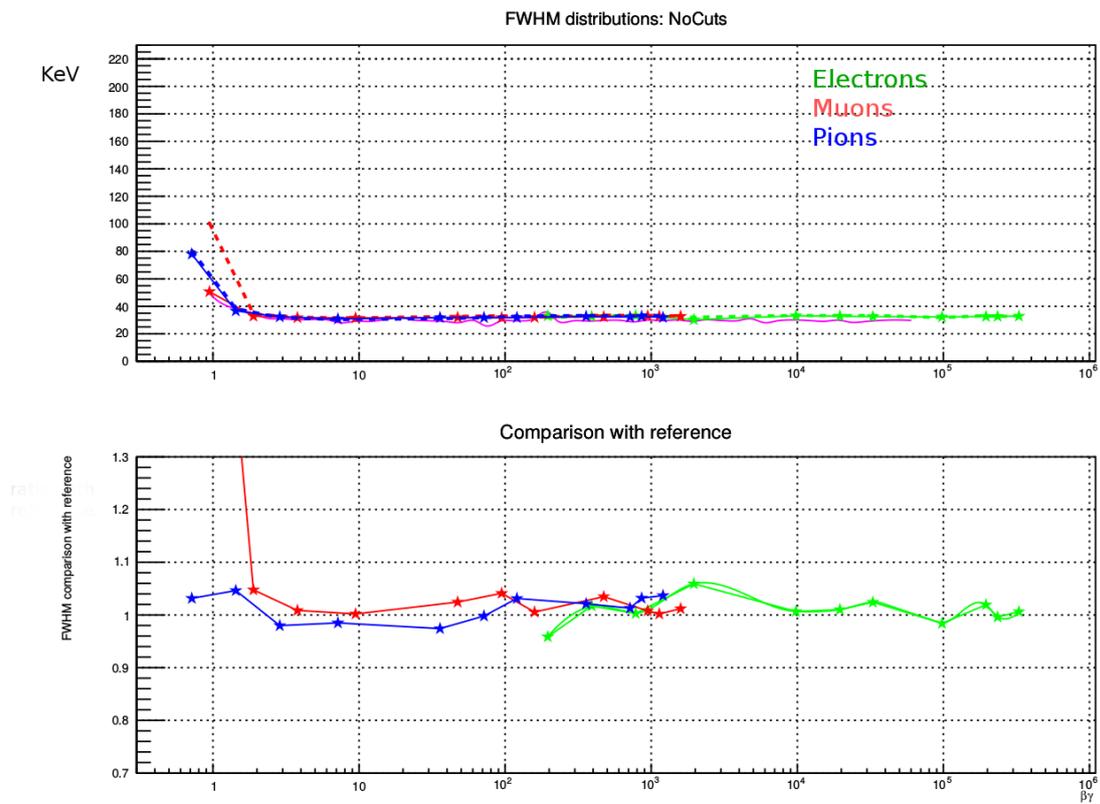
Figure 2.27: MPV comparison for v9.6 `EmOpt1` and ratio with v9.5 `EmNoCuts` (bottom).

Summary

A set of tools has been developed for systematic validation and testing of EM physics for LHCb, specifically for energy loss and bremsstrahlung radiation in the VELO silicon. In the latest validation run, a comparison of the various electromagnetic physics lists available in GEANT4 v9.6 have been compared in detail and cross-checked with v9.5 within the Gauss framework. Previously, the $dEdx$ tool was also used to validate the update from GEANT4 v9.4 to v9.5, and after further investigation of apparent disparities that became apparent from the studies, it is clear there is little difference in the dE/dx behaviour of these lists in silicon. The FWHM discrepancies in v9.5 `EmNoCuts` presentand a m an area for further study, and more accurate methods to determine the FWHM could be considered.

Future plans

Development of the validation tools will continue within the University of Birmingham LHCb group. The front-end of the EM validation software handles the submission of simulation jobs, analysis of the output files and the generation of figures and numerical results. It is developed in pure python code to allow seamless integration into LHCbPR, which employs the python web framework, Django. In LHCbPR, the tools will be run periodically and in future versions, automatically. Developing meaningful numerical comparisons for the minimum ionisation energy, mean bremsstrahlung photon energy and multiplicity would be practical for automatic comparison between simulation run conditions in LHCbPR.

Chapter 3

First observation and branching fraction measurement of the decay $\Lambda_b^0 \to pK\mu^+\mu^-$

3.1 Introduction

The rare decay $\Lambda_b^0 \to pK\mu^+\mu^-$ is a flavour changing neutral current decay, in which the quark level transition, $b \to s\mu^+\mu^-$ occurs through the SM processes of electroweak penguin and W box decays, as described in Section 1. This makes it sensitive to new physics. Figure 3.1 depicts the feynman diagrams for possible penguin and box decays.

There is much interest around the decays of Λ_b^0 particles due to several factors. As a baryon, the Λ_b^0 has non-zero spin, and can be used to probe the helicity structure of the underlying Hamiltonian [71,72]. The existence of a di-quark spectator system compared to the single spectator in similar meson decays means the hadronic physics is very different. Information gained from studying these decays may aid the treatment of hadronic physics in both the meson and baryon sectors
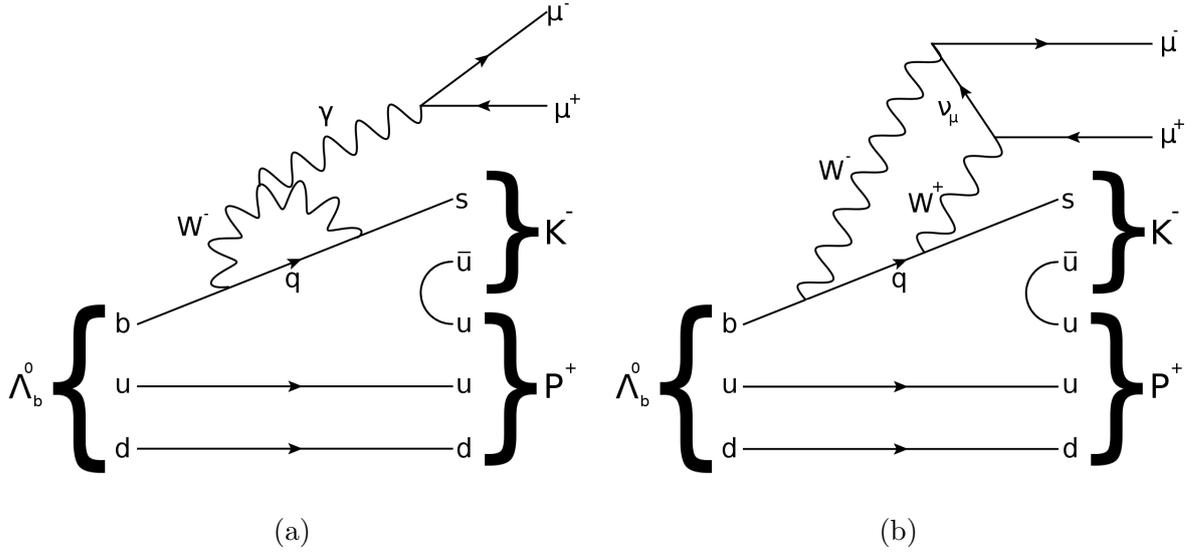
Figure 3.1: Penguin (a) and w-box (b) diagrams for the $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ decay via a $b \rightarrow s$ transition.

In this particular decay, the relatively high mass system of the final state hadrons restricts the Λ_b^0 to cascade down through numerous Λ^* resonances. The structure of the pK mass spectrum is not well known and $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ decays provide an opportunity for further study. The branching fraction measurement is performed in q^2, the square of the dimuon invariant mass, this allows comparison with theoretical predictions and parameters of the effective Hamiltonian. It is important to note that the high hadronic mass allows one to probe the lower q^2 range, below the J/ψ resonance. This region can have high sensitivity to new physics and offers complimentary measurements to rare Λ_b^0 decays that proceed via the ground state Λ^0 baryon, which are distributed in the higher q^2 range.

3.1.1   Signal blinding

As the total and q^2-differential branching fractions of $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ have not been measured to date, signal candidates in the data sample are blinded to reduce the potential

bias in the selection procedure. All variables related to the analysis are blinded in the $pK^-\mu^+\mu^-$ mass window between 5.5 and 5.7 GeV/c^2. Furthermore the signal shape, *i.e.* the overall fit shape within the blinding region and χ^2 value of the full fit, remains obscured. This analysis is currently under review by the LHCb Collaboration, after which unblinding and publication of the results will quickly follow.

The branching fraction of the resonant $\Lambda_b^0 \to J/\psi pK$ was recently measured by the LHCb Collaboration as $(3.17 \pm 0.04 \pm 0.07 \pm 0.34^{+0.45}_{-0.28}) \times 10^{-4}$ [73], where the uncertainties are statistical, systematic, and due to the knowledge of the normalisation channel $B^0 \to J/\psi K^*(892)^0$ and the relative production rate of Λ_b^0 to $\overline{B}^0$ ($f_{\Lambda_b^0}/f_d$). We use this decay as a control channel, performing the branching ratio measurement of,

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to J/\psi pK)} \tag{3.1}$$

to control systematic uncertainties belonging to both decays. The resonant mode, $\Lambda_b^0 \to J/\psi pK$ is also used to study kinematic distributions in data, to evaluate quantitatively the impact of possible mismodelling of data by simulation, and to investigate peaking backgrounds.

Measuring the branching fraction of $\Lambda_b^0 \to pK\mu^+\mu^-$ comes with several challenges. Currently, theoretical descriptions of the decay are not well developed. In $\Lambda_b^0 \to J/\psi pK$, the proton and kaon come predominantly through $\Lambda_b^0 \to (\Lambda^* \to pK)$ of which there are numerous Λ^* states. Predictions for $\Lambda_b^0 \to \Lambda^* l^+ l^-$ for several of these states have been calculated [74] but without detailed knowledge of the decay structure and amplitudes, these cannot be easily transformed into a prediction of $\mathcal{B}(\Lambda_b^0 \to pK\mu^+\mu^-)$. With the lack of predictions, the MC samples of $\Lambda_b^0 \to pK\mu^+\mu^-$ (and $\Lambda_b^0 \to J/\psi pK$) are produced with a phase-space model only and will not correctly describe the real decay structure, and as the analysis is blind, MC cannot be corrected for data-MC discrepencies either.

These challenges are tackled by performing an event by event correction of the observed signal candidates in data using a multi-dimensional efficiency model, while the production kinematics of the Λ_b^0 are corrected using a technique that is independent of the Λ_b^0 kinematics in $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\, pK$.

3.2   Simulation and Software

Monte Carlo samples are used in this analysis to evaluate efficiencies, to develop candidate selections as well as understand the shapes and yields of the signal, control channel and peaking backgrounds reconstructed under the $\Lambda_b^0 \to J/\psi\, pK$ or $\Lambda_b^0 \to pK\mu^+\mu^-$ hypotheses. The official LHCb samples used are listed in Table 3.1. The $\Lambda_b^0 \to \widetilde{\psi} pK$ sample involves a

Sample	Model
$\Lambda_b^0 \to pK\mu^+\mu^-$	phase space
$\Lambda_b^0 \to J/\psi\, pK$	phase space
$B^0 \to J/\psi\, K_s^0$	w. radiative
$B_s^0 \to J/\psi\, \phi$	w. radiative
$B^0 \to K\pi\mu\mu$	phase space
$B_s^0 \to KK\mu\mu$	w. transition form factors [75]
$\Lambda_b^0 \to \widetilde{\psi} pK$	custom

Table 3.1: Simulated samples used in this analysis

custom model that generates a high-q^2 pseudo-resonance, which is forced to decay into two muons. This is to avoid imprecise efficiency modelling at high q^2 that would otherwise arise from the small population of such candidates. The q^2 distributions of both phase-space $\Lambda_b^0 \to pK\mu^+\mu^-$ decays alone, and of the combination of phase space decays with the pseudo-resonance sample, are shown in Figure 3.2.

This sample is only used to construct the efficiency model introduced in Section 3.8. The unphysical q^2 distribution can be used for this purpose because the method largely removes the dependence of the calculated selection efficiency on the shape of the mass
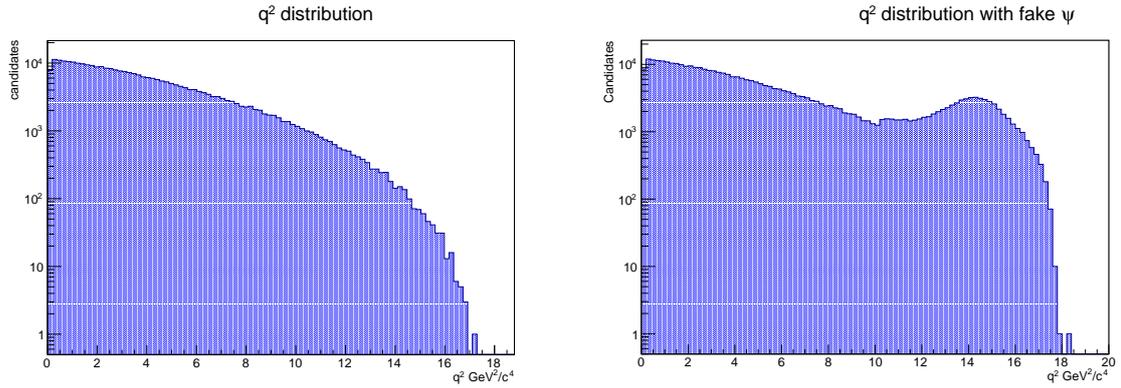
Figure 3.2: Distribution of q^2 in phase-space $\Lambda_b^0 \to pK\mu^+\mu^-$ MC (left) and the same sample combined with the pseudo-resonance sample, $\Lambda_b^0 \to \widetilde{\psi}pK$ (right)

distributions.

Specific packages used

Several LHCb-specific and third party software packages are used in this analysis. The n-tuples of candidates are created using `DaVinci` [76]. For geometric acceptance studies, generator level[1] MC is produced privately for $\Lambda_b^0 \to pK\mu^+\mu^-$, $\Lambda_b^0 \to J/\psi\, pK$ and $\Lambda_b^0 \to \widetilde{\psi}\, pK$ where events in a 'forward' 2π hemisphere are generated. These use the same parameters and models as the official productions and are produced using `Gauss` [77]. For the particle identification studies (Section 3.6.1), `Urania` with the `PIDCalib package` [78] is used. The kinematic weighting procedure (Section 3.2.1) uses the `python` analysis tool `Bender` [79]. For the neural network selection, the `Phi-T` software `NeuroBayes` [80] is used.

3.2.1 Monte Carlo weighting

The MC samples are generated under the phase-space hypothesis, and it is clear from Figure 3.3 that there are kinematic discrepancies between $\Lambda_b^0 \to J/\psi\, pK$ data and MC. These may give rise to inaccurate evaluations of the efficiency for both the kinematic and

---

[1]Simulation without any modelling of the material interactions, resolution or acceptance of the LHCb detector.

neural network selections. Due to correlations between kinematic variables, evaluating the efficiency as a function of q^2 alone cannot be assumed to be correct.

The two main causes of disagreement between MC and data are the modelling of the production of the Λ_b^0 baryon in MC, and the structure of the decay to the final state. The efficiency is parametrised in five kinematic dimensions as described in Section 3.6, to reduce dependence of the integrated efficiency on the simulated decay model.

Simulated, generator-level samples of $B^0 \to J/\psi K_S^0$ and $\Lambda_b^0 \to J/\psi pK$ decays, and the measured ratio of production rates of Λ_b^0 and $\overline{B}^0$, are used to correct potential discrepancies in the production kinematics of Λ_b^0 baryons.

The channel $B^0 \to J/\psi K_S^0$ is used because it is a relatively clean channel in data and as it is described accurately in MC with a fixed decay structure. We do however, re-weight the $B^0 \to J/\psi K_S^0$ for the small data-MC discrepancy. Figure 3.4 shows a comparison of $B^0 \to J/\psi K_S^0$ data and MC, illustrating a reasonable level of agreement. Each event in the generator level $B^0 \to J/\psi K_S^0$ MC is re-weighted by the ratio of data to MC after both reconstruction and selection have been carried out. This is done by parameterising the p_{T} and η distributions using kernel density estimation PDFs and evaluating the ratio of PDFs at the given p_{T} and η of the generator level MC event.

The procedure to correct the Λ_b^0 production kinematics is as follows: We determine for each Λ_b^0 MC event, the values of $f_{\Lambda_b^0}/f_d\,(p_{\mathrm{T}})$ (and $f_{\Lambda_b^0}/f_d(\eta)$) as measured by LHCb [81] and form an event-by-event weight by taking the product of this with the ratio of $B^0 \to J/\psi K_S^0$ and $\Lambda_b^0 \to J/\psi pK$ kinematics at the given point in p_{T} (η). More explicitly, this ratio is calculated by taking the PDFs corresponding to the p_{T} (η) distributions of $B^0 \to J/\psi K_S^0$ and $\Lambda_b^0 \to J/\psi pK$ at generator level and evaluating them at the p_{T} (η) value of the given Λ_b^0, and we take the ratio of these values. This gives two separate expressions, one for p_{T} and one for η. We correct simply by the product of these two expressions as the correlation is unmeasured in [81].
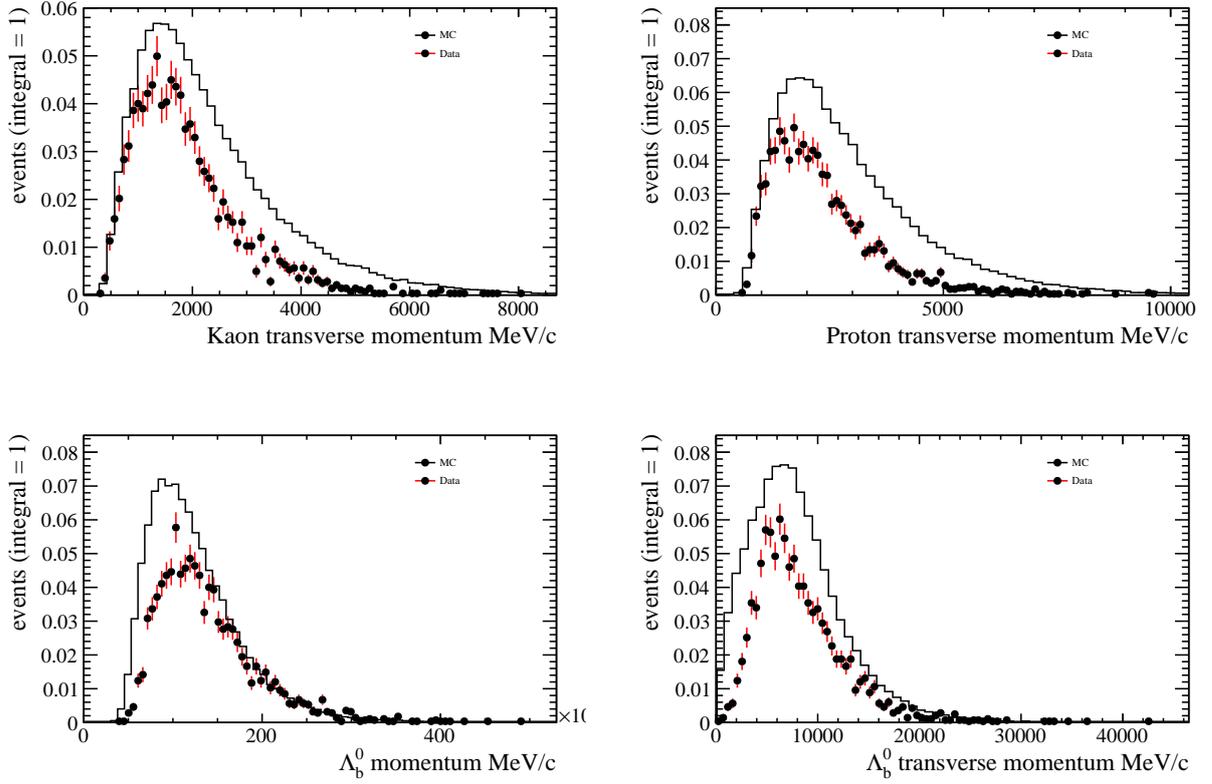
Figure 3.3: Data-MC comparison for $\Lambda_b^0 \to J/\psi\, pK$, PID selection of `h_ProbNNh` > 0.8 applied for each hadron species, `h`.

This gives the complete expression for the event-by-event weight of a candidate with transverse momentum, $p_{\mathrm{T}i}$ and pseudo-rapidity, η_i,

$$ w_i = \frac{f_{\Lambda_b^0}}{f_d}(p_{\mathrm{T}i}) \cdot \frac{PDF^{B^0 \to J/\psi\, K_{\mathrm{S}}^0}(p_{\mathrm{T}i})}{PDF^{\Lambda_b^0 \to J/\psi\, pK}(p_{\mathrm{T}i})} \cdot \frac{f_{\Lambda_b^0}}{f_d}(\eta_i) \frac{PDF^{B^0 \to J/\psi\, K_{\mathrm{S}}^0}(\eta_i)}{PDF^{\Lambda_b^0 \to J/\psi\, pK}(\eta_i)}, \qquad (3.2) $$

where $PDF^{B^0 \to J/\psi\, K_{\mathrm{S}}^0}(p_{\mathrm{T}i})$ represents the normalised PDF of the $B^0 \to J/\psi\, K_{\mathrm{S}}^0$, evaluated at the transverse momentum value of the Λ_b^0 candidate at generator level, $p_{\mathrm{T}i}$, and similarly for the corresponding PDF of the $\Lambda_b^0 \to J/\psi\, pK(p_{\mathrm{T}i})$ and for their equivalent PDFs as a function of η.

For the value of $f_{\Lambda_b^0}/f_d$ as a function of p_{T} and η, we take values according to the

(a)

(b)
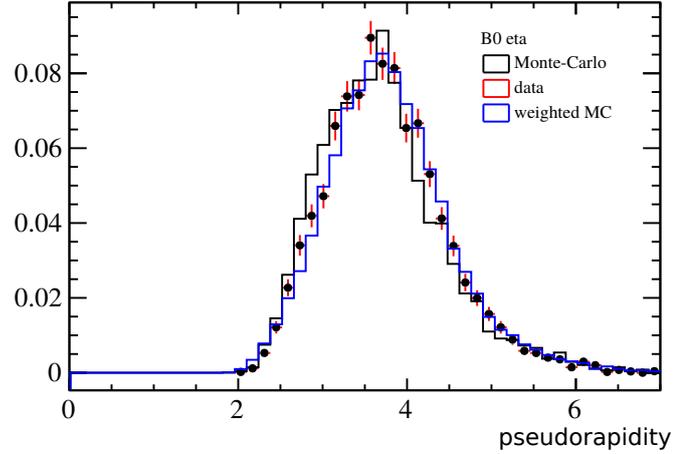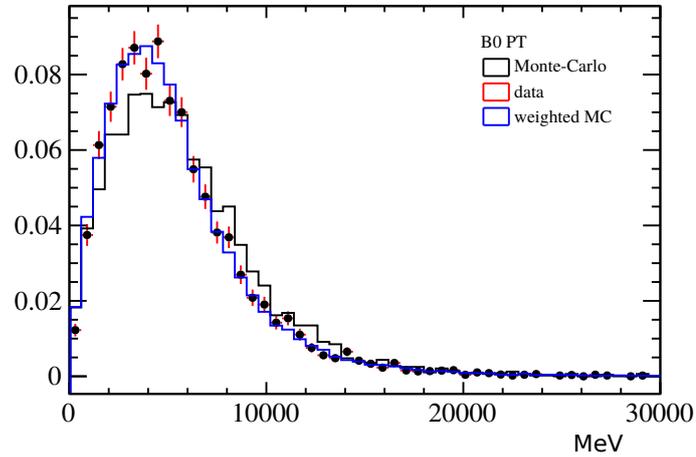
Figure 3.4: p_{T} and η MC-data comparison for B^0 using the $B^0 \to J/\psi K_{\mathrm{S}}^0$ channel. MC is reweighted by $\frac{data}{MC}$

expressions shown in Equations 3.3 and 3.4. We normalise each weight to preserve $\sum_{i=0}^{N}(w_i) \equiv N$, where N is the total number of candidates in the sample. The kinematic re-weighting is performed on MC candidates after stripping selection, and when applied to $\Lambda_b^0 \to J/\psi\, pK$ MC gives the best overall agreement with $\Lambda_b^0 \to J/\psi\, pK$ candidates in data.

The $B^0 \to J/\psi K_{\mathrm{S}}^0$ candidates, which have background substracted using the sPlot technique [82] are used to characterise the kinematics of the B^0 mesons produced at LHCb.

With loose selection to reduce excessive kinematic bias, we thus know the distributions of pseudorapidity and transverse momentum for the B^0 mesons. In [81], the behaviour of $f_{\Lambda_b^0}/f_d$ as a function of pseudo-rapidity and transverse momentum is found to be,

$$\frac{f_{\Lambda_b^0}}{f_d}(\eta) = \overbrace{\left(0.834^{+0.064}_{-0.067}\right)}^{\mathcal{S}}\left[\left(0.464 \pm 0.003^{+0.008}_{-0.010}\right) + \left(0.081 \pm 0.005^{+0.013}_{-0.009}\right) \times (\eta - 3.198)\right],$$

$$(3.3)$$

$$\frac{f_{\Lambda_b^0}}{f_d}(p_{\mathrm{T}}) = \overbrace{\left(0.834^{+0.064}_{-0.067}\right)}^{\mathcal{S}}\left[\left(0.181 \pm 0.018 \pm 0.026\right) \right. \qquad\qquad (3.4)$$

$$\left. + \exp\left\{\left(-0.391 \pm 0.023^{+0.069}_{-0.067}\right) + (-0.095 \pm 0.007 \pm 0.014) \times p_{\mathrm{T}}(\mathrm{GeV})\right\}\right],$$

where $\mathcal{S}$ is a constant scale factor. This enables us to relate the kinematics of the Λ_b^0 to the kinematics of the B^0 through the p_{T} and η dependent $f_{\Lambda_b^0}/f_d$ relationships. These distributions are shown in Figure 3.5, taken from Ref. [81].

This method is also used to evaluate the ratio of B^0 and Λ_b^0 generator-level kinematics in the calculation of the event weight (Equations 3.3 and 3.4).

The loose cuts presented in Table 3.2 are applied to the B^0 MC and data samples to ensure a sufficiently pure sample of $B^0 \to J/\psi K_{\mathrm{S}}^0$ candidates is obtained in data, without significantly biasing the kinematics of the B^0. Only LL tracks are selected while the trigger and stripping selections remain the same as those used for $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi pK$ decays (Tables 3.3 and 3.5). The samples used for this correspond to $29\,472$ (post-reconstruction and stripping) MC events, $20\,000$ generator level events, and 5870 B^0 candidates in data.
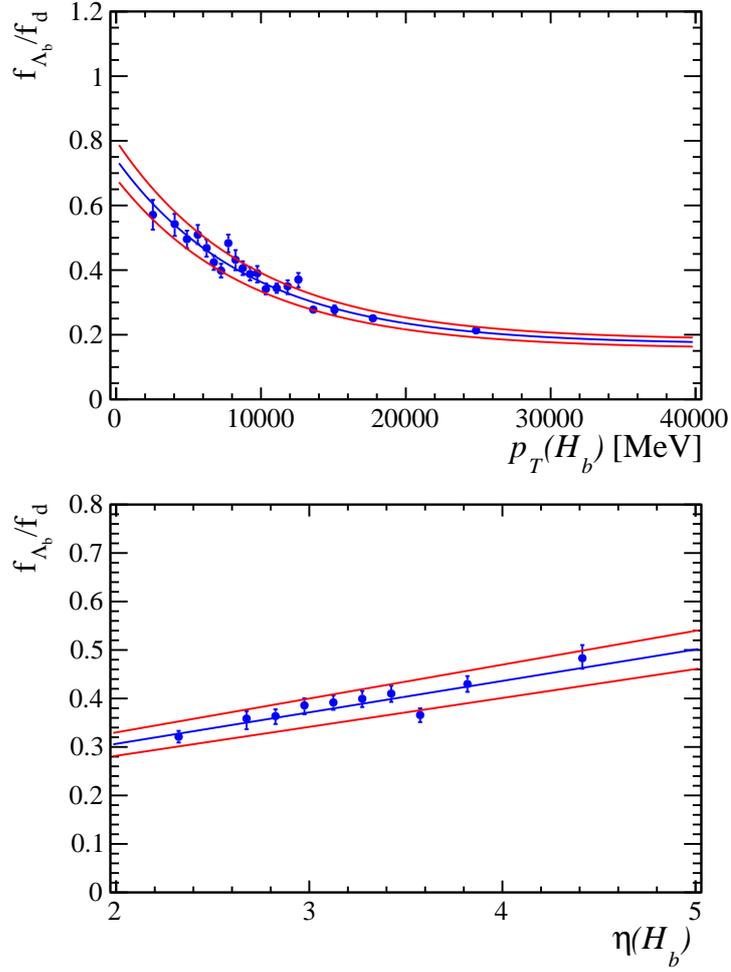
Figure 3.5: (Included from Ref. [81] for completeness.) Hadronic $f_{\Lambda_b^0}/f_d$ vs. p_{T} (a) and η (b) dependencies, where the data points and blue fit functions are scaled by the factor $\mathcal{S}$, the red lines indicate the upper and lower bounds of the total uncertainty on $\mathcal{S}$. The uncertainties on the vertical axes are the combined statistical and systematic errors of efficiency-corrected yields. The uncertainties on the horizontal axes are the standard error on the mean (and are too small to be visible).

Particle	Variable	Requirement
B^0	Mass	$5.17 < m < 5.4$ GeV
Daughter hadrons	Track type	3
	Mass	< 5.05 GeV

Table 3.2: Cuts to B^0 candidates after stripping

3.2.2   Model validation

Candidate $\Lambda_b^0 \to J/\psi pK$ decays in data are used to evaluate the result of re-weighting the MC kinematics.

As this procedure is only correcting for MC discrepancies in the production kinematics of the Λ_b^0 baryon, we do not expect perfect agreement between data and the re-weighted MC as these samples are generated using a generic phase-space model, and we anticipate significant differences in the decay structure. Figure 3.7 shows the Λ_b^0 p_{T} for a small slice in the pK mass around the $\Lambda(1520)$ resonance. There is a small improvement in the description of data by simulation, as expected when considering a restricted mass interval.
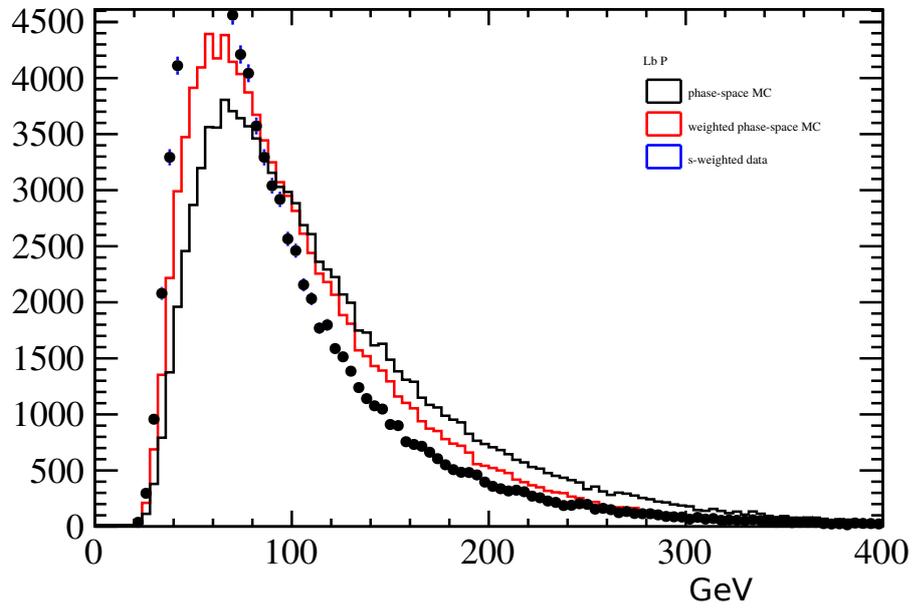
(a)
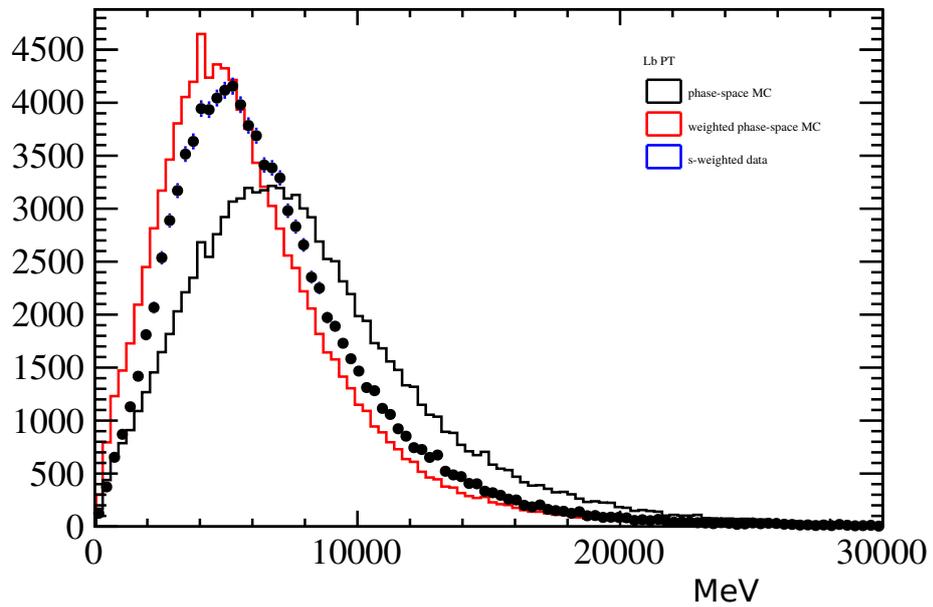
(b)

Figure 3.6: Comparison of 2012 data with both weighted and default simulation for $\Lambda_b^0 \to J/\psi\, pK$ candidates, as a function of Λ_b^0 (a) momenta, and (b) p_{T}.
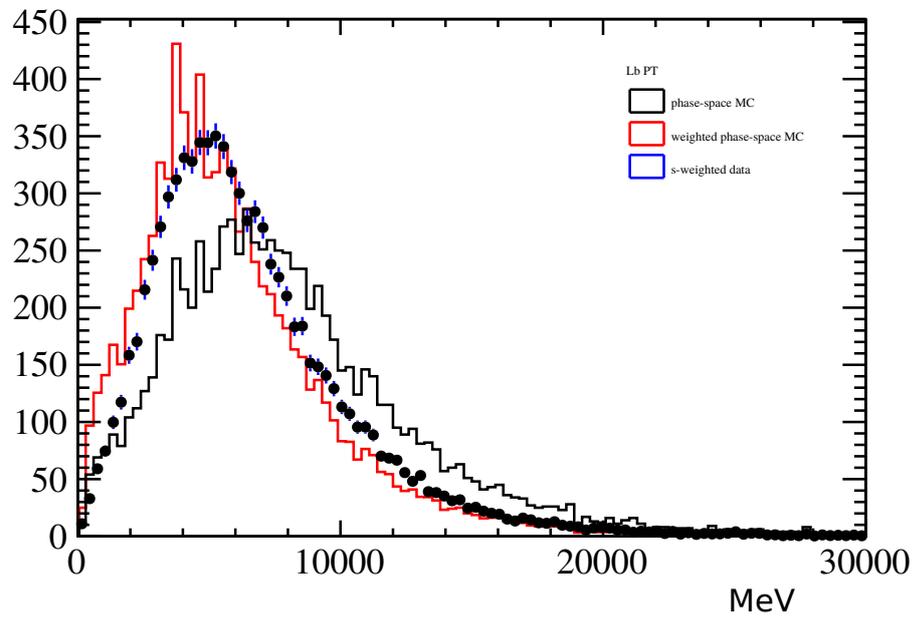
Figure 3.7: Comparison of Λ_b^0 baryon p_{T} in 2012 data with both weighted and default simulation, for candidates in which the invariant mass of the pK is in the interval 1450–1600 MeV/c^2.

3.3   Selection

We analyse data collected by LHCb in 2011 and 2012 corresponding to approximately $3\,\mathrm{fb}^{-1}$ in total, with $1\,\mathrm{fb}^{-1}$ collected in 2011 and the further $2\,\mathrm{fb}^{-1}$ in 2012. The DIMUON dataset is used, run through the stripping line B2XMuMu. Stripping versions 20r1 and 20 are used for the 2011 and 2012 data sets respectively, in which the stripping selection is the same in both versions. We use the same stripping selection for $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,pK$ samples, the requirements of which are listed in Table 3.3. The $\Lambda_b^0 \to pK\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi\,pK$ channels have almost identical selection criteria to reduce systematic effects. The two channels are distinguished from each other by their q^2 values.

We search for $\Lambda_b^0 \to pK\mu^+\mu^-$ in the q^2 range of 0.1–17.5 GeV^2/c^4 and select decays in the $pK^-\mu^+\mu^-$ final state, predominantly proceeding through through the J/ψ and $\psi(2S)$ $c\bar{c}$ states. We also restrict the pK^- invariant mass to be below $4.6\,\mathrm{GeV}/c^2$. This exludes a small region of phase space where it is difficult to parameterise efficiency as our simulation does not produce sufficiently many events.

For $\Lambda_b^0 \to pK\mu^+\mu^-$ decays, candidates with q^2 values in the regions $q^2 : [8.0, 11.0]\,\mathrm{GeV}^2$ and $q^2 : [12.5, 15.0]\,\mathrm{GeV}^2$ are rejected to veto resonant J/ψ and $\psi(2S)$ modes. For $\Lambda_b^0 \to J/\psi\,pK$, the J/ψ mode is exclusively selected within the $q^2 : [8.0, 11.0]\,\mathrm{GeV}^2$ window.

We consider the rare decay q^2 region in five bins: three below the J/ψ exclusion with the boundaries [0.1, 2.0, 4.0, 8.0] GeV^2/c^4, one between the J/ψ and $\psi(2S)$ from [11.0–12.5] GeV^2/c^4 and one above the $\psi(2S)$ at [15–17.5] GeV^2/c^4.

Additionally, the pre-selection criteria listed in Table 3.4 are applied after stripping. The decision to reject low p_T protons is a result of preliminary studies into the separation of signal and background in kinematic variables, and was investigated with a multi-layer perceptron (MLP) neural network (NEUROBAYES). A $\Lambda_b^0 \to pK\mu^+\mu^-$ MC sample was used for signal, and for background, the upper side-band of data from the B2XMuMu stripping line, with no further cuts applied to either sample. This allows a clear visualisation on

Subject	Variable	Requirement
Global	SPD hits	< 600
Tracks	χ^2/nDof	< 5.0
	Ghost probability	< 0.4
	min PV IPχ^2	> 9
Λ_b	Mass	$4.6 < m < 7.0$ GeV
	BDIRA	> 0.999968
	IP χ^2	< 9.0
	FD χ^2	> 100.0
	Vtx χ^2/DOF	< 8.0
Muons	PIDmu	> -3
	isMuon	True
	IP χ^2	> 16.0
Dimuon	Vtx χ^2/DOF	< 9.0
	Mass	< 7.1 GeV

Table 3.3: Summary of stripping requirements in B2XMuMu

Particle	Variable	Requirement
Λ_b	Vtx χ^2/DOF	< 5.0
Proton	p_T	> 500 MeV
Dimuon	q^2	< 17.5 GeV^2

Table 3.4: Post-stripping selection criteria.

the separation of signal and background, as seen by the NN. The separation in proton p_T can be seen in Figure 3.8. The trigger selection used is shown in Table 3.5 and we require 'trigger on signal' (TOS), meaning the the trigger has fired on a reconstructed signal candidate.

3.3.1 PID selection

The particle identification system of LHCb is used to suppress peaking backgrounds that appear through the misidentification of one or more of final state particles. The studies conducted in Section 3.4 suggest that both $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ decays are present in the post-stripping selection of $\Lambda_b^0 \to J/\psi pK$ candidates.
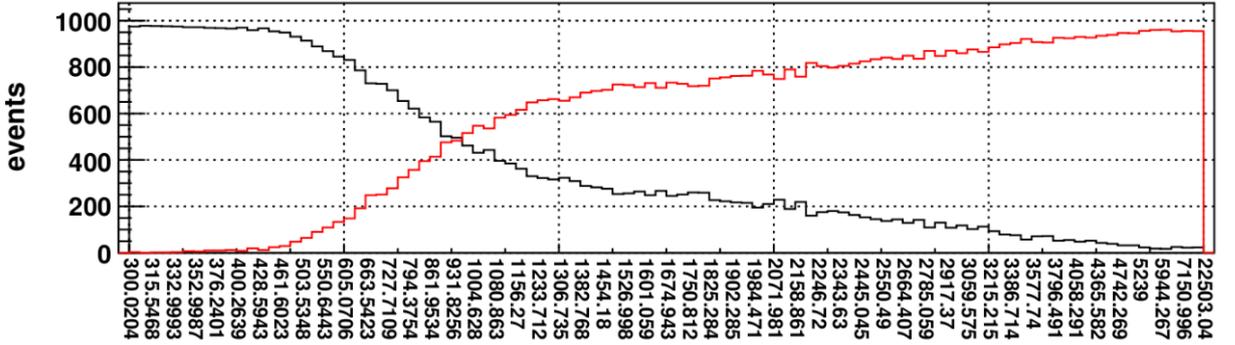
Figure 3.8: Signal-background separation in proton p_{T}, after stripping with signal in red and background in black.

Trigger Level	Lines
L0	L0MuonDecision_TOS
	L0DiMuonDecision_TOS
Hlt1	Hlt1DiMuonHighMassDecision_TOS
	Hlt1MuTrackDecision_TOS
	Hlt1TrackAllL0Decision_TOS
	Hlt1TrackMuonDecision_TOS
Hlt2	Hlt2Topo2BodyBBDTDecision_TOS
	Hlt2Topo3BodyBBDTDecision_TOS
	Hlt2Topo4BodyBBDTDecision_TOS
	Hlt2TopoMu2BodyBBDTDecision_TOS
	Hlt2TopoMu3BodyBBDTDecision_TOS
	Hlt2TopoMu4BodyBBDTDecision_TOS
	Hlt2SingleMuonDecision_TOS
	Hlt2DiMuonDetachedDecision_TOS

Table 3.5: Summary of trigger lines used by the analysis.

The $B^0 \to J/\psi K^{*0}$ decays in which $K^* \to K^+\pi^-$ may be selected as signal decays through misidentification of the pion as a proton, or through double misidentification of the pion as a kaon and the kaon as a proton. $B_s^0 \to J/\psi\phi$, in which $\phi \to K^+K^-$ can have one of the kaons misidentified as a proton, while $\Lambda_b^0 \to J/\psi\, pK$ itself can be mis-reconstructed through the double misidentification of the proton and kaon as each other. We therefore expect the decays of $B^0 \to K^{*0}\mu^+\mu^-$, $B_s^0 \to \phi\mu^+\mu^-$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ with the same misidentification of the final state hadrons to appear in the selection of

$\Lambda_b^0 \to pK\mu^+\mu^-$ candidates.

The final PID cuts were chosen after conducting several studies into their effectiveness. From the neural network studies noted previously, we see that there is an abundance of pions in the combinatorial background, leading to significant $\pi \to p$ misidentification. Therefore a cut of `p_probNNpi` < 0.7 was chosen to suppress this.

Symmetric cuts on the proton and kaon `ProbNN` variables are then applied to mainly suppress background from $p \to K$ and $K \to p$ mis-identifications, as shown in Table 3.6.

Particle	cut
p	`ProbNNp > 0.2 & ProbNNK < 0.8 & ProbNNpi < 0.7`
K	`ProbNNK > 0.2 & ProbNNp < 0.8`

Table 3.6: Chosen PID variable cuts for $\Lambda_b^0 \to pK\mu^+\mu^-$

3.3.2 Multivariate selection

Finally, a multivariate analysis is used to discriminate between signal candidates and combinatorial background. This is achieved with a NEUROBAYES neural network. We classify a signal and background sample for the network training. The signal sample is a selection of truth-matched $\Lambda_b^0 \to pK\mu^+\mu^-$ events consisting of 12500 candidates, and a background sample of comparable size. The MC sample includes simulation of detector and trigger, and event reconstruction. We apply all the selection criteria described above. For the background sample, we take events from data after the full $\Lambda_b^0 \to pK\mu^+\mu^-$ selection, with the additional criteria of $m(pK^-\mu^+\mu^-) > 6\,\mathrm{GeV}/c^2$. This is to give a pure background sample from the high mass side-band.

We do not include PID variables in the training due to the discrepancies between MC and data for the PID variables. The multivariate selection for $\Lambda_b^0 \to pK\mu^+\mu^-$ is performed on $\Lambda_b^0 \to pK\mu^+\mu^-$ MC with stripping and pre-selection cuts applied without any additional

cuts applied to the q^2 mass window. The MC is truth-matched via the `BKGCAT`[2] variable and requiring the true ID's of all particles. The data sample has the same pre-selection applied but with the additional high mass side-band cut of $6\,\mathrm{GeV}/c^2$ as mentioned above.

Due to the kinematic MC-data disagreements, the MC sample is weighted via the process explained in Section 3.2.1 before performing the training. This process only accounts for the production kinematics of the Λ_b^0 baryon, with the 5-D parameterisation of mass and angular variables accounting for the decay structure (see Section 3.8). The NN output weight for signal and background samples, as well as signal purity as a function of NN cut are shown in Figure 3.9. The rankings of each variable are shown in Table 3.7.
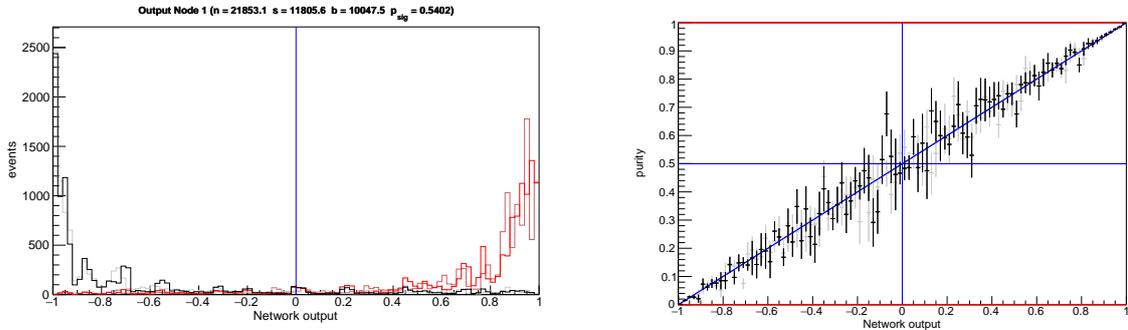
Figure 3.9: (a) NN output for signal (red) and background (black) samples. (Fainter lines for pre-boosted output). (b) Signal purity as a function of NN cut

Variables 9–12 are automatically pruned by NEUROBAYES and only the first eight are used. The comparison between MC and data for $\Lambda_b^0 \to J/\psi\, pK$ for these eight variables are shown in Figure 3.12

Applying the network assigns a weight between -1 and $+1$ to each candidate, depending on how signal-like it is perceived to be. To optimise the cut on this weight at which we select candidates, we scan the variable in increments of 0.01 and evaluate the Punzi figure of merit, $\frac{\varepsilon}{\sqrt{B}+a/2}$ [83], where ε and B are the signal efficiency and the expected number of background decays, and a is the target significance; a value of $a = 5$ is used. This is an

---

[2] categorises MC candiates into signal, full reconstructed background, partially reconstructed background, etc.

Phi-T

NeuroBayes® Teacher

correlation matrix of input variables
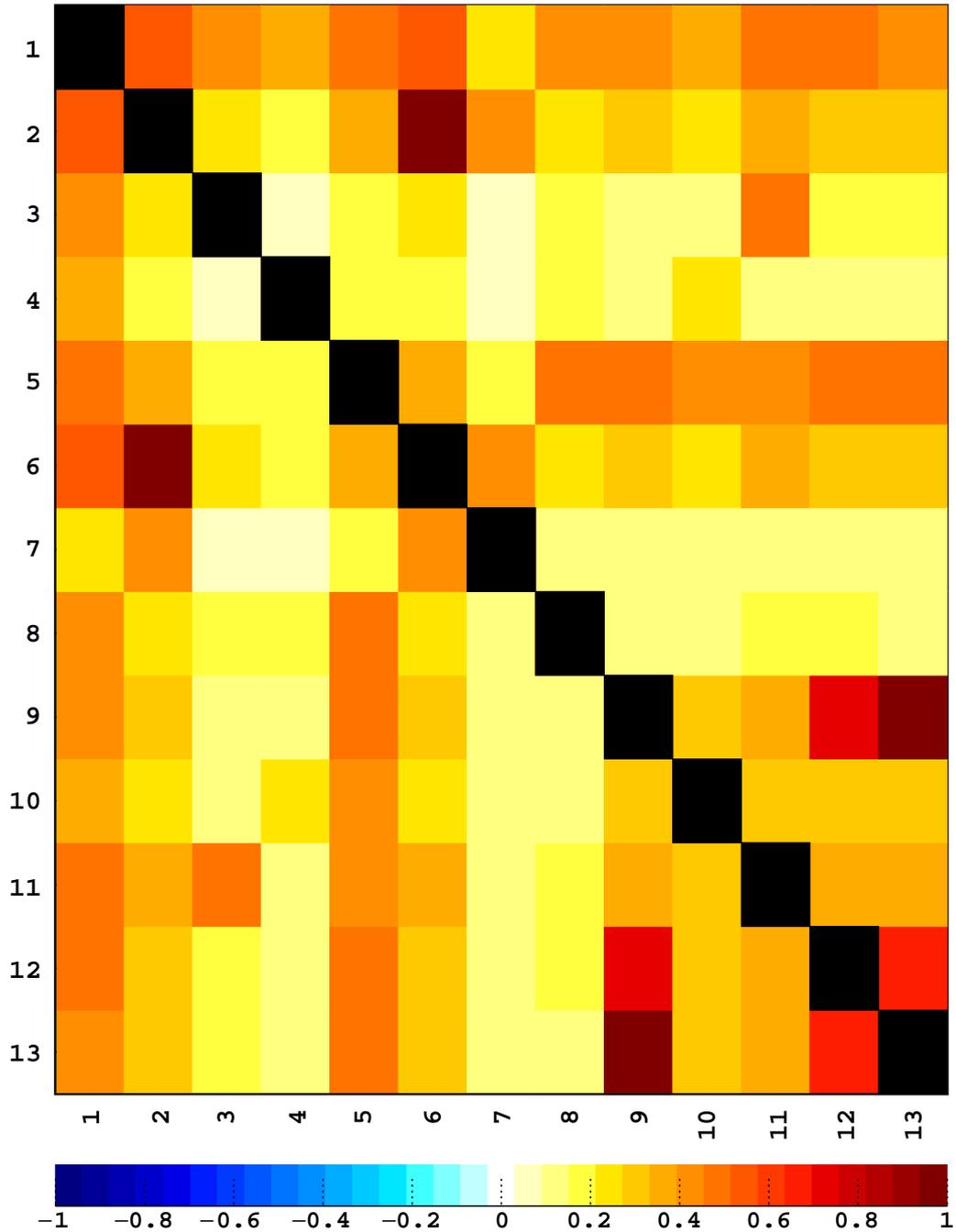
Figure 3.10: Correlation matrix between input variables

Input	rank	id	added	isolated	lost
Λ_b^0 end vertex χ^2	1	6	82.23	82.23	4.49
$\min(\mu\ \chi^2_{\mathrm{IP}})$	2	12	46.65	70.58	14.22
proton p_{T}	3	4	39.04	57.54	31.60
kaon p_{T}	4	3	36.49	60.69	20.64
$\Lambda_b^0\ \chi^2_{\mathrm{IP}}$	5	8	30.96	63.83	25.52
kaon χ^2_{IP}	6	11	19.39	73.88	14.57
$\max(\mu\ \chi^2_{\mathrm{IP}})$	7	13	12.29	64.89	3.44
proton χ^2_{IP}	8	10	8.70	51.51	7.85
Λ_b^0 Dec. Tree Fit χ^2	9	2	1.94	81.77	1.98
Λ_b^0 DIRA	10	5	1.77	74.21	1.76
dimuon end vtx χ^2	11	7	0.11	31.48	0.12
dimuon χ^2_{IP}	12	9	0.07	66.72	0.07

Table 3.7: Summary of inputs to the neural network. Rank gives internal ranking of the input according to NEUROBAYES (based on adds column). The column 'id' is the index of variables used in the correlation matrix, as shown in Fig. 3.10 'Added' shows the correlation significance added by that particular variable. 'Isolated' is the power provided by the given variable alone and loss shows how much information is lost when removing this particular variable.

appropriate figure of merit because it does not depend on the signal to background ratio to determine the optimal working point for the selection, given there are no measurements or predictions for the $\Lambda_b^0 \rightarrow J/\psi\, pK$ or $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ production rates. The results of the optimisation can be seen in Figure 3.11, from which the requirement on the neural network response of 0.91 was chosen.
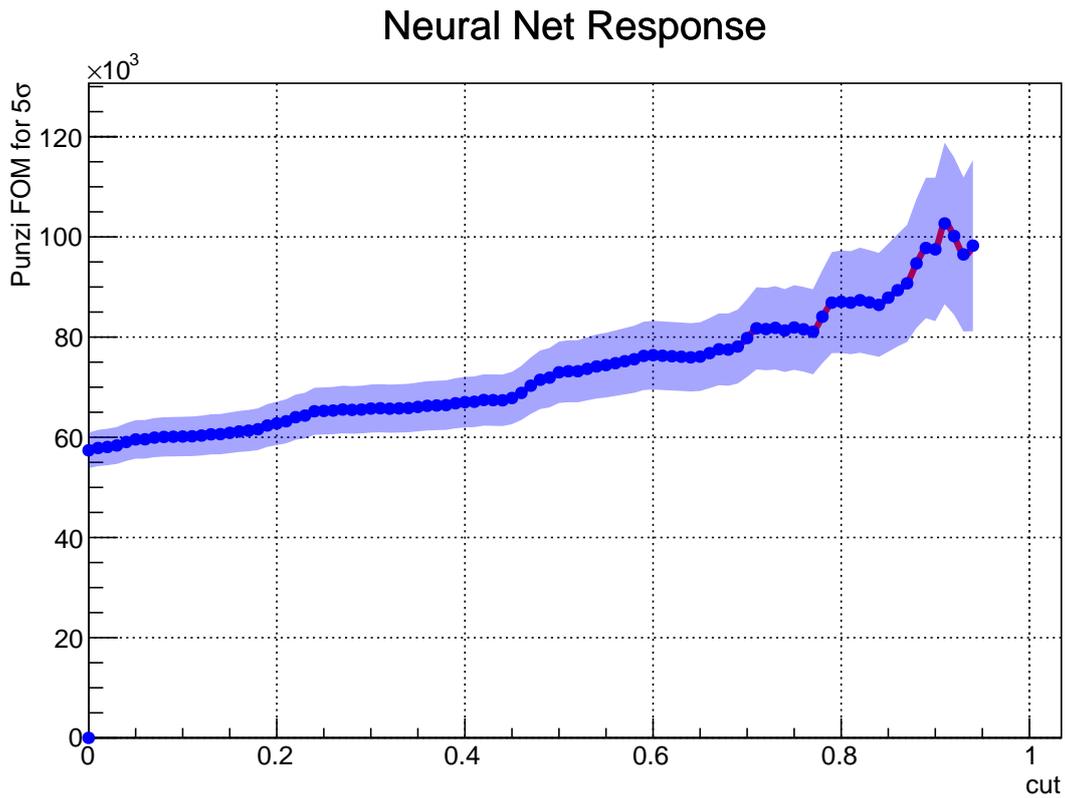
Figure 3.11: Figure of merit as a function of neural network response. Despite the large uncertainty on the figure of merit when the cut is tight, we choose the cut from the peak at 0.91. The fluctuations at higher cut values are a result of insufficient candidates remaining in the side band sample with which to perform a stable fit.
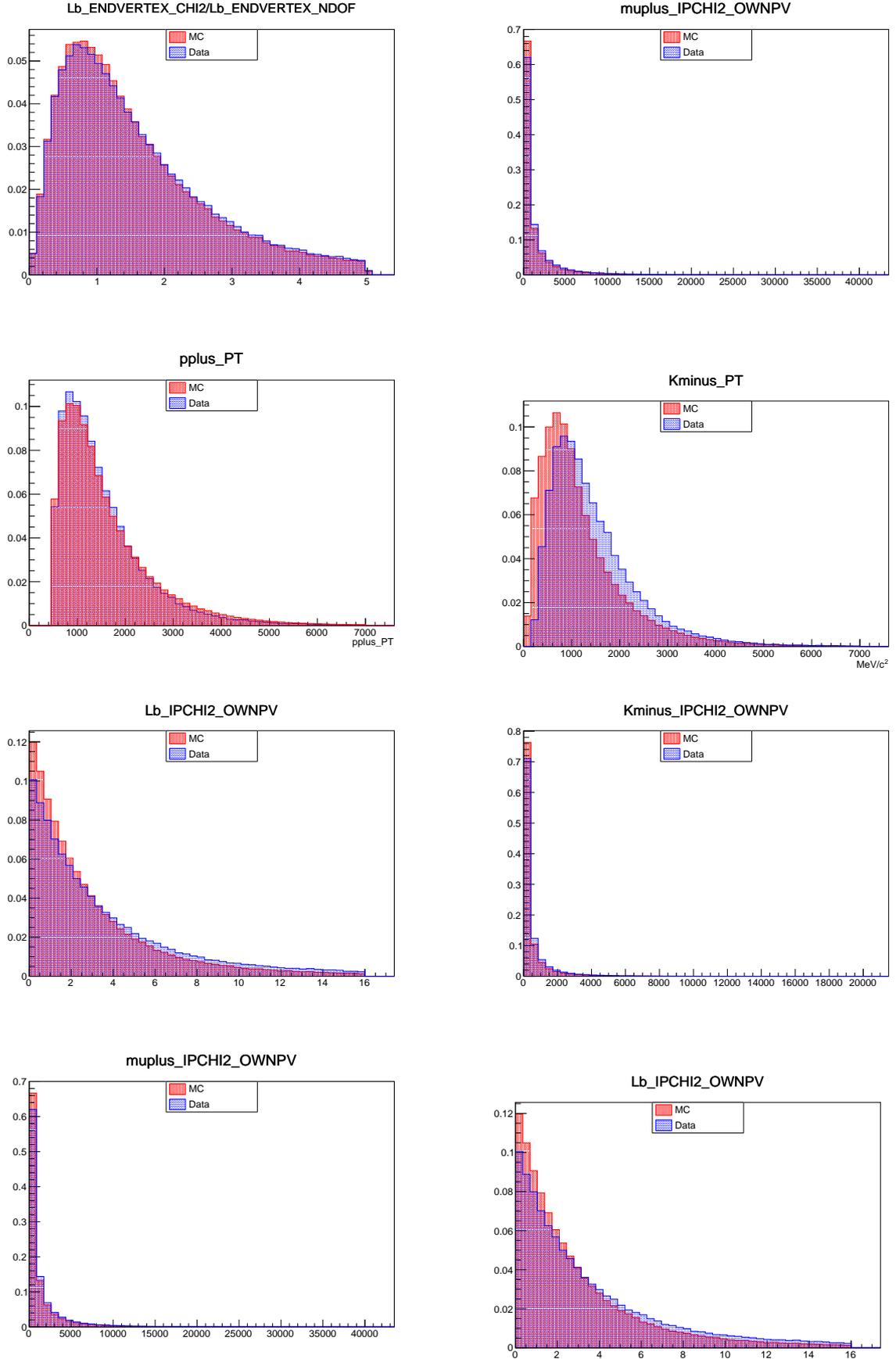
Figure 3.12: Comparison between MC and data of the variables used in the training on the neural network for $\Lambda_b^0 \to J/\psi pK$. We split the muon χ^2 variables into minimum and maximum values for each event. For these comparison plots, the μ^+ and μ^- variables are shown individually

3.4   Peaking Backgrounds

To investigate which decays happen to be mis-reconstructed and contribute to the background of $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$, we select a sub-sample of the data using the $\Lambda_b^0 \to J/\psi pK$ selection in q^2. We then search for peaks in the mass spectrum of the reconstructed final state system under various combinations of mass hypotheses of the individual final state particles. When exchanging the mass of reconstructed proton tracks with the kaon mass, a peak around the nominal B_s^0 mass is clearly visible. Similarly, with $[p \to \pi]$ or $[p \to K, K \to \pi]$ reflections, we see peaking at the B^0 mass. It is clear there is a significant contribution from $B_s^0 \to KK\mu\mu$ and $B^0 \to K\pi\mu\mu$ decays, primarily from $B^0 \to J/\psi K^*$ and $B_s^0 \to J/\psi \phi$, with their respective non-resonant modes assumed to also be present in the $\Lambda_b^0 \to pK\mu^+\mu^-$ due to the identical final states and similar kinematics. Additionally, the double reflection $[p \to K, K \to p]$ will result in mis-reconstructed $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$. Figure 3.13 shows a sample of the $\Lambda_b^0 \to J/\psi pK$ data sample with pre-selection applied, under different mass hypotheses.

We also check for the possibility of partially constructed $\Lambda_b^0 \to \Lambda_c^+ \mu\nu$ decays in $\Lambda_b^0 \to pK\mu^+\mu^-$ where the Λ_c^+ decays to $pK\pi$ and the pion is misidentified as a muon. This is done by swapping one of the muon mass hypotheses for a pion in a blinded data sample with the $\Lambda_b^0 \to pK\mu^+\mu^-$ selection applied and seeing if a peak at $M_{\Lambda_c^+}$ (2.29 GeV) is observed in the $pK\pi\mu$ mass interval in the combined upper and lower mass side-band of $pK\mu\mu$. We can see a clear contribution to the background from the Λ_c^+ in Figure 3.14, but looking at where these candidates lie in the $pK\mu\mu$ mass frame (Figure 3.15), it is clear that they lie far from the Λ_b^0 and we expect the contribution to be negligible, thus no further PID selections or mass cuts are considered necessary for Λ_c^+ suppression. We rely predominantly on the PID selection to reduce contamination from misidentified final states and good modelling in the yield extraction fit of the significant B_s^0 pollution that remains (Section 3.5) rather than further suppression by mass dependent cuts.
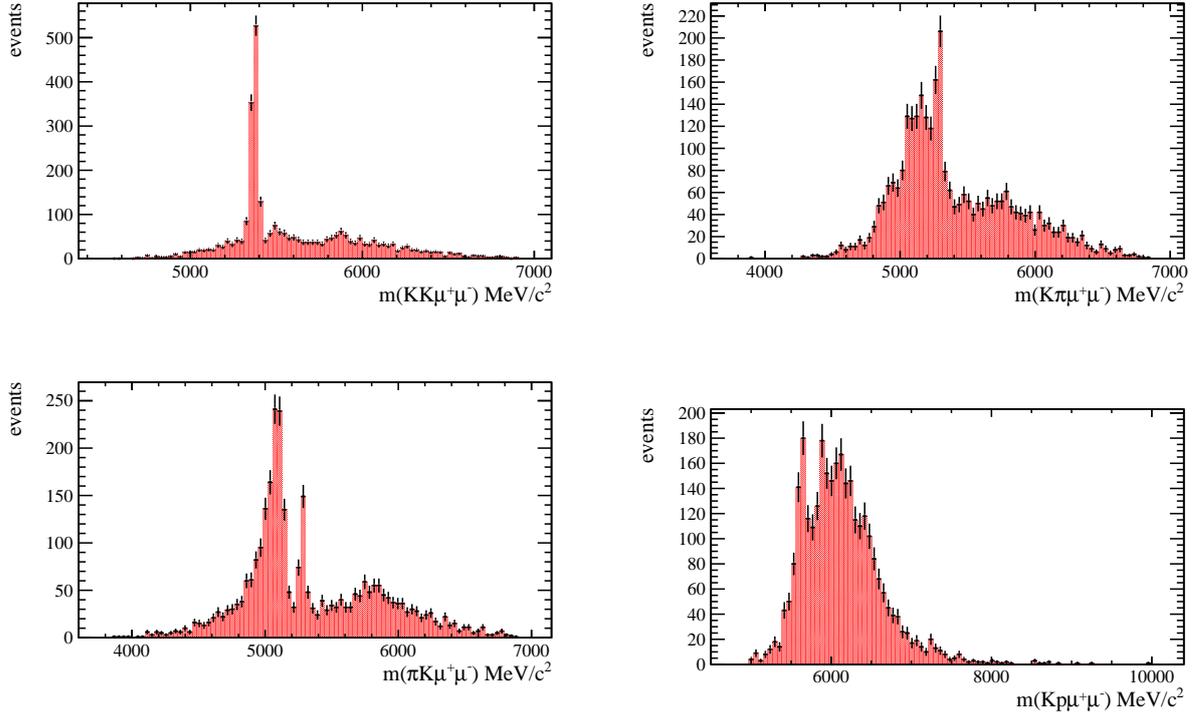
Figure 3.13: The invariant mass of the mother particle under different hypotheses for the final state hadrons, imposing $m(pK^-\mu^+\mu^-) > 6\,\text{GeV}/c^2$.

$$\pi \to \mu$$
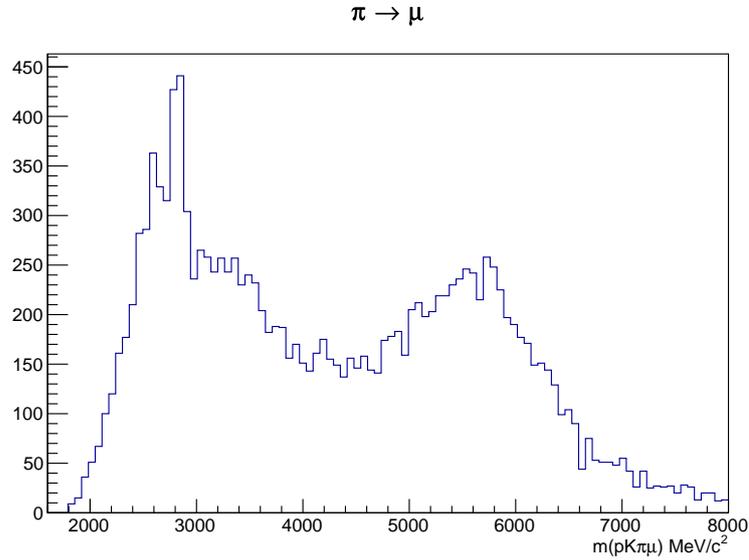
Figure 3.14: Combined lower and upper mass side-band of blinded $\Lambda_b^0 \to pK\mu^+\mu^-$ data under the $pK\pi\mu$ hypothesis, showing partially reconstructed $\Lambda_b^0 \to \Lambda_c^+\mu\nu$ decays
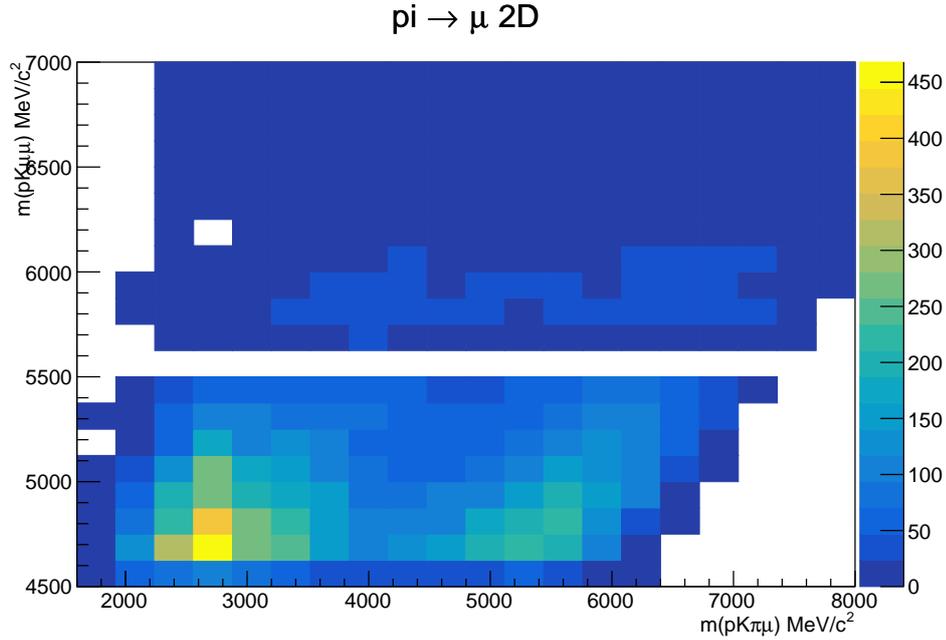
Figure 3.15: Two dimensional mass plot showing the location of the Λ_c^+ peak in $M(pK\mu\mu)$

For $\Lambda_b^0 \rightarrow J/\psi pK$, the pollution from B^0 and $\Lambda_b^0 \rightarrow J/\psi pK$ with double reflection is negligible due to the powerful suppression from the PID selection on mis-reconstructed B^0 events (1.7 % efficiency) and a combination of relatively good PID suppression (14% efficiency) and shifting of the peak far below the Λ_b^0 peak for the $\Lambda_b^0 \rightarrow J/\psi pK$ double reflection. The high mass tail of the double reflection remains under the correctly reconstructed $\Lambda_b^0 \rightarrow J/\psi pK$ peak and this is included in the fits.

3.5   Yield Extraction

Yields are determined using unbinned extended maximum likelihood fits, which have the form

$$\mathcal{L} = e^{-(N_S+N_B+N_{pbkg})} \times \frac{1}{N} \prod_{i=1}^{N} [N_S P_S(m_i) + N_B P_B(m_i) + N_{pbkg} P_{pbkg}(m_i)], \qquad (3.5)$$

where N_S represents the number of signal candidates, while N_B and N_{pbkg} correspond to the numbers of candidates from combinatorial and peaking backgrounds, respectively. Each $P(m_i)$ is the corresponding probability density function. The $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ mass peaks are described by the sum of two Crystal Ball (DCB) functions that share common means (m) and tail parameters (σ and n) but have independent widths.

A single Crystal Ball function is described in [84], consists of a Gaussian peak and power-law tail of slope n. The threshold between the Gaussian and power law tail is determined by a parameter α.

The RooFit package [85] is used to fit the Λ_b^0 mass spectrum in the rare and normalisation selections. In the nominal fit, we constrain the shapes of the B_s^0 and Λ_b^0 double reflection from their shapes in mis-reconstructed MC. The full selection is applied to all samples from which the shapes are obtained. However due to the poor modelling of the PID variables, when fits are performed on MC to extract shape parameters, we perform fits without the PID selection while evaluating fit model systematics (Section 3.9.4). For extracting the $\Lambda_b^0 \to J/\psi pK$ yield, fits are performed separately on the 2011 and 2012 data sets.

In the $\Lambda_b^0 \to J/\psi pK$ case, the combinatorial background contribution is modelled using a Chebyshev polynomial of order three, while for $\Lambda_b^0 \to pK\mu^+\mu^-$ an exponential function is used.

3.5.1   Modelling peaking backgrounds

B_s^0 appears to be the dominant peaking background due to the relatively large $p \to K$ reflection probability. Taking the side-band defined as $m(pK^-\mu^+\mu^-) > 5.8\,\mathrm{GeV}/c^2$ but in the invariant mass $m(K^+K^-\mu^+\mu^-)$ we expect mis-reconstructed $B_s^0 \to J/\psi \phi$ decays from $p \to K$ reflections to peak at the nominal B_s^0 mass. We can then extract the yield of B_s^0 in the side-band with DCB fits to the now 'correctly' reconstructed B_s^0 peak and a Chebyshev

polynomial fit for the combinatorial background. This is shown in Figure 3.16. We fix all the parameters except the yield of the now correctly reconstructed B_s^0 candidates in the sideband fit to those obtained from a fit to $B_s^0 \to J/\psi\phi$ MC (see Figure 3.19)

The B_s^0 mass distribution when reconstructed under the Λ_b^0 hypothesis is also modelled in simulation with a separate DCB (see Figure 3.20). This fixes the shape under the $\Lambda_b^0 \to J/\psi pK$ fit and also allows the yield of B_s^0 candidates in the $\Lambda_b^0 \to J/\psi pK$ fit mass window to be calculated. More explicitly, we can express the integral of the mis-reconstructed B_s^0 mass distribution in the Λ_b^0 hypothesis as

$$I \equiv \int_{m_a}^{m_c} f(m_{\Lambda_b^0})_{B_s^0}\, \mathrm{d}m_{\Lambda_b^0} \equiv \int_{m_a}^{m_b} f(m_{\Lambda_b^0})_{B_s^0}\, \mathrm{d}m_{\Lambda_b^0} + \int_{m_b}^{m_c} f(m_{\Lambda_b^0})_{B_s^0}\, \mathrm{d}m_{\Lambda_b^0}, \qquad (3.6)$$

where m_a and m_c define the full fit range for the extraction of the $\Lambda_b^0 \to J/\psi pK$ yield and m_b is the side-band cut. Therefore, after extracting the side-band yield N_{sb} from the fit we can estimate a total B_s^0 yield, $N_{B_s^0}$ from,

$$N_{B_s^0} = N_{sb} \times \frac{\int_{m_a}^{m_c} f(m_{\Lambda_b^0})_{B_s^0}\, \mathrm{d}m_{\Lambda_b^0}}{\int_{m_b}^{m_c} f(m_{\Lambda_b^0})_{B_s^0}\, \mathrm{d}m_{\Lambda_b^0}}, \qquad (3.7)$$

where in this case $m_a = 5.4\,\mathrm{GeV}/c^2$, $m_b = 5.8\,\mathrm{GeV}/c^2$ and $m_c = 6.2\,\mathrm{GeV}/c^2$. This allows us to fix the yield of B_s^0 candidates using the mis-reconstruced B_s^0 shape from MC (Figure 3.20) and the yield of B_s^0 extracted from the high mass side-band in data (Figure 3.16) For this reason, we extend the mass window of the nominal $\Lambda_b^0 \to J/\psi pK$ fit into the high mass side-band. The full fits to the $\Lambda_b^0 \to J/\psi pK$ mass can be seen in Figure 3.17 (2011 data) and Figure 3.18 (2012 data). The fit parameters are shown in Tables 3.8 and Tables 3.9, while yield and description summaries are in Tables 3.10 and 3.11.

The $\Lambda_b^0 \to J/\psi pK$ decays under double reflection is also included in the fit. The peak
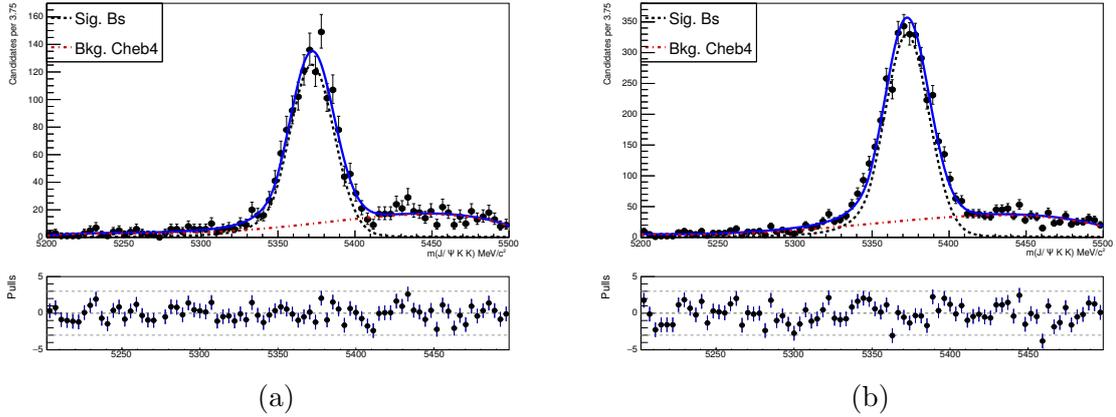
Figure 3.16: B_s^0 fit in $\Lambda_b^0 \to J/\psi\, pK$ side-band $[5.7{:}6.0]\,\mathrm{GeV}$ to fix yield, for 2011 (a) and 2012 (b)

is shifted into the lower mass side-band with the high mass tail contributing to the fit region. We fix the shape (exponential) from a fit to MC candidates and scale the yield by the ratio of integrals, in the same manner as we do for B_s^0 (eqn. 3.7), but constrain the yield in the fit region, N_{pK_swap}, to be,

$$N_{pK\_\mathrm{swap}} = N_{sig}\,\epsilon_{\mathrm{PID}}^{pK\_\mathrm{swap}}\,f_{fit} \tag{3.8}$$

where f_{fit} is the fraction of the shape expected to be in the fit region, N_{sig} is the signal yield of $\Lambda_b^0 \to J/\psi\, pK$ and $\epsilon_{\mathrm{PID}}^{pK\_\mathrm{swap}}$ is the double reflection efficiency (14%) which we evaluate using `PIDCalib`. This leaves only 399 expected candidates under the full fit region (2011 and 2012 combined).

We expect a negligible contribution from B^0 modes as no significant contribution is seen in the high mass side-band under the single and double mis-identifications required for B^0 to pollute the signal, which has only a 1.4% efficiency, and so this background shape is not included in the nominal fit. For systematic studies we add a background shape with floating yield and all other parameters constrained to a Crystal Ball fit on mis-reconstructed B^0 MC.
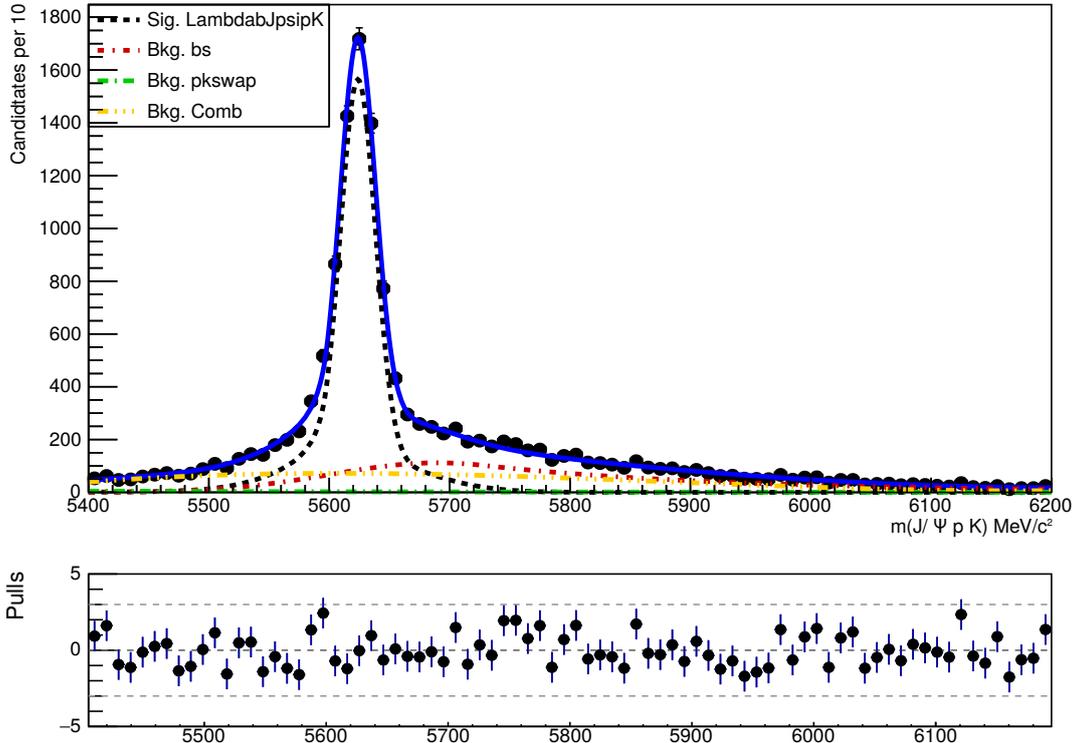
Figure 3.17: Fit to $\Lambda_b^0 \to J/\psi pK$ 2011 data, together with the contributions in the fit for B_s^0 (red dotted line), B^0 (green dotted line). The corresponding yields are presented in Table 3.8.

For the nominal fit, we constrain only the tail/slope parameters of the $\Lambda_b^0 \to J/\psi pK$ shape to that obtained from MC as there is a noticeable discrepancy in width and mean mass. In the nominal fit, we leave mean mass and Gaussian widths of the signal DCB floating. All fits to MC for $\Lambda_b^0 \to J/\psi pK$ are shown in Figures 3.19-3.20

Using the fit models in Figures 3.17 and 3.18, two further fits to the data are performed with all parameters (except yields) fixed from these nominal fits, enabling us to generate signal weights (s-weights) with the sPlot technique. Using the models described in Section 3.8, the $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ yields can be corrected event by event
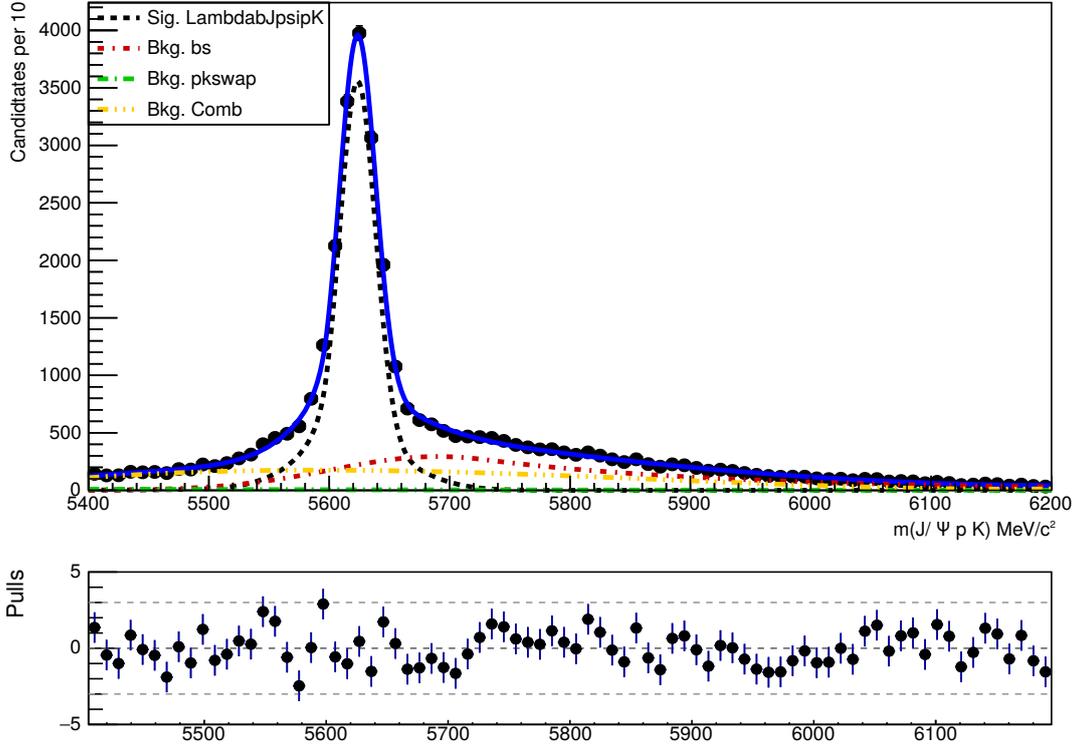
Figure 3.18: Fit to $\Lambda_b^0 \to J/\psi\, pK$ 2012 data, together with the contributions in the fit for B_s^0 (red dotted line), B^0 (green dotted line). The corresponding yields are presented in Table 3.9.

for acceptance effects, detector and selection efficiency. This gives a corrected yield of

$$N = \sum_{i=0}^{n} \frac{sw_i}{\epsilon_i}, \tag{3.9}$$

where sw_i and ϵ_i are the s-weight and absolute efficiency, evaluated for the ith event.
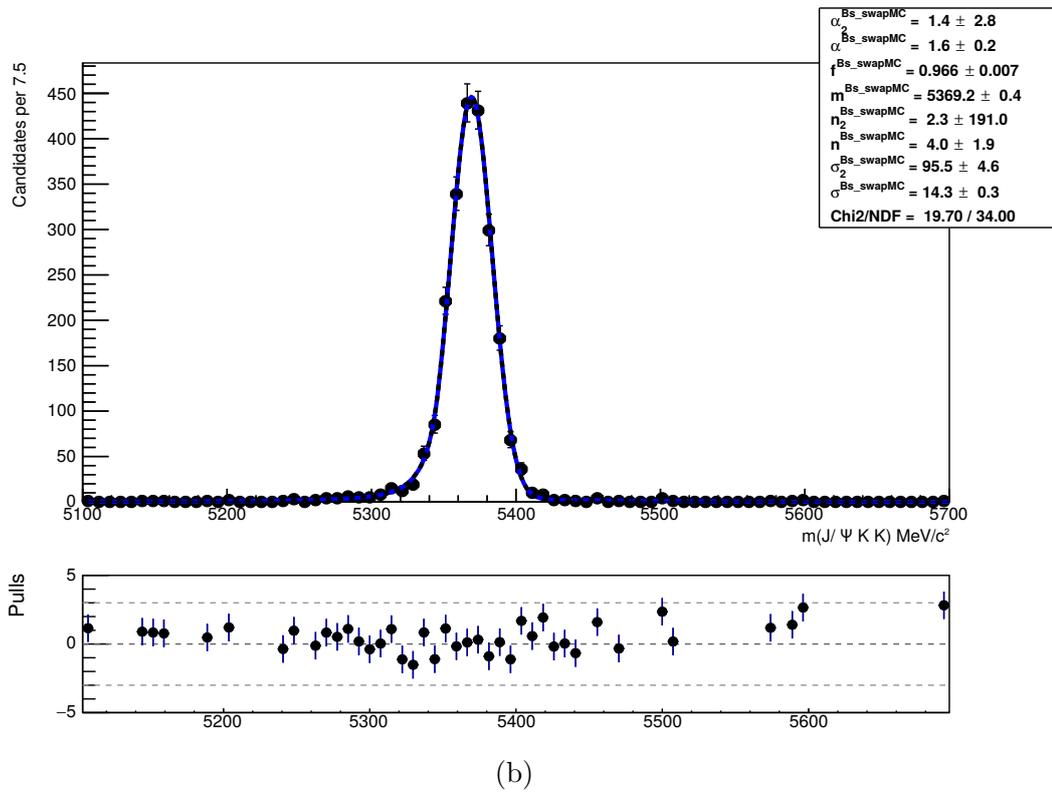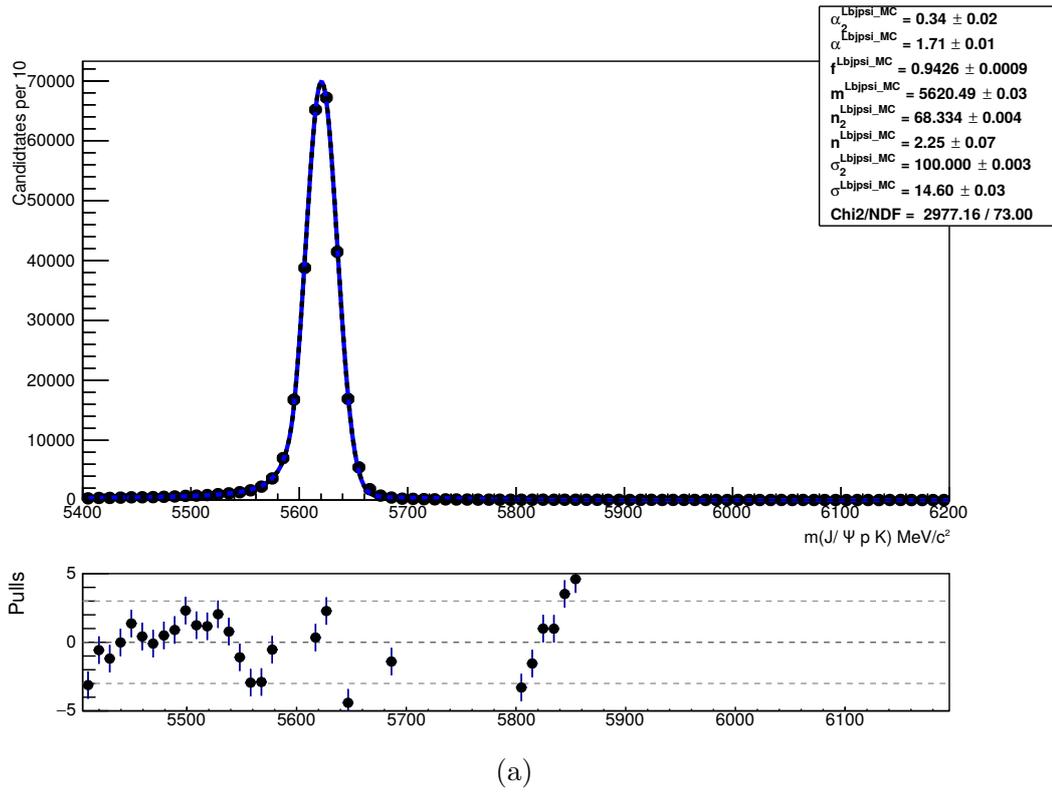
Figure 3.19: (a) Fit to $\Lambda_b^0 \to J/\psi pK$ MC sample. (b) Fit to $B_s^0 \to J/\psi\phi$ MC under correct mass hypotheses.

parameter	value
Combinatorial background	
c_0^{comb}	-0.739 ± 0.035
c_1^{comb}	-0.62 ± 0.04
c_2^{comb}	0.40 ± 0.04
N_{comb}	3526 ± 127
B_s^0 background	
$\sigma^{B_s^0}$	$89.1(\pm 3.1)$
$\alpha^{B_s^0}$	$-0.4430(\pm 0.0000)$
$m^{B_s^0}$	$5690.0(\pm 1.1)$
$n^{B_s^0}$	$2.8458(\pm 0.8256)$
$N^{B_s^0}$	3226.8013 (fixed)
$\Lambda_b^0 \to J/\psi\, pK$ double reflection	
b^{swap}	$0.0035(\pm 0.0002)$
N^{swap}	123.9839 (fixed)
$\Lambda_b^0 \to J/\psi\, pK$ signal	
$\alpha^{\Lambda_b^0}$	$1.76(\pm 0.25)$
$\alpha_2^{\Lambda_b^0}$	$2.23(\pm 0.17)$
$m^{\Lambda_b^0}$	5624.0 ± 0.3
$n^{\Lambda_b^0}$	$2.50(\pm 1.63)$
$n_2^{\Lambda_b^0}$	$0.034(\pm 0.844)$
$N^{\Lambda_b^0}$	7426 ± 129
$\chi^2/NDOF$	1.2

Table 3.8: Fit parameters for $\Lambda_b^0 \to J/\psi\, pK$ components on 2011 data. Brackets on uncertainties denote parameter that is fixed to MC fit with uncertainty from MC fit.

parameter	value
Combinatorial background	
c_0^{comb}	-0.799 ± 0.022
c_1^{comb}	-0.6150 ± 0.0001
c_2^{comb}	0.47 ± 0.03
N_{comb}	9129.0 ± 183
B_s^0 **background**	
$\sigma^{B_s^0}$	89.1 ± 3.0
$\alpha^{B_s^0}$	$-0.4430(\pm 0.0000)$
$m^{B_s^0}$	$5690.0(\pm 1.1)$
$n^{B_s^0}$	$2.85(\pm 0.83)$
$N^{B_s^0}$	8489.5185 (fixed)
$\Lambda_b^0 \to J/\psi pK$ **double reflection**	
b^{swap}	$0.0035(\pm 0.0002)$
N^{swap}	276.5531 (fixed)
$\Lambda_b^0 \to J/\psi pK$ **signal**	
$\alpha^{\Lambda_b^0}$	$1.76(\pm 0.24)$
$\alpha_2^{\Lambda_b^0}$	$2.22(\pm 0.17)$
$m^{\Lambda_b^0}$	5623.8 ± 0.2
$n^{\Lambda_b^0}$	$2.50(\pm 1.63)$
$n_2^{\Lambda_b^0}$	$0.03(\pm 0.84)$
$N^{\Lambda_b^0}$	16299.0 ± 183
$\chi^2/NDOF$	1.9

Table 3.9: Fit parameters for $\Lambda_b^0 \to J/\psi pK$ components on 2012 data. Brackets on uncertainties denote parameter that is fixed to MC fit with uncertainty from MC fit.

component	shape	yield 2011	yield 2012	total yield
$\Lambda_b^0 \to J/\psi pK$	DCB	7426	16299	23725
$B_s^0 \to J/\psi \phi$	CB	3227	8489	11716
$\Lambda_b^0 \to J/\psi pK$ (refl.)	Exponential	122	277	399
combinatorial	Chebyshev $\mathcal{O}(3)$	3526	9129	12655

Table 3.10: Fit model and yields obtained in data for $\Lambda_b^0 \to J/\psi pK$ candidates for the full fit range
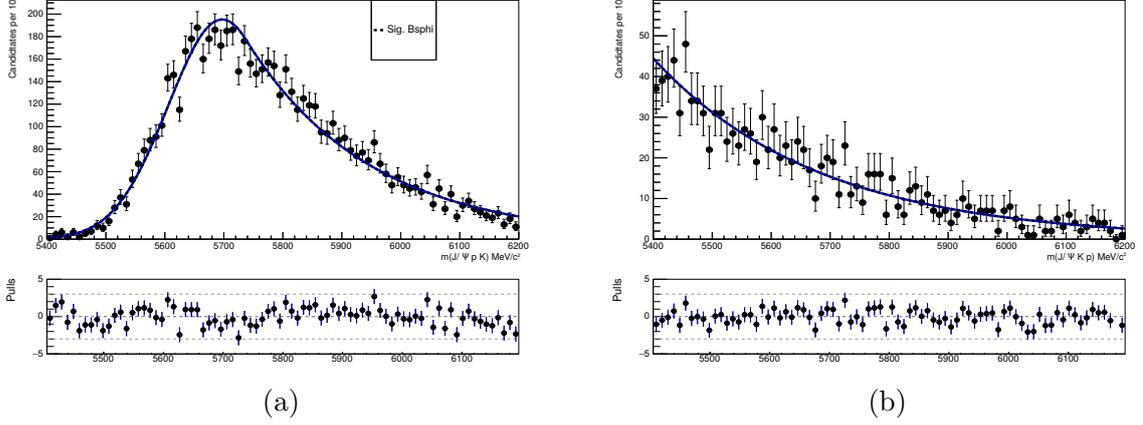
Figure 3.20: (a) Fit to $B_s^0 \to J/\psi\phi$ MC misreconstructed as $\Lambda_b^0 \to J/\psi pK$. (b)Fit to $\Lambda_b^0 \to J/\psi pK$ under double hadron reflection

component	Summary
$\Lambda_b^0 \to J/\psi pK$	Mass, widths, yield free floating. Other parameters fixed to MC
$B_s^0 \to J/\psi\phi$	All shape parameters fixed to MC. Yield fixed from side-band B_s^0 fit
$\Lambda_b^0 \to J/\psi pK$ (refl.)	Shape fixed to MC. Yield fixed using PID mis-id efficiency
combinatorial	All parameters floating

Table 3.11: Summarised description of fits in the nominal $\Lambda_b^0 \to J/\psi pK$ fits.

3.6   Efficiency

At generator level, the phase-space MC is flat in all angular variables and the mass distributions of q^2 and $m(pK)$ are described by just the kinematics of a four body decay. The efficiency is not flat in angle and mass distributions and there is also significant correlation between them. Therefore we attempt to describe the efficiency in a way that minimises dependence on the decay structure of the MC sample used to estimate the efficiency.
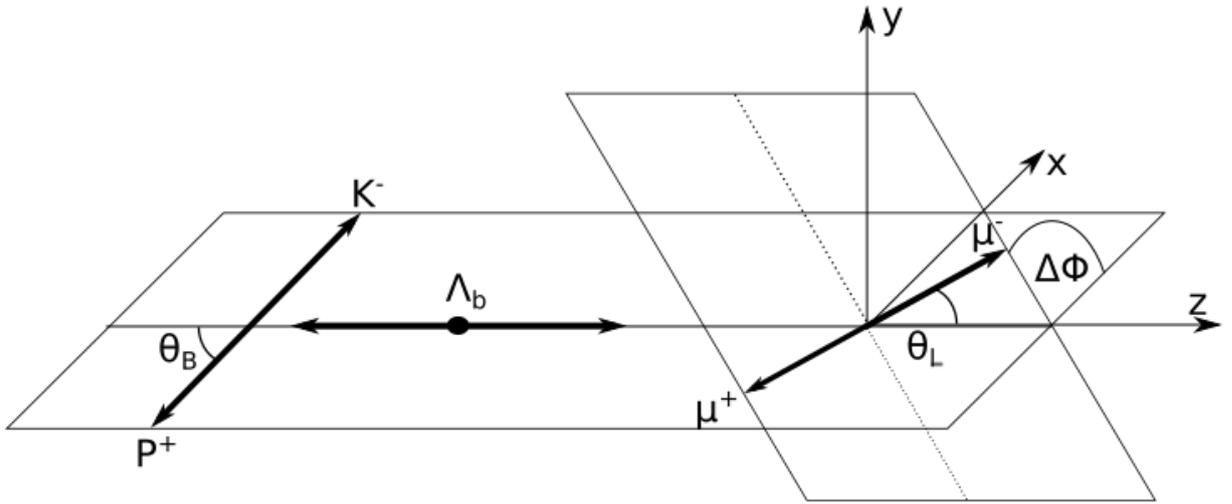
Figure 3.21: Characteristic decay angles of $\Lambda_b^0 \to pK\mu^+\mu^-$.

The decay structure of $\Lambda_b^0 \to pK\mu^+\mu^-$ can be described using three angles[3], q^2 and $m(pK)$. The angles, shown in Figure 3.21 are the characteristic decay angles of the daughter hadrons and dimuons. θ_L is the angle between the negative muon and the Λ_b^0 in the dimuon rest frame, θ_B is the angle between the proton and the Λ_b^0 in the rest frame of the pK system, and $\Delta\phi$ is the angle between these two decay planes.

---

[3]This is assuming negligible production polarisation of the Λ_b^0. Measurements of the production polarisation at LHCb show the polarisation to be consistent with zero [86]

The chosen method for efficiency modelling that has been implemented still requires us to know the integrated detection efficiencies evaluated on the phase-space MC samples. To gain a more complete understanding of the behaviour of the efficiencies, we examine the efficiency for each component of Equation 3.10 in bins of q^2, and present the results in Section 3.6.1. For the actual efficiency correction, we use the efficiency model described in Section 3.8, normalised by the phase-space efficiency (see Section 3.8.4). Other parametrisation techniques that were considered are discussed in sections 3.7.1 and 3.7.2. All three modelling techniques aim to parameterise the distortions of the five variables due to detector acceptance and selection. Despite the $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ decays in MC being governed only by phase-space kinematics, they populate the full physically allowed phase-space. Therefore, if we are able to evaluate the efficiency bias at any given point in the 5D angle-mass space, the observed $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ candidates in data can be corrected for this bias.

3.6.1   Efficiency components

The integrated efficiency can be written as

$$\epsilon = \epsilon(geo) \cdot \epsilon(sel + reco|geom) \cdot \epsilon(trig|sel + reco) \cdot \epsilon(PID|trig)\epsilon(NN|PID), \quad (3.10)$$

where $\epsilon(geo)$ is the efficiency to have the final state contained within the geometric acceptance of the LHCb detector. The selection term refers to the reconstruction, stripping and selection efficiency and is calculated for events within the geometric acceptance of LHCb. $\epsilon(trig|sel)$ is the trigger efficiency, evaluated for the events that satisfy the offline selection process. $\epsilon(PID|trig)$ refers to the efficiency to identify correctly the hadrons in the final state, and is calculated with respect to the events passing the trigger. The final term is the efficiency of the multivariate selection using a neural network.

Geometry

The geometric acceptance is evaluated using generator level Monte Carlo samples with events produced in the forward hemisphere only. All daughter particles are then required to be within the LHCb angular acceptance of between 10 and 400 mrad. Before weighting the events the acceptance is given by $\epsilon(geo) = N_{pass}/N_{gen}$. For this evaluation we used 10^6 generated events.

Reconstruction and Stripping

The reconstruction and stripping selection includes the efficiency to reconstruct correctly signal events. The efficiency of the stripping selection shown in Table 3.3, while that of the pre-selection criteria are given in Table 3.4

Trigger

Trigger lines are simulated in the Monte Carlo and the efficiency evaluated by applying the TOS requirement and comparing the number of reconstructed candidates after stripping with those that also pass the TOS selection.

Particle identification

The PID efficiency depends on the kinematics of the particle traversing the particle identification systems. The PIDCalib package is used to evaluate PID efficiency using calibration samples from data. A binning is chosen in momentum, P and pseudorapidity, η to ensure both sufficiently many events in the calibration samples and also a relatively even population distribution across the bins in the $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ MC test samples. Combined efficiencies for the correct identification of both hadrons are evaluated with PIDCalib by comparing each event's daughter hadrons to the efficiency in the bin corresponding to the hadron's momentum and η in MC. These results are then combined

to produce event weights corresponding to the multi-track PID efficiency, and we perform an event by event weighting of the MC from the PID efficiency bins in P and η. The
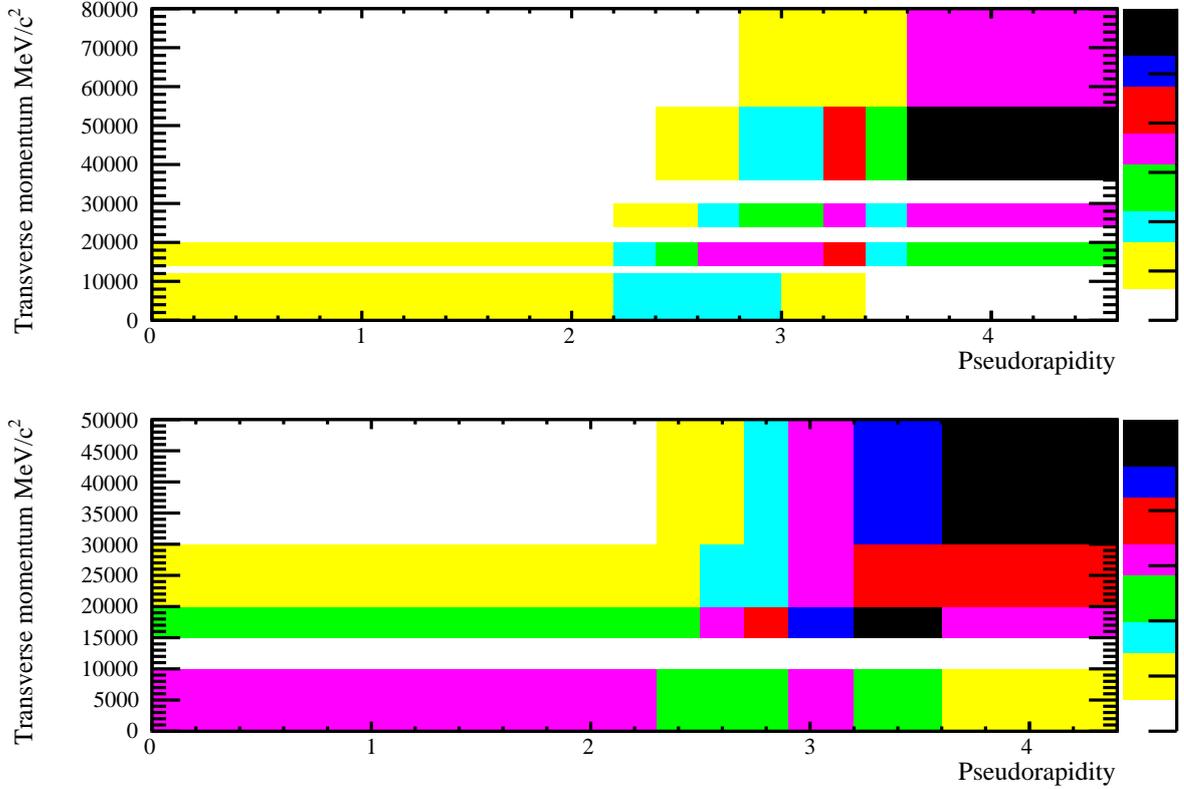
Figure 3.22: Phase-space MC kinematics for the proton (top) and kaon (bottom) in $\Lambda_b^0 \to pK\mu^+\mu^-$.

PID efficiency is evaluated separately for both of the magnet polarities and also for the 2011 and 2012 data, to account for potential variation in the PID performance with time. Figure 3.23 shows how the efficiency, integrated over q^2, changes with each running period. The PID efficiency for all four sets of magnet polarity and run period are presented in Tables 3.12–3.15, while the efficiencies for $\Lambda_b^0 \to J/\psi\, pK$ are presented in Table 3.16. The uncertainty quoted is statistical only and determined by the size of the calibration and MC samples in each of the p_{T} and η bins. This does not include the (typically larger) effect of binning, which are taken into account in the evaluation of the PID systematic uncertainty, as summarised in Section 3.9.2.

q^2 (GeV/c^2)	PID Efficiency (%)	Uncertainty
0.1-2.0	77.934	0.045
2.0-4.0	77.488	0.048
4.0-6.0	76.557	0.055
6.0-8.0	75.334	0.069
11.0-12.5	71.510	0.017
15.0-17.5	68.39	0.058
Integrated	76.397	0.024

Table 3.12: PID efficiency for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagUp' polarity, 2011

q^2 (GeV/c^2)	PID Efficiency (%)	Uncertainty
0.1-2.0	76.871	0.027
2.0-4.0	79.368	0.029
4.0-6.0	75.381	0.033
6.0-8.0	74.230	0.041
11.0-12.5	69.670	0.100
15.0-17.5	66.880	0.400
Integrated	75.258	0.015

Table 3.13: PID efficiency for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagDown' polarity, 2011

q^2 (GeV/c^2)	PID Efficiency (%)	Uncertainty
0.1-2.0	79.650	0.021
2.0-4.0	79.325	0.022
4.0-6.0	78.449	0.025
6.0-8.0	77.019	0.031
11.0-12.5	74.121	0.078
15.0-17.5	70.970	0.290
Integrated	78.264	0.011

Table 3.14: PID efficiency for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagUp' polarity, 2012

q^2 (GeV/c^2)	PID Efficiency (%)	Uncertainty
0.1-2.0	80.937	0.021
2.0-4.0	80.722	0.022
4.0-6.0	79.995	0.025
6.0-8.0	78.882	0.031
11.0-12.5	75.104	0.079
15.0-17.5	72.58	0.330
Integrated	79.749	0.011

Table 3.15: PID efficiency for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagDown' polarity, 2012

Run (GeV/c^2)	PID Efficiency (%)	Uncertainty
2011 MagUp	71.944	0.011
2011 MagDown	71.620	0.011
2012 MagUp	74.417	0.011
2012 MagDown	75.605	0.011

Table 3.16: PID efficiencies for $\Lambda_b^0 \to J/\psi\, pK$
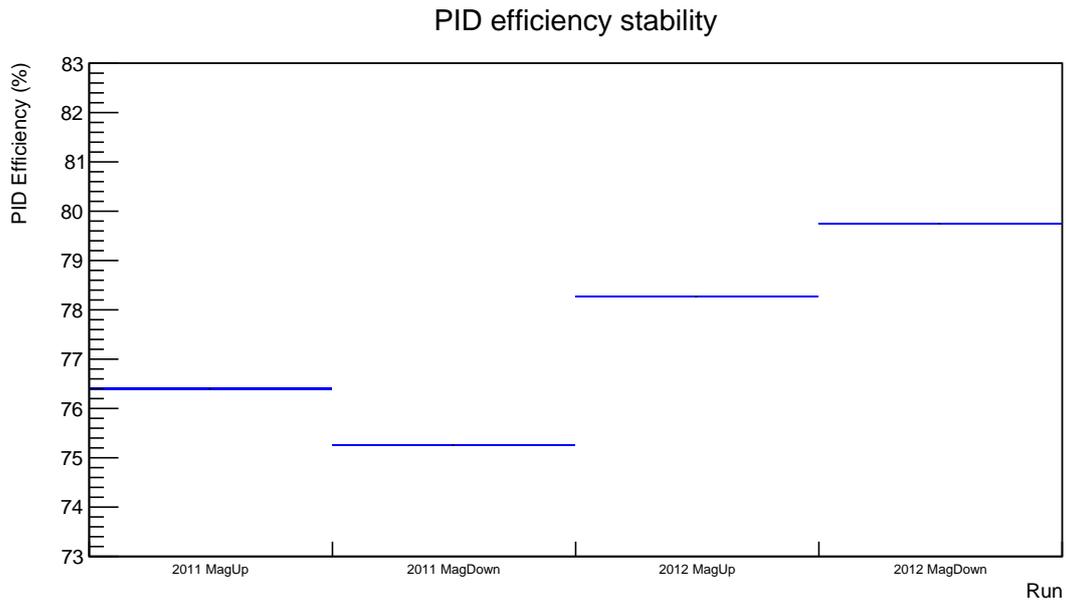
Figure 3.23: PID efficiency for each magnet polarity in 2011 and 2012, showing $\sim 5\%$ variation over all runs. The bar thickness corresponds to the uncertainty on the efficiency due to the statistics of the calibration and MC samples

3.7   Multi-dimensional density estimation

Parameterising the full detector and selection efficiency in the five dimensions of $\cos\theta_L$, $\cos\theta_B$, $\Delta\phi$, $m(pK)$ and q^2 is non-trivial. Finite MC statistics can cause issues with accuracy, especially at the phase-space boundaries.

Investigations were carried out into the viability of several different non-parametric modelling techniques, which included relative kernel density estimation, supervised training of a neural network and density estimation with orthogonal (Legendre) polynomials. Each method was found to have its merits as well as disadvantages. Ultimately the method of density estimation using Legendre polynomials is used for yield correction of $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ candidates for the branching fraction measurement, however a discussion of all three methods is presented below.

3.7.1   Four dimensional relative kernel density estimation

Kernel density estimation can be used to estimate the true distribution (PDF_{true}) that is characterised by a random variable, for instance $\mathbf{x} = (x_1, x_2, ...x_n)$. The kernel ($K(x)$) is typically a weighting function that has an integral of 1, and can be used to generate an estimation (PDF_{KDE}), of PDF_{true} as [87],

$$PDF_{KDE}(x) = \frac{1}{N} \sum_{i=1}^{N} K(x - x_i). \tag{3.11}$$

A kernel function is generated for each data point with the PDF_{KDE} being the normalised sum of these.Figure 3.24 shows a simple example of KDE using Gaussian kernels.

A free parameter in $K(x)$ is the bandwidth. A large bandwidth is ideal for data with limited statistics due to the distribution smoothing, however this can cause fine structures to be smeared away, or lead to mis-modelling of sharp boundaries as shown in Figure 3.25. On the contrary if the bandwidth is too narrow, this can cause 'over-parametrisation' of
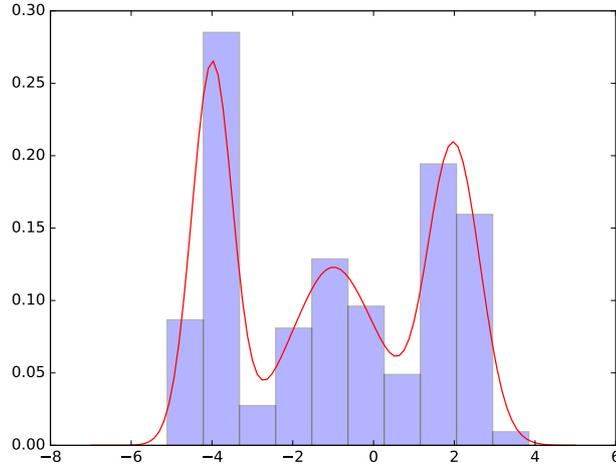
Figure 3.24: Simple one dimensional KDE example using a Gaussian kernel.
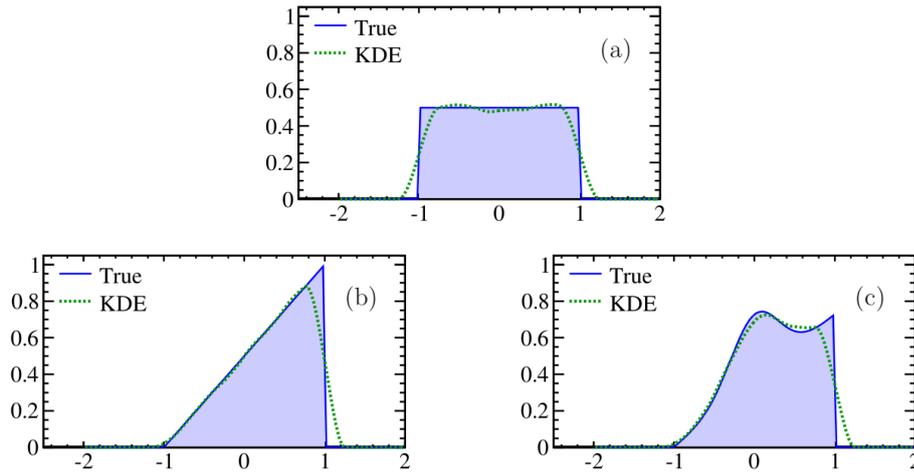
Figure 3.25: True and KDE PDF's for uniform (a), linear (b) and linear+Gaussian (c) distributions without boundary corrections. [88]

statistical fluctuations.

For the parametrisation of $\Lambda_b^0 \to pK\mu^+\mu^-$, the LHCb MeerKat [88] package is used which employs a boundary correction technique. MeerKat uses the Epanechnikov kernel

defined as,

$$K(x) = \begin{cases} \frac{3}{4\sigma}\left(1 - \frac{x^2}{\sigma^2}\right), & \text{for } x \geq (-\sigma, \sigma) \\[2ex] 0 & \text{otherwise} \end{cases} \tag{3.12}$$

The boundary correction technique is to create a corrected PDF $PDF_{corr}(x)$ which incorporates an approximation function $F(x)$ which is equal to 1 for $x \geq X$ and zero elsewhere,

$$PDF_{corr}(x) = \begin{cases} \dfrac{\frac{1}{N}\sum\limits_{i=1}^{N} K(x-x_i)}{(U\otimes K)(x)} & x \geq X \\[2ex] 0 & \text{otherwise} \end{cases} \tag{3.13}$$

This approximation function, which can be factorised 1D KDE's, analytical descriptions of phase-space coverage, a top hat function etc. allows the boundaries and fine structures to be accurately parameterised as relative fluctuations to the approximation function.

The variables $\cos\theta_L$, $\cos\theta_B$, $\Delta\phi$ and $m(pK)$ are used to construct a four dimensional relative kernel density parameterisation in each q^2 bin. The choice to bin in q^2 rather than incorporate it as a fifth dimension is due to CPU time considerations. At four dimensions the process takes $\mathcal{O}(1day)$ on a typical desktop machine, and does not increase linearly with extra dimensionality. However, each parameterisation in q^2 can be run in parallel.

The quality of the fit is evaluated by comparing toy Monte-Carlo events created from the kernel PDF to an independent sample of simulated events. This is done in 24 bins of the 4D space and the χ^2 of the difference between simulated and PDF events with respect to a fit of y=1 is used as the figure of merit. These results are shown in Figure 3.26. In these plots, the four dimensions are unfolded into one by scanning through each of the 24 bins of the model and the MC and the integrals of those bins compared, imposing that the integrals over the entire phase space of MC and model are equal.

Overall, the modelling is of exceptional quality. However, the desired technique to

incorporate systematic uncertainty estimations on MC statistics, PID weighting, kinematic weighting and $\varLambda_b^0$ lifetime uncertainty require that toy MC experiments be run in which uncertainties on each of these sources are used to randomly vary each individual event weight before being applied to MC used to generate the efficiency parameterisation, each of which is applied to data to recalculate the corrected yields. This requires the very computationally intensive model generation to be run $\mathcal{O}(1000) \times N^{sources} \times N^{q^2bins}$ times, where $N^{sources}$ is the number of sources of uncertainty being considered and N^{q^2bins} is the number of q^2 bins. This is unfeasible and other methods for estimating the systematic uncertainties would have to be investigated.

(a) q^2:[0.1,2] GeV^2/c^4

(b) q^2:[2,4] GeV^2/c^4

(c) q^2:[4,6] GeV^2/c^4

(d) q^2:[6,8] GeV^2/c^4

(e) q^2:[11,12.5] GeV^2/c^4

(f) q^2:[15,17.5] GeV^2/c^4
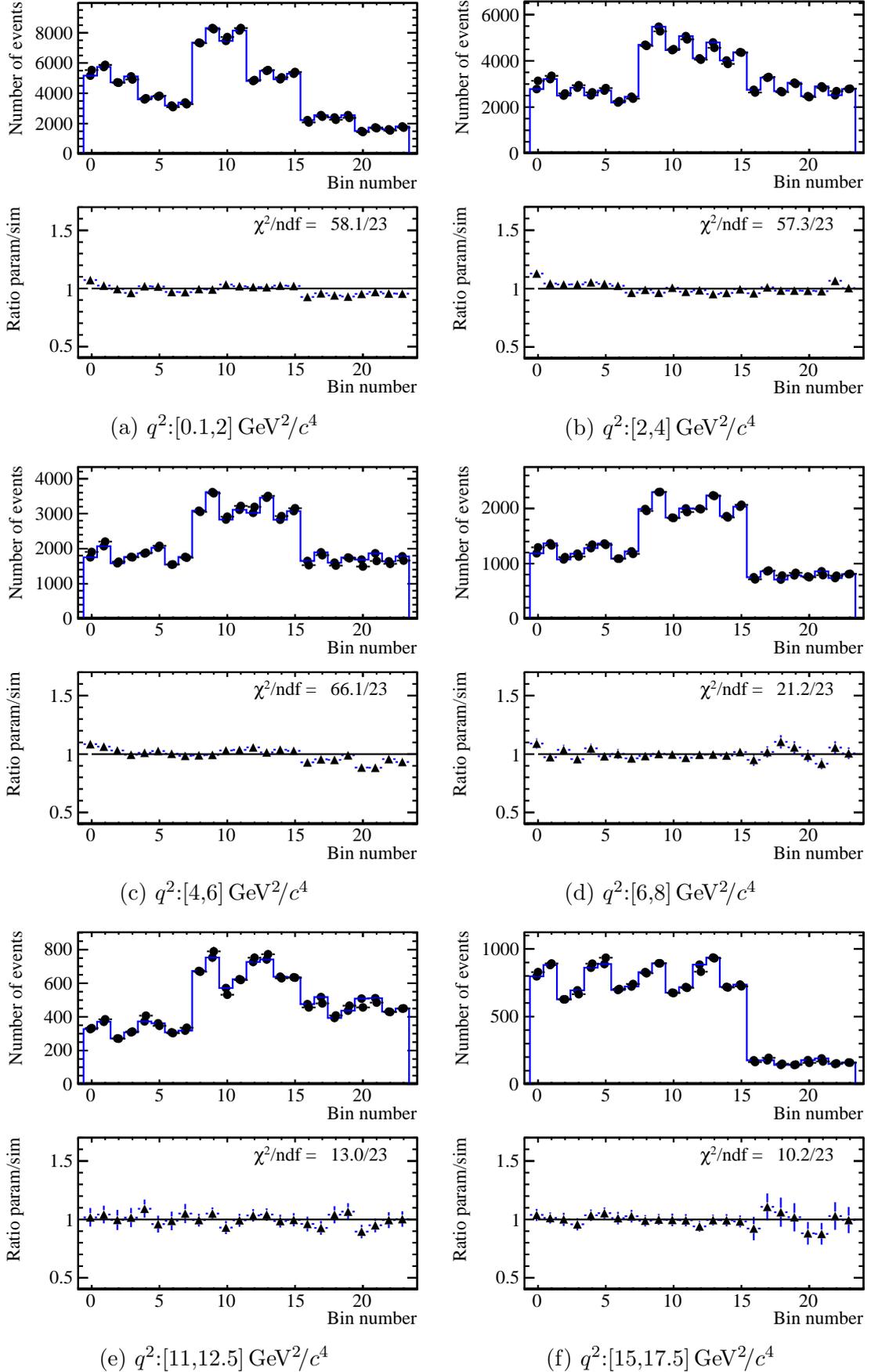
Figure 3.26: PDF quality evaluation after full selection for six q^2 bins. These share the same bin boundaries as the five standard bins defined in 3.3 but with a split bin for finer granularity.

3.7.2   Efficiency estimation with a Neural Network weighting method

The viability of the the same NEUROBAYES framework used for the selection of $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ candidates (section 3.3.2) in performing multi-dimensional density estimation for event-by-event efficiency correction is investigated. Instead of using the neural network as a classifier to discriminate between signal and background candidates, the network is trained on the five variables of $\cos\theta_L$, $\cos\theta_B$, $\Delta\phi$, $m(pK)$ and q^2 for two MC samples (A and B) of $\Lambda_b^0 \to pK\mu^+\mu^-$ candidates. Sample A is a set of MC-truth generator level events in which no selection is performed, while sample B is a set of reconstructed events passed through full detector simulation and the full $\Lambda_b^0 \to pK\mu^+\mu^-$ selection. The output weight, usually used to discriminate between signal and background is rescaled from -1–+1 to 0–1, where it akin to corresponds to a Bayesian probability. Weighting the generator level sample event by event should reproduce the kinematic biasing caused by the reconstruction and selection, including their correlations. It can therefore be treated as an efficiency correction weight. The integrated efficiency is still required in order to correctly normalise. This is because the weight may reproduce the relative changes in the kinematic phase-space, but has an arbitrary mean which depends largely on the relative sizes of the two training samples.

The normalisation is found by scaling the efficiency weights to give the integrated phase-space efficiency, i.e,

$$\epsilon_{phsp} \equiv \frac{N_{obs}}{N_{gen}} = I \cdot \frac{\sum_{i=1}^{n} w_i}{n} \tag{3.14}$$

where N_{gen} is the number of generated candidates, N_{obs} is the number of those which are observed after the full selection, I is the normalisation quantity to be found, w_i is the weight for the ith candidate and n is the total number of weighted candidates. The results

are shown in Table 3.17

Component	Value	Uncertainty
ϵ_{phsp}	0.0368	0.0004
Mean weight	0.496	0.001
Correction factor	0.0742	0.0007

Table 3.17: Values for the integrated phase-space MC efficiency, mean weight and corresponding correction factor.

The quality of the modelling is evaluated by applying the trained neural network to a sample of generator level MC. This has the effect of weighting each event by the efficiency, as determined by the neural network, for the point in the kinematic phase space in which that particular event lies. If the modelling is of good quality, the weighted generator level sample will appear kinematically similar to an MC sample in which the full LHCb detector, reconstruction and trigger is simulated, and full selection applied. Figure 3.27 displays the comparison of the weighted generator level events with an MC sample passed through full simulation and selection, as well as the original unweighted generator level. The results are projected into the five, one dimensional components of the kinematic phase-space and this gives a visual representation of how well the biases are modelled. Figure 3.28 instead displays the 5D phase-space unfolded into one dimension using the same technique as with the relative kernel density estimation technique. The angular variables are folded over into positive values and binned into two equal bins. The two masses are binned into three approximately equal bins, where the boundaries are chosen to ensure sufficient statistics in each.

From figure 3.27 we can see that this method attempts to approximate the acceptance effects, but the quality is far from what is achieved with the relative KDE technique. The advantage of this method is that it is ~two orders of magnitude faster than with using KDE's, but the degree of accuracy is not sufficient for efficiency correction of the $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ yields.
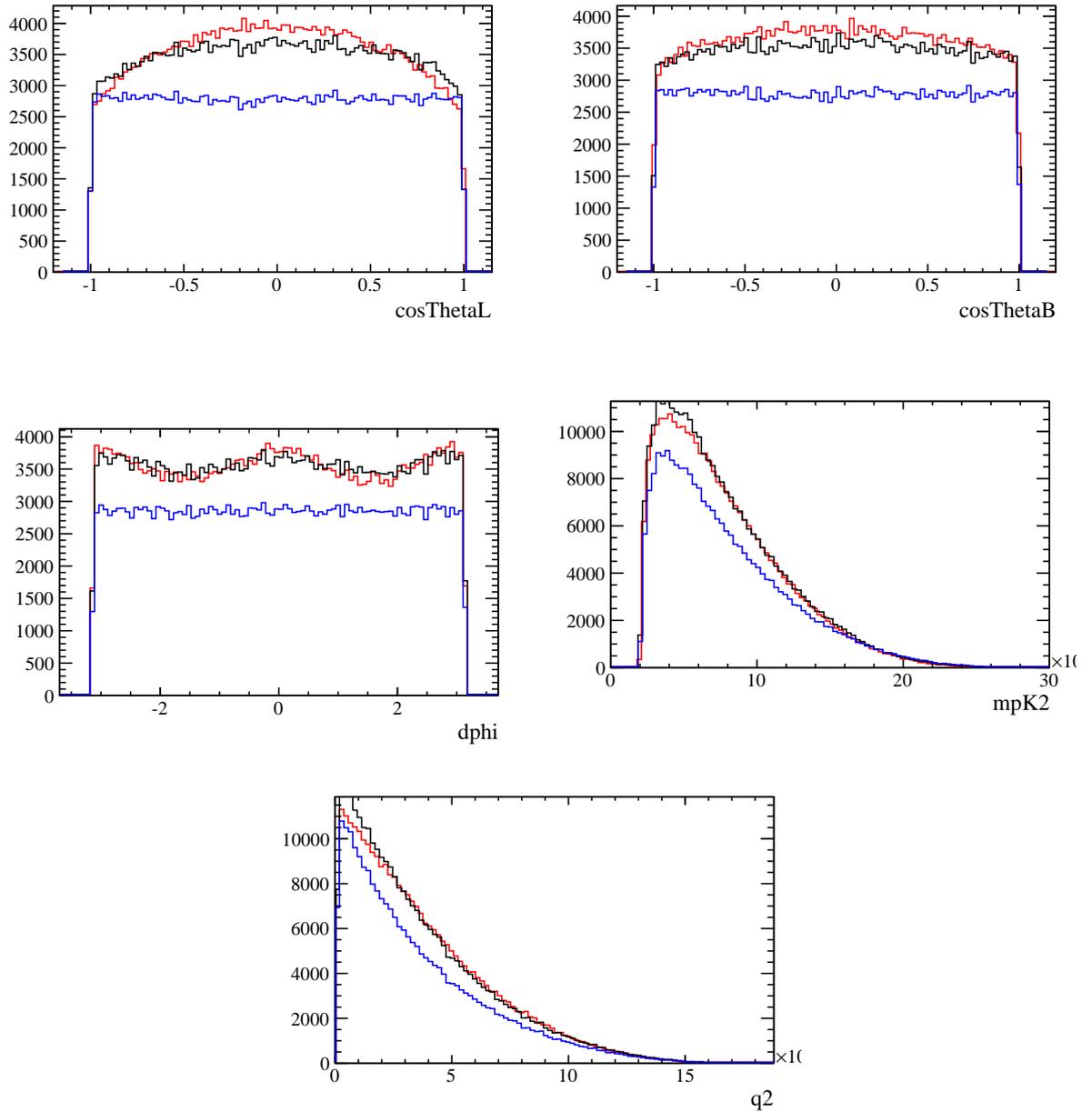
Figure 3.27: 1D projections from 5D phase-space. Blue: Unweighted generator level. Red: full selection sample. Black: weighted generator level. Scaling of 0.8 on unweighted generator level for visual clarity
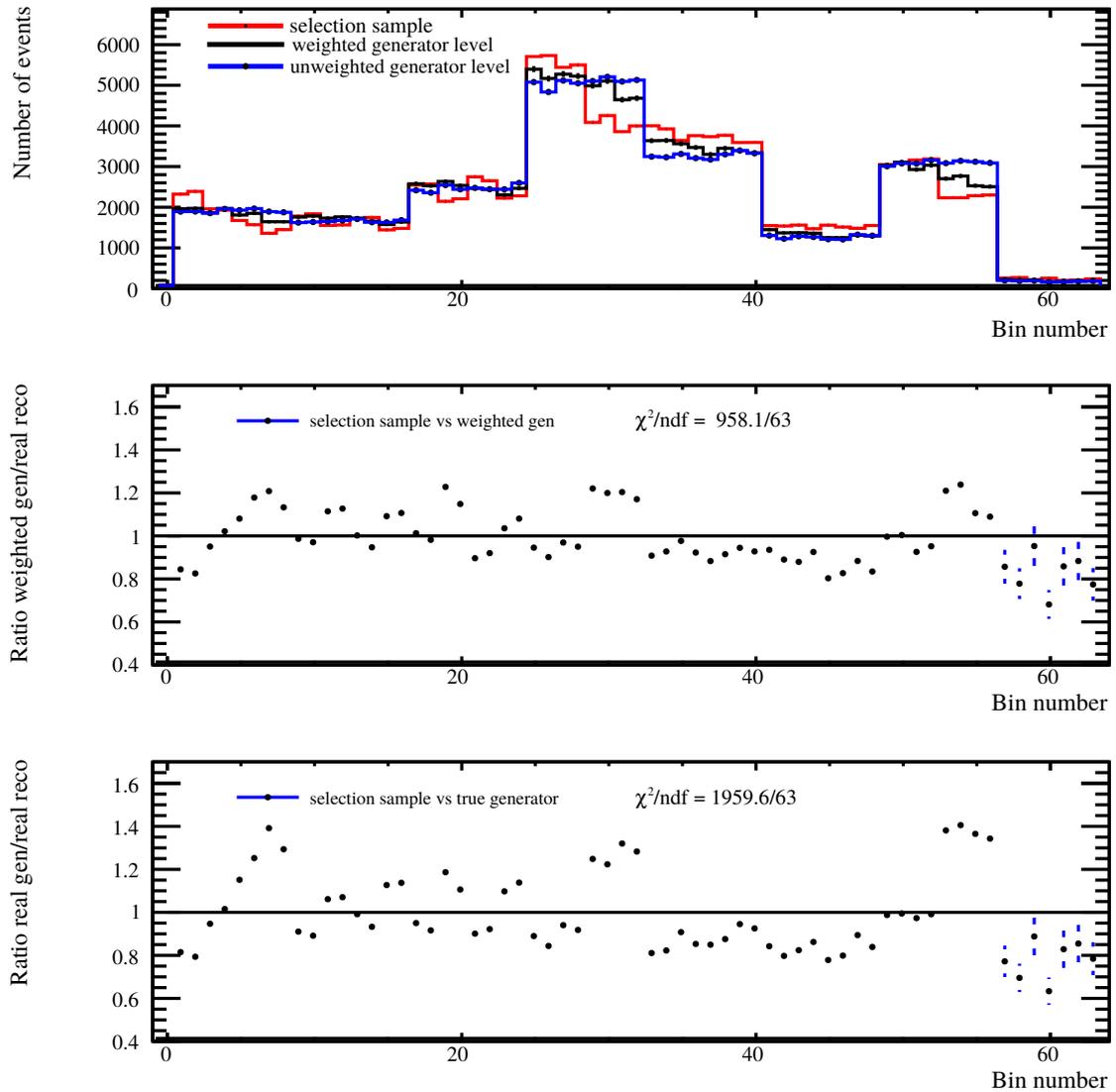
Figure 3.28: Test of efficiency modelling, showing very large discrepencies. The statistical errors are too small to see in all but the highest q^2 bins (bin 58 - 64).

3.8   Density estimation with Legendre polynomials

The following method is used for the final efficiency correction model for both $\Lambda_b^0 \to J/\psi\, pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$. The parametrisation is constructed using Legendre polynomials.

Taking the one dimensional case of wishing to parameterise the efficiency in only the variable variable $\cos\theta_L$,

$$\epsilon(\cos\theta_L) = \sum_i c_i P_i(\cos\theta_L), \tag{3.15}$$

where c_i are coefficients that need to be found and P_i are Legendre polynomials, with i denoting the order of the polynomial. Legendre polynomials are orthogonal, satisfying the condition,

$$\int_{-1}^{+1} P_n(\cos\theta_L)P_m(\cos\theta_L)d\cos\theta_L = \frac{2}{2n+1}\delta_{mn}, \tag{3.16}$$

where δ_{mn} is the Kronecker delta. To calculate the coefficients, c_i, we must first calculate the normalised moments,

$$M_i = \frac{1}{N_{events}} = \sum_j^{N_{events}} P_j \cos\theta_{L_j} \equiv \frac{1}{N}\int_{-1}^{+1} \epsilon(\cos\theta_L)P_i(\cos\theta_L)d\cos\theta_L \tag{3.17}$$

where N_{events} is the total number of events in the sample used to generate the model, and the normalisation constant $N = \int_{-1}^{+1} \epsilon(\cos\theta_L)d\cos\theta_L$. Therefore, we can see from substituting Equation 3.15 into Equation 3.17 and using the orthogonality condition, that $N = 2c_0$.

The coefficients c_i can be evaluated from the normalised moments by expanding Equation 3.15,

$$M_i = \frac{1}{2c_0}\frac{2}{2j+1}\delta_{ij}c_j = \frac{1}{c_0}\frac{c_i}{2i+1}, \tag{3.18}$$

which gives,

$$c_i = c_0 M_i(2i+1), \tag{3.19}$$

with c_0 being a free parameter that controls the overall normalisation of the model. For simplicity we choose $c_0 = 1/2$ to give a normalisation of unity.

Elaborating this into the five dimensional case for $\Lambda_b^0 \to pK\mu^+\mu^-$, it takes the form,

$$
\epsilon\left(\cos\theta_L, \cos\theta_B, \Delta\phi, m(pK), q^2\right) = \sum_{i,j,k,l,m=0} c_{i,j,k,l,m} P_i\left(\cos(\theta_L)\right) P_j\left(\cos(\theta_B)\right)
$$
$$
\cdot P_k\left(\Delta\phi'\right) P_l\left(m(pK)'\right) P_m\left((q^2)'\right),
\tag{3.20}
$$

where the primed variables, $\Delta\phi'$, $m(pK)'$ and $(q^2)'$ have been transformed to between -1 and +1 to preserve orthogonality. The sum runs up to chosen order in each of the orders of each Legendre polynomial. Such parametrisation includes correlations amongst variables and thus fully parametrises all dependencies. The coefficients $c_{i,j,k,l,m}$ are given as

$$
c_{ijklm} = c_0 M_{ijklm}(2i+1)(2j+1)(2k+1)(2l+1)(2m+1),
\tag{3.21}
$$

with M_{ijlkm} calculated using the method of moments as

$$
M_{ijlkm} = \frac{1}{N_{evts}} \sum_{\text{events}} P_i\left(\cos(\theta_L)\right) P_j\left(\cos(\theta_B)\right) \cdot P_k\left(\Delta\phi'\right) P_l\left(m(pK)'\right) P_m\left((q^2)'\right)
\tag{3.22}
$$

3.8.1   Efficiency modelling for $\Lambda_b^0 \to J/\psi\, pK^-$ decay

For $\Lambda_b^0 \to J/\psi\, pK$, the model takes the same form but is performed in four dimensions by treating q^2 as δ-function and integrating over q^2 as the J/ψ has very narrow width and signal is distributed over narrow q^2 interval.

As this method describes only the shapes of the distributions of variables in MC, it is only valid under the assumption that they are flat at generator level. The angular variables are flat in the generator level MC but with distribution in q^2 and $m(pK)^2$ governed by the phase-space kinematics. Therefore, we model the distributions of both generator level and

selected candidates using the method of moments and correct using the true phase-space distributions in 1D (m(pK)). The model is then parametrising only the relative difference between the true and observed variables. We choose Legendre polynomial orders for each variable as shown in Table 3.18.

Variable	order
$\mathcal{O}\left(\Delta\left(\phi\right)\right)$	8
$\mathcal{O}\left(\cos\left(\theta_L\right)\right)$	2
$\mathcal{O}\left(\cos\left(\theta_B\right)\right)$	2
$\mathcal{O}\left(m\left(pK\right)^2\right)$	8

Table 3.18: Order of the Legendre polynomial for the different variables driving the $\Lambda_b^0 \to J/\psi pK$ efficiency model

3.8.2   Efficiency modelling for $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decay

For $\Lambda_b^0 \to pK\mu^+\mu^-$, the case is more complicated. The inclusion of q^2 creates a two dimensional mass space that needs to corrected by the inverse of the phase-space distribution, and which is sparsely populated towards the boundaries. A full parameterisation across the whole distribution requires going to high orders of polynomial and with so few events at the edges of the phase space, such a high order model begins to oscillate below zero at these points. Before developing the final model, several methods were investigated in an attempt to suppress negative weighting:

- Division of the phase space into two dimensional bins of $m(pK)$ and q^2.

- Generating factorised one dimensional Legendre polynomial models of the variables before performing the full 5D parameterisation. Similar to the use of approximation functions in the relative KDE model.

- Generate toy MC's where the MC is resampled and the efficiency model regenerated on each resampled set. The distributions of each moment are extracted, and moments

within $\alpha\sigma$ of zero are killed, where σ is the standard deviation of the moment's distribution and α is scaled to minimise the number of negative weights

Ultimately, all methods saw improvement and the final model uses a combination of all of the above excluding the killing of moments as these toys revealed that the majority of the negative efficiency weights do not come from small moments with high uncertainties, but from strong oscillations below zero.

For the full $\Lambda_b^0 \to pK\mu^+\mu^-$ efficiency model we firstly derive two parameterisations, one to cover a q^2 range of 0.3 GeV^2/c^4 to maximum q^2, and one for the q^2 region between 0.1 and 0.3 GeV^2/c^4. For each of the two models, the variables are parameterised assuming factorisation into lower dimensional spaces. This allows the strong distortions of the variables to be described by simpler expressions before the 5D model of Equation 3.20 incorporates correlations using fewer orders of polynomials. A analytically determined four body phase-space factor is used to weight the 2D $m(pk) - -q^2$ space to be flat and this weighted distribution is used in the calculation of the moments. To transform $m(pK^-)$ and q^2 into the range -1 and 1, $m(pK^-)$ is first transformed based on the full $m(pK)$ range. Then for each event, for q^2 we set maximum based on the value of $m(pK^-)$. This avoids empty regions in the phase-space. Details on this method are specified in Appendix .3.

Analytic functions are used to parametrise the angular variables in one dimensional projections, while Legendre polynomials up to order 10 are used to parameterise $m(pK)$ between the q^2 range of 0.05–17.6 GeV^2/c^4.

We then parameterise the 3D angular space and 2D mass space separately with Legendre polynomials before finally performing the full 5D parameterisation of Equation 3.20. Each step is done on the sample corrected from the last step, so the final efficiency model is the product of each step.
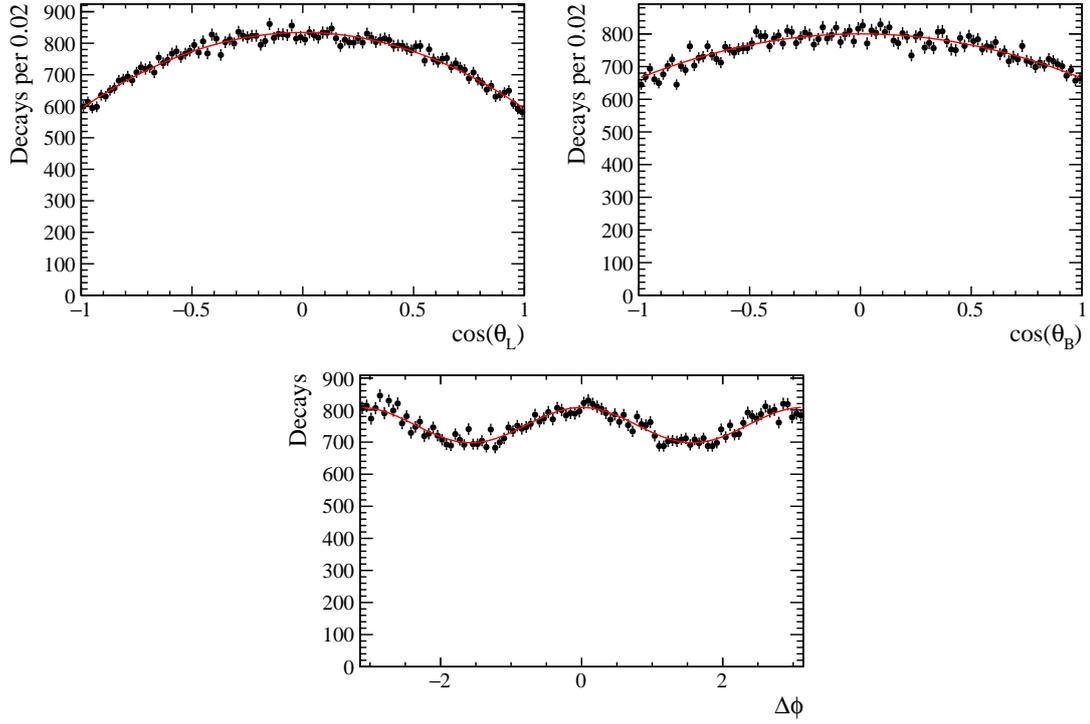
Figure 3.29: One dimensional distributions of $\Lambda_b^0 \to pK\mu^+\mu^-$ MC events with full selection in $q^2 > 0.2$ GeV$^2/c^4$ region with the 1-dimensional parametrisation of the efficiency overlayed. The lower q^2 boundary is extended down to 0.2 GeV$^2/c^4$ to avoid boundary effects by overlapping with the low q^2 model.

For the model in $q^2 > 0.3$ GeV$^2/c^4$, the three angular projections are parametrised by,

$$\epsilon(\Delta\phi) = 1 + 0.07308 \cos(2\Delta\phi), \tag{3.23}$$

$$\epsilon(\cos(\theta_L)) = 1 - 0.29(\cos\theta_L)^2, \tag{3.24}$$

$$\epsilon(\cos(\theta_B)) = 1 - 0.17(\cos\theta_B)^2. \tag{3.25}$$

These are superimposed over fully selected MC events (Figure 3.29), weighted for PID efficiencies and corrected for the Λ_b^0 production kinematics. The 1-dimensional angular
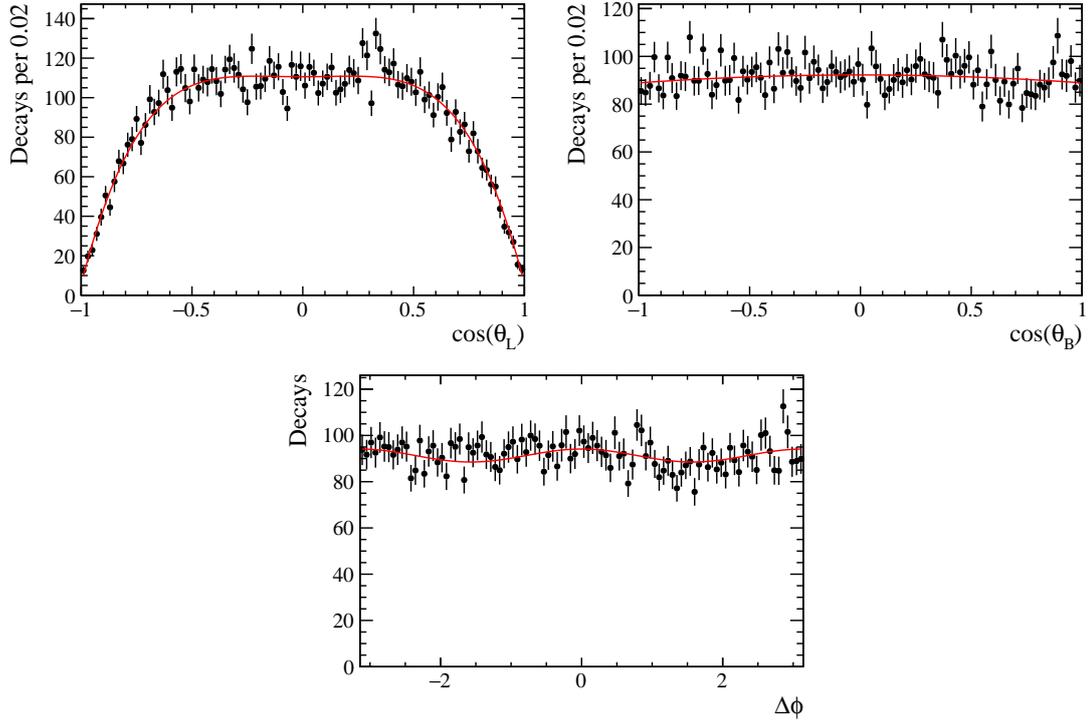
Figure 3.30: One dimensional distributions of $\Lambda_b^0 \to pK\mu^+\mu^-$ MC events with full selection in the $q^2 < 0.8$ GeV$^2/c^4$ region, with the 1-dimensional parametrisation of the efficiency overlayed.

projections for $q^2 < 0.3\,\text{GeV}^2/c^4$ are parameterised by,

$$\epsilon(\Delta\phi) = 1 + 0.0307\cos(2\Delta\phi), \tag{3.26}$$

$$\epsilon(\cos(\theta_L)) = 1 + 0.101(\cos\theta_L)^2 - 1.05(\cos\theta_L)^4, \tag{3.27}$$

$$\epsilon(\cos(\theta_B)) = 1 - 0.036(\cos\theta_B)^2. \tag{3.28}$$

Shown in Figure 3.30 overlayed on fully selected MC events with PID and kinematic weightings. The parameterisation of $\Delta\phi$ requires a high order of polynomials. Performing the one dimensional projections significantly lowers this order requiremnt in the full 5D model.

The 2D $m(pK^-)-q^2$ parameterisation uses Legendre polynomials up to order 8 and 5 in $m(pK^-)$ and q^2 respectively. The 3D angular parameterisation uses Legendre polynomials

up to second order in all three angles. The final 5D correction uses Legendre polynomials up to second order in all 5 variables.

3.8.3   Efficiency Model Validation

The efficiency model is tested in a similar way to the $\Lambda_b^0 \to J/\psi\, pK$ model. MC candidates with the full selection as well as PID and kinematic correction weights applied, is weighted by the inverse of the efficiency from the full parametrisation. If the model correctly describes the non-trivial dependencies, they will be removed, leaving flat distributions, assuming they are flat at generator level. Comparison of these distributions with a line of zero gradient are shown for $\Lambda_b^0 \to pK^-\mu^+\mu^-$ in Figure 3.31. The χ^2/ndf for the $\cos\theta_L$ is 96.378/99, 122.91/99 for $\cos\theta_B$ projection and 97.4077/99 for the $\Delta\phi$ projection. We therefore determine these to be sufficiently flat.

The $m(pK^-)$-q^2 distribution is also reasonably uniform, with significant fluctuations at some boundary areas due to the small bin size, chosen to decrease possible edge effects on the region.

Similar plots for the $\Lambda_b^0 \to J/\psi\, pK$ decay are shown in Figure 3.32, which shows good modelling of the variables. The boundary behaviour is mainly due to binning effects.
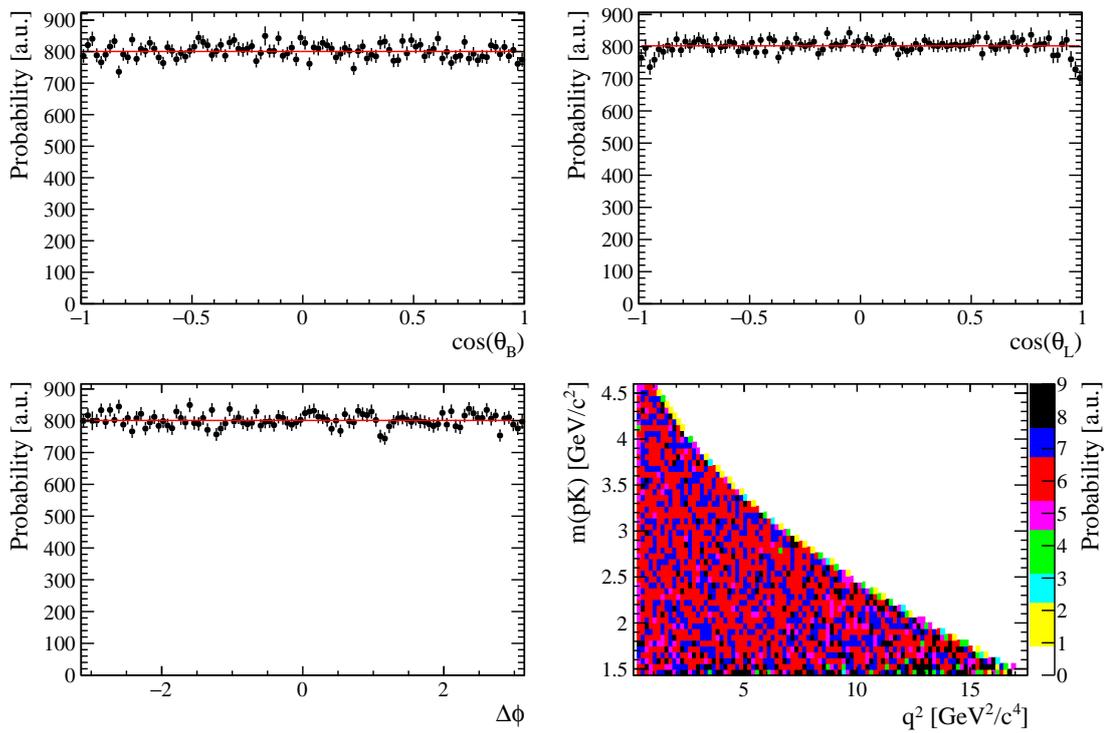
Figure 3.31: Validation of the efficiency model on fully reconstructed and selected MC, weighted by inverse of the efficiency model. We expect the distributions to be uniformly flat.
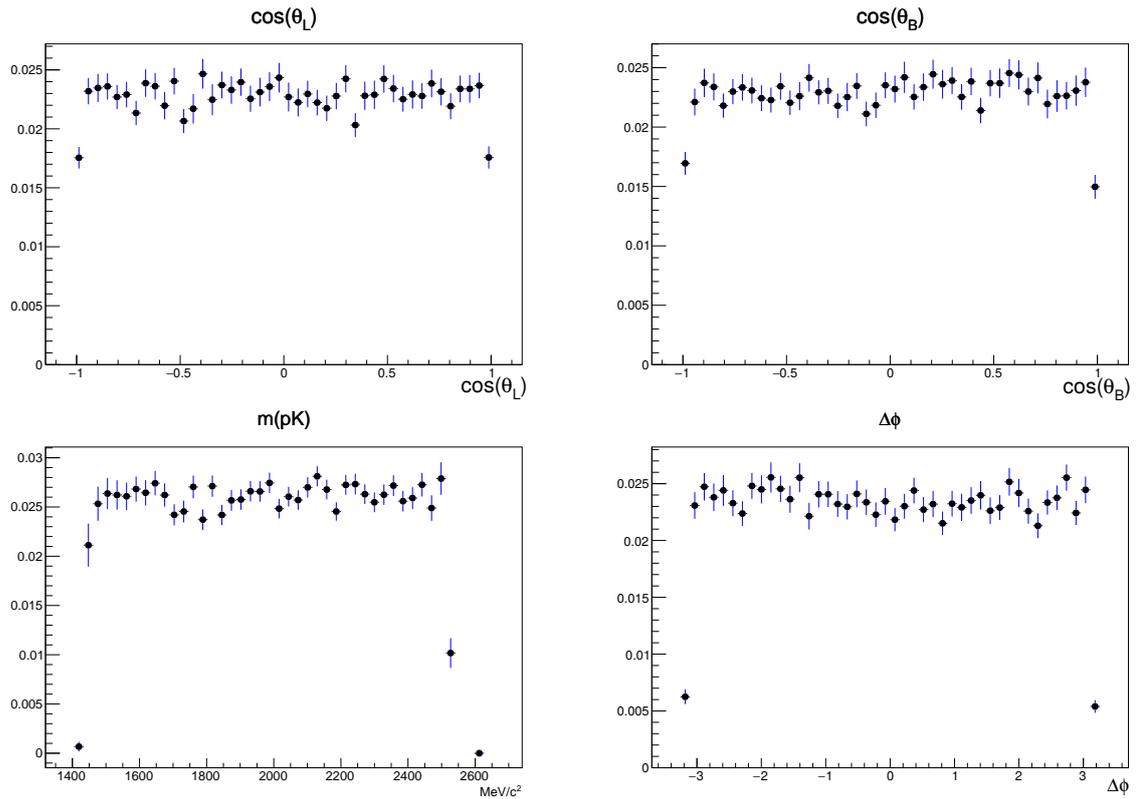
Figure 3.32: 1D validation checks on $\Lambda_b^0 \to J/\psi\,pK$ model, where an MC of selected $\Lambda_b^0 \to J/\psi\,pK$ events is corrected by the efficiency weight. Satisfactory modelling of the observables should give flat distributions when efficiency corrected.

3.8.4   Model Normalisation

As these efficiency models for $\Lambda_b^0 \rightarrow J/\psi pK$ and $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ correct only for the relative difference, we must correctly normalise it using the absolute phase-space efficiency calculated from MC. The phase-space efficiency, ϵ^{phsp} is,

$$\epsilon^{phsp} = \frac{N^{observed}}{N^{generated}}, \tag{3.29}$$

for an MC sample, and the observed yields on $\Lambda_b^0 \rightarrow J/\psi pK$ and $\Lambda_b^0 \rightarrow pK\mu^+\mu^-$ (N)can be corrected for efficiency to N^{corr} by the ratio of the phase space efficiency and average model efficiency,

$$N^{corr} = \sum_i \frac{sW_i}{\epsilon_i} \alpha, \tag{3.30}$$

where ϵ_i is the per correction from the model and,

$$\alpha = \frac{1}{\epsilon^{phsp}} \sum_{i=0}^{n} \frac{\epsilon_i}{n}. \tag{3.31}$$

The phase-space efficiencies and normalisation for each year and magnet polarity are shown in Table 3.19. All MC samples used have the kinematic re-weighting applied.

Year/polarity	ϵ_{tot}^{phsp}	normalisation factor
2011 MagDown	0.00340 ± 0.00008	0.00348 ± 0.00008
2011 MagUp	0.00340 ± 0.00008	0.00348 ± 0.00008
2012 MagDown	0.00339 ± 0.00007	0.00346 ± 0.00008
2012 MagUp	0.00334 ± 0.00007	0.00341 ± 0.00007

Table 3.19: Phase space efficiencies for $\Lambda_b^0 \rightarrow J/\psi pK$ with corresponding efficiency normalisation factor

3.9   Uncertainties

Statistical uncertainties

For the statistical uncertainty, the procedure is a somewhat similar to what is used in [89]. We have the uncertainty on the corrected yield, $N^{corrected}$ which is simply,

$$\sigma\left(N^{corrected}\right) = \sqrt{\sum_i^n \left(\frac{sW_i}{\epsilon_i}\right)^2},$$ (3.32)

where sW_i and ϵ_i are the s-weight and efficiency correction respectively, for the ith candidate. However, the s-weights are extracted from a fit where only the yields are floated, and the rest of the parameters fixed to their values from the nominal fit. We must also take into account the statistical uncertainty coming from the nominal fit itself.

To evaluate this, we perform toy MC experiments where all floating parameters of the fit are varied within their (Gaussian) uncertainties taking into account correlations between the parameters. This can be evaluated without refitting the distributions in data using Cholesky matrix decomposition. Taking our correlation matrix for the fit parameters, $\mathbf{C}$, we can write this as the lower triangular matrix and its conjugate transpose,

$$\mathbf{C} = \mathbf{L}\mathbf{L}^\intercal$$ (3.33)

where a vector of correlated random variables, $\mathbf{v}$ can be obtained by using the vector of uncorrelated random variables $\mathbf{u}$ and the lower triangular matrix,

$$\mathbf{L}\mathbf{u} = \mathbf{v}.$$ (3.34)

Specifically, for each toy iteration we generate a vector of random variables from the fit

parameters but with zero mean, perform the product in Equation 3.34 to obtain a vector of correlated random variables and then extract the random variable corresponding to N^{sig} from **v**. We also offset by the nominal value of N^{sig} to obtain a Gaussian distribution with mean N^{sig}.

We perform Gaussian fits to these distributions, which are obtained individually for the 2011 and 2012 datasets, and take the standard deviation as the contribution to the statistical uncertainty from the nominal fit. A full statistical uncertainty on the corrected yield is then the quadrature sum of this uncertainty with the uncertainty from (Equation 3.32) scaled by the efficiency. This gives,

$$\sigma\left(N^{corrected}\right) = \sqrt{\sum_i \left(\frac{w_i}{\epsilon_i}\right)^2 + \left(\sigma^{fitstat}(N)\frac{N^{corrected}}{N}\right)^2} \qquad (3.35)$$

where $\sigma^{fitstat}(N)$ is the standard deviation obtained from the Cholesky toy MC's. The distributions from the toys are shown in Figure 3.33, while the matrices **C** and **L** are shown in the Appenxix, Figures 2-5

(a)   (b)
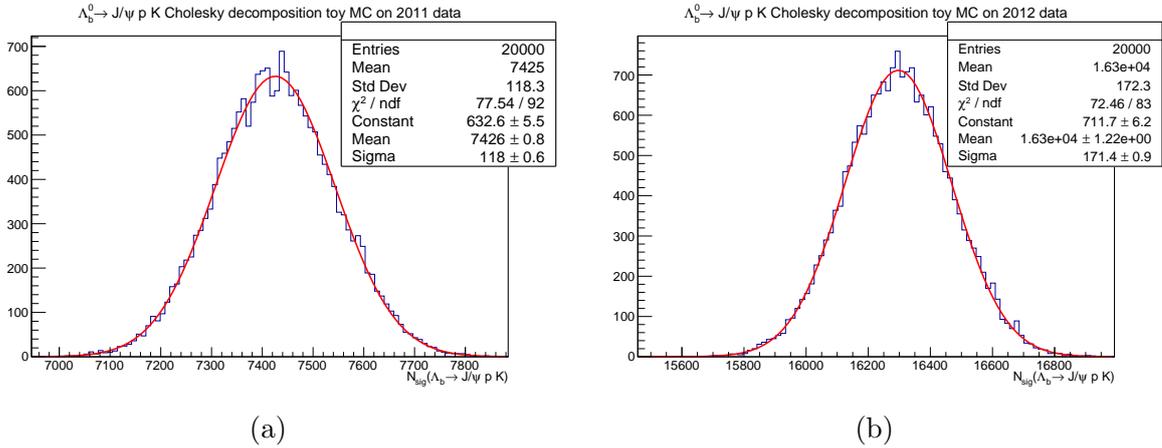
Figure 3.33: Covariant fluctuations on $\Lambda_b^0 \to J/\psi pK$ yield extracted from fits to 2011 (a) and 2012 (b) data

3.9.1   Systematic uncertainties

Several potential sources of systematic effects that could influence the efficiency determination are considered. These include the uncertainty on the measured lifetime of the Λ_b^0, the finite size of the MC samples used to calculate efficiencies, the stability of the efficiency modelling method (Section 3.8), the kinematic correction procedure (Section 3.2.1), the statistical distributions of MC and calibration samples as well as binning effects in the evaluation of PID efficiency, and in the choice of fit models for yield extraction (Section 3.5)

Systematic uncertainties from processes that involve event-by-event weighting, *e.g.* PID and kinematic re-weighting, are evaluated using toy Monte Carlo within the bootstrapping process described in detail in Section 3.9.3. The MC candidates from which the acceptance models are generated are re-sampled 1000 times and a new model generated for each re-sample. The corrected yield is calculated with each iteration and so allows us to study directly systematic effects on the measured branching fraction associated with the modelling and the sample size.

Within this process we can also incorporate the lifetime, PID and kinematic systematics by selecting a different model for each of these for each bootstrapping iteration. These models are themselves generated with toy MC methods, for instance, the various kinematic models are generated by randomly fluctuating the parameters of the $f_{\Lambda_b^0}/f_d$ functions under the assumption that they follow a Gaussian distribution.

3.9.2   Investigations on phase-space MC

The relative magnitudes of the individual systematic uncertainties as well as their contributions in each stage of the selection is not particularly discernible with the methods proposed in section 3.9.3 for the evaluation of the total uncertainty on the yields. Thus, we investigate each source individually in bins of q^2

Monte Carlo statistics

We propagate the statistical uncertainty on the MC samples used in the efficiency calculations to give an uncertainty on the efficiency. The results from this are presented in Table 3.20.

q^2 (GeV/c^2)	Efficiency	Uncertainty
0.1-2.0	0.03385	0.00057
2.0-4.0	0.04517	0.00069
4.0-6.0	0.05333	0.00086
6.0-8.0	0.05792	0.00110
8.0-11.0	0.06032	0.00127
11.0-12.5	0.05790	0.00270
12.5-15.0	0.05440	0.00334
15.0-17.5	0.04276	0.00945

Table 3.20: Systematic uncertainty from finite size of MC samples for full selection relative to candidates within geometric acceptance of LHCb.

Particle identification

The calibration data and MC samples are binned into relatively large bins in p_T and η where the change in efficiency across the bin boundaries may not necessarily be gradual. To account for the uncertainty on the PID efficiency arising from this, we vary the binning by take the binning scheme chosen in section 3.6.1 and merge adjacent bins. The PID efficiency is recalulated with this new scheme, and the one σ uncertainty taken to be the difference between the efficiencies from the two different binnings. This is combined with the uncertainty from sample statistics per bin, giving our full PID systematic to be,

$$\sigma\left(\epsilon_{PID}^{total}\right) = \sqrt{\left[\epsilon_{PID}^{\alpha} - \epsilon_{PID}^{\beta}\right]^2 + \sigma\left(\epsilon_{PID}^{\alpha}\right)^2}, \tag{3.36}$$

where ϵ_{PID}^{α} and ϵ_{PID}^{β} refer to efficiencies from the original and merged binning schemes respectively. The efficiency in p_T and η for the merged binning can be seen in Figure 3.34.

Tables 6 - 9 present the efficiencies for both binnings, the resulting uncertainties and total systematic error in bins of q^2.
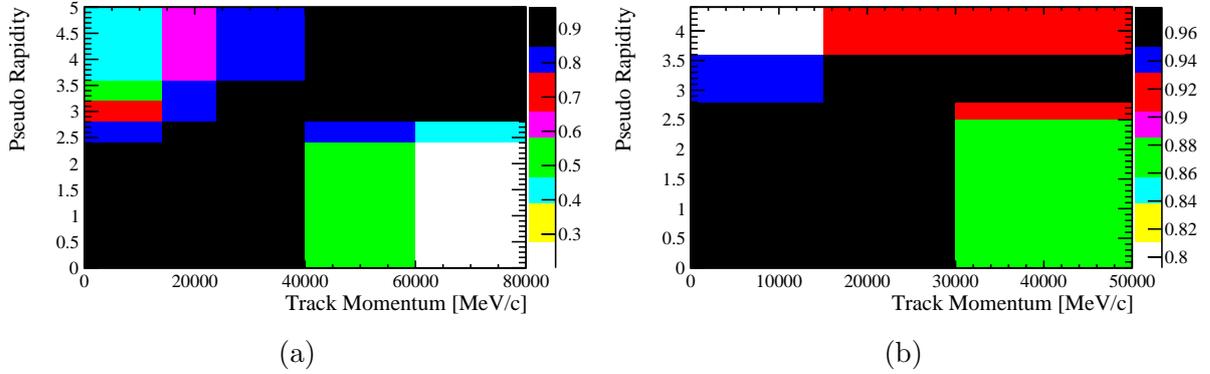
Figure 3.34: PID efficiency in merged bins for the proton (a) and kaon (b) in $\Lambda_b^0 \to pK\mu^+\mu^-$.

3.9.3   Systematic uncertainties evaluated with Monte Carlo toys

The systematic uncertainties related to the various MC weighting procedures are evaluated with toys, in which a different weighting model is chosen for each iteration and the efficiency model is regenerated with the new weights. In general, the creation of a new weighting model involves randomly varying any parameters in the weighting model that have known uncertainties. For instance, in the case of the kinematic re-weighting, in each toy iteration the experimentally measured [81] values that describe the $f_{\Lambda_b^0}/f_d$ functions are randomly varied to generate a new kinematic weighting model.

The weighting procedures for which we perform these toys are:

- The systematic uncertainty on the multi-dimensional efficiency models

- The re-weighting of Λ_b^0 production kinematics.

- The application of PID efficiency via event by event weighting.

- The evaluation of the systematic effect of the Λ_b^0 lifetime uncertainty

For each of these cases we perform 1000 toys. The systematic uncertainty for each case is determined by recalculating the branching ratio every iteration, by correcting the observed yields of $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ using the regenerated efficiency model.

The procedure for each case is described below. Figures 3.37-3.39 show the distributions from these toys on s-weights from the yield extraction fits. ϵ_{model} refers to the weights that each toy model produces.

As this method requires a set of real $\Lambda_b^0 \to pK\mu^+\mu^-$ candidates, it will be performed once the $\Lambda_b^0 \to pK\mu^+\mu^-$ set is unblinded. However, the procedure is carried out on the $\Lambda_b^0 \to J/\psi pK$ candidates, and the uncertainties on the corrected $\Lambda_b^0 \to J/\psi pK$ yield are presented.

Efficiency model

The sensitivity of the efficiency model's stability to the statistical size of the MC samples used to generate it is determined using a bootstrapping technique. The MC samples used to generate the model are re-sampled into new sets each iteration. With bootstrapping, sampling with replacement is explicitly allowed, and in this case, the size of the re-sampled sets is determined by a Poisson law with a mean equal to the original sample size. The results of these toys on $\Lambda_b^0 \to J/\psi pK$ can be seen in Figure 3.38.

Kinematic weighting

The uncertainties on the parameters in Equations 3.4 and 3.3 are propagated to an uncertainty on the kinematic weight. These weights are then randomly fluctuated in the toy MC experiment, reapplied to the efficiency model and the efficiency recalculated to determine the systematic on the efficiency deriving from the $f_{\Lambda_b^0}/f_d$ functions' uncertainties. All uncertainties are assumed to be Gaussian, and the results are displayed in Figure 3.35.

Particle identification

For the PID we evaluate the proton and kaon ID efficiencies for each candidate using the efficiency tables produced by `PIDCalib`, which bin the estimated single track efficiencies in momentum and pseudorapidity. We use the values from the fine binned histograms as the mean of a Gaussian and the Gaussian's sigma value is calculated using equation 3.36, where ϵ_{PID}^{α} and ϵ_{PID}^{β} are taken from the bins corresponding to the track's momentum and pseudorapidity, for the fine binned and course binned histograms respectively. A random value is chosen from the Gaussians for each bin to generate new, fluctuated PID histograms for each iteration of the toy MC. The distributions from the toys can be seen in Figure 3.36

Lifetime uncertainty

We take the uncertainty $(\sigma(\tau_{lhcb}))$ on the recent measurement performed by LHCb of the Λ_b^0 lifetime (τ_{lhcb}) and fluctuate the lifetime by the uncertainty for each toy iteration, re-weighting the MC sample by

$$\tau^{\Lambda_b^0} = \exp\left\{ -\frac{\tau^i}{\tau_{lhcb} \pm \sigma(\tau_{lhcb})} \right\} \tag{3.37}$$

where the measured value, $\tau_{lhcb} = 1.482 \pm 0.018 \pm 0.012\,\text{ps}$ [90]. The results of the toys can be seen in Figure 3.37
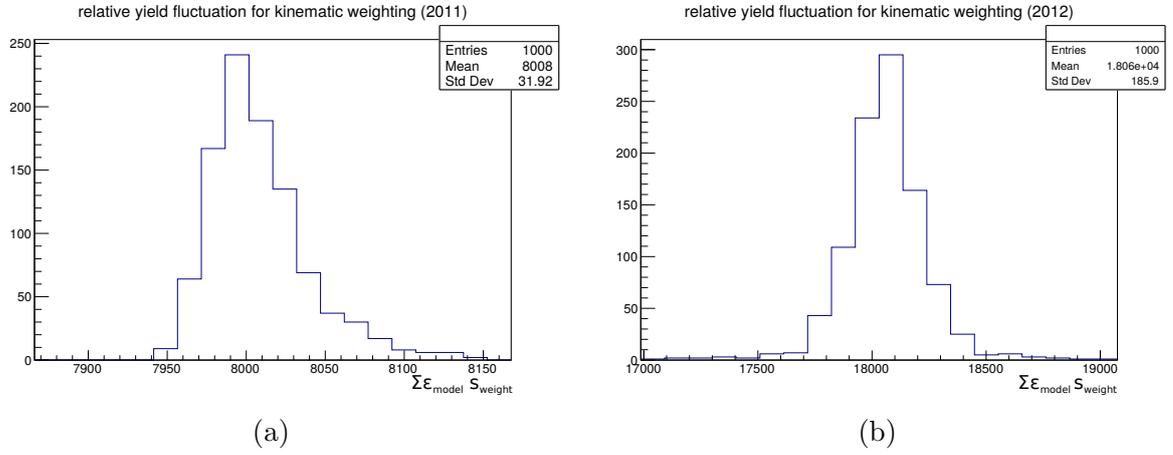
Figure 3.35: Toy MCs on efficiency model for $\Lambda_b^0 \to J/\psi pK$ to evaluate systematic uncertainty from kinematic weighting.
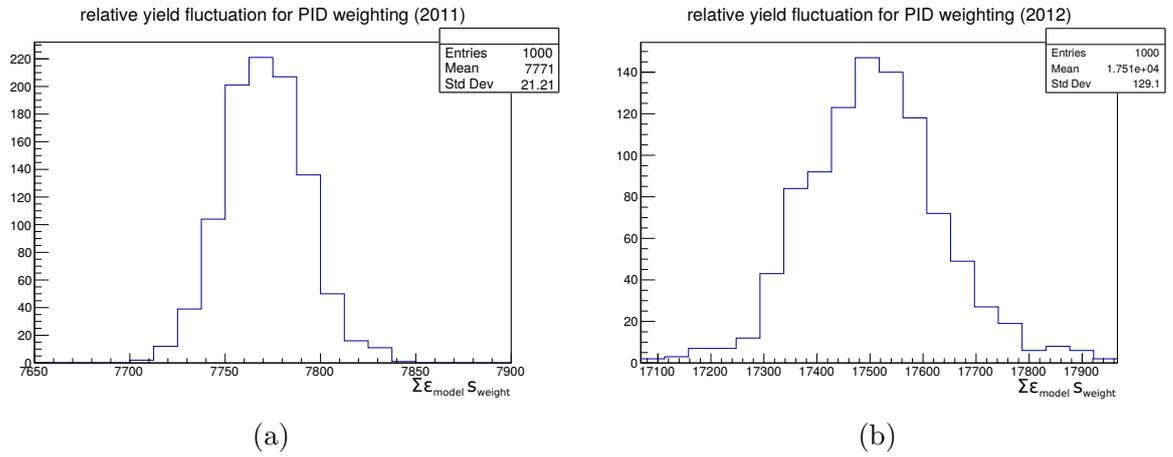
Figure 3.36: Toy MCs on efficiency model for $\Lambda_b^0 \to J/\psi pK$ to evaluate systematic uncertainty from PID weighting.
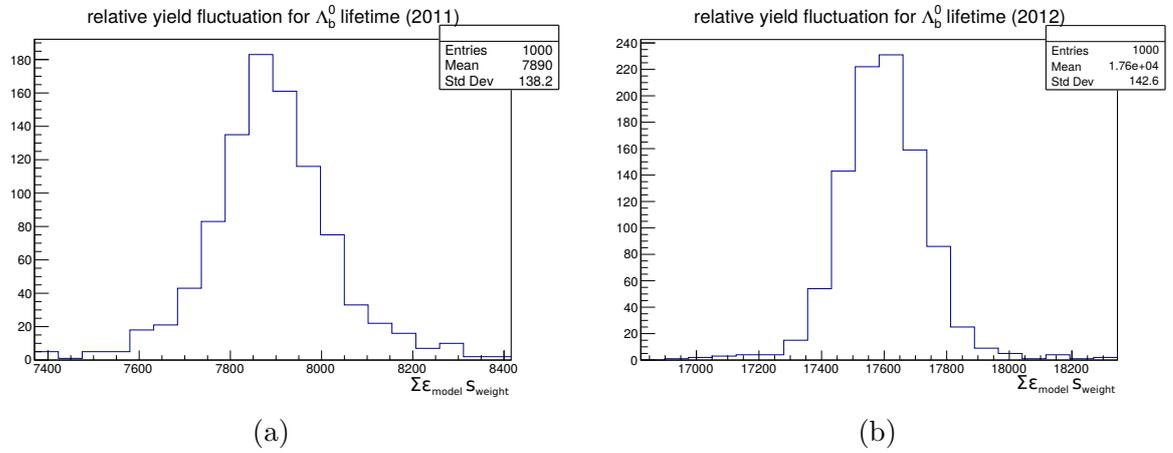
Figure 3.37: Toy MCs on efficiency model for $\Lambda_b^0 \to J/\psi pK$ to evaluate systematic uncertainty from Λ_b^0 lifetime.
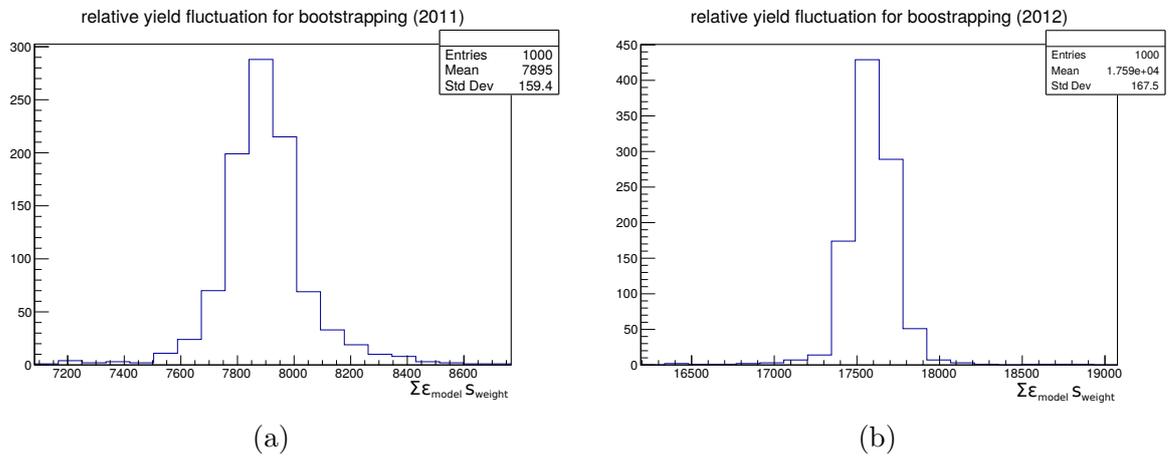
Figure 3.38: Bootstrapping toy MCs to evaluate systematic uncertainty from signal MC statistics for the generation of the $\Lambda_b^0 \to J/\psi pK$ efficiency model.
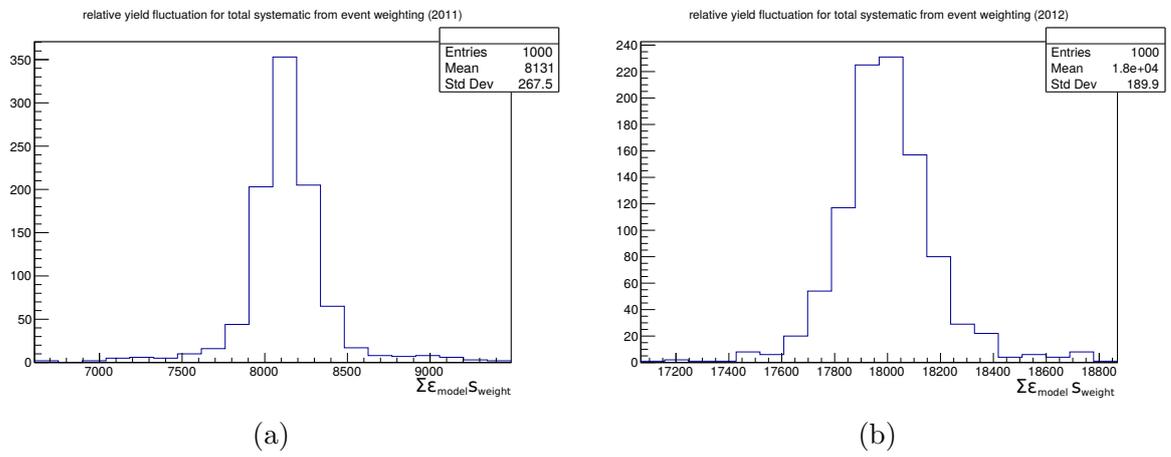
Figure 3.39: Combined systematic uncertainties from toy MC's on $\Lambda_b^0 \to J/\psi\, pK$ efficiency model.

3.9.4   Fit model

We investigate several sources of potential systematic, to account for differences in modelling of widths, background shape and level.

The default fit for $\Lambda_b^0 \to pK\mu^+\mu^-$ data will consist of a: DCB with yield free, central value constrained to that from $\Lambda_b^0 \to pK\mu^+\mu^-$ MC, with an additive offset determined by comparison of data/MC for $\Lambda_b^0 \to J/\psi pK$; the width is constrained to $\Lambda_b^0 \to pK\mu^+\mu^-$ MC, with multiplicative scaling applied from comparison of data/MC for $\Lambda_b^0 \to J/\psi pK$. The cut-off and slope parameters for the DCB tails are fixed according to $\Lambda_b^0 \to pK\mu^+\mu^-$ MC.

The cases considered for the $\Lambda_b^0 \to J/\psi pK$ fit are:

- Case one: Signal pdf where only the tail parameters are fixed.

- Case two: Changing of the Chebyshev polynomial used to model the combinatorial background from order three to order five.

- Case three: Modelling the B_s^0 background shape without applying PID cuts. This is to account for any biasing of the shape from mis-modelling of the PID variables in MC.

- Case four: Changing the B_s^0 pdf from a single Crystal Ball to a double.

- Case five: Removing the $\Lambda_b^0 \to J/\psi pK$ proton-kaon double reflection background component

- Case six: Adding a background component for $B^0 \to J/\psi K^{*0}$

For the $\Lambda_b^0 \to pK\mu^+\mu^-$ decays, the default background parametrisation is an exponential function, with contributions from $B_s^0 \to KK\mu\mu$ (shape parametrised according to MC using a Crystal Ball) and a much smaller component from $B^0 \to K\pi\mu\mu$ (shape fixed also

parametrised using a Crystal Ball). For the latter two components, their relative yields in the fit are fixed by their relative efficiencies obtained from simulation.

The $\Lambda_b^0 \to pK\mu^+\mu^-$ channel is still blinded and the systematics cannot be evaluated yet. Possible cases for the evaluation of $\Lambda_b^0 \to pK\mu^+\mu^-$ fit model systematics are:

- Case one: Fixing the Λ_b^0 mass from the result obtained by the $\Lambda_b^0 \to J/\psi pK$ fit rather than letting it float.

- Case two: Changing the signal pdf from a DCB to a double Gaussian.

- Case three: Replacing the exponential function to model the combinatorial background with a polynomial.

- Case four: Removing PID requirements on MC for modelling B_s^0 background shape.

In table 3.21 we detail the results from each $\Lambda_b^0 \to J/\psi pK$ case and how it differs from the nominal fit.

Case	Relative uncertainty		Description
	2011	2012	
1	4.98%	1.77%	All but tail parameters floating in signal pdf
2	0.98%	2.75%	Combinatorial changed to Chebyshev order five
3	11.5%	7.16%	PID requirements removed on B_s^0
4	1.52%	6.96%	B_s^0 pdf changed to DCB
5	0.05%	0.02%	pK swap shape removed
6	$< 10^{-2}\%$	$< 10^{-2}\%$	B^0 shape added
total	12.6%	10.5%	

Table 3.21: Fit systematics for $\Lambda_b^0 \to J/\psi pK$. These are added in quadrature to estimate the total contribution.

3.10   Results

3.10.1 $\Lambda_b^0 \to J/\psi\, pK$ corrected yield

The systematic effects associated with Λ_b^0 lifetime, particle identification, kinematic weighting and efficiency model are evaluated event by event and propagated through to the yields of $\Lambda_b^0 \to J/\psi\, pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$ via the aforementioned toy MC's and bootstrapping (Section 3.9.3). These uncertainties are combined in quadrature with the statistical uncertainty on the model normalisation (Section 3.8.4) and from the systematic uncertainties relating to the choices of fit models and parameter constraints (Section 3.9.4). Our total systematic uncertainty is therefore,

$$\sigma_{sys} = \sqrt{\sum \sigma_{toys}^2 + \sigma_{fitsys}^2 + \sigma_{effnorm}^2}, \tag{3.38}$$

where $\sum \sigma_{toys}^2$ sums over the uncertainty squared for each weighting toy, σ_{fitsys}^2 is the total fit model systematic and $\sigma_{effnorm}^2$ is the statistical uncertainty on the efficiency model normalisation. We use relative errors in this calculation, which gives us a total systematic uncertainty of 8.28% on the combined 2011 and 2012 $\Lambda_b^0 \to J/\psi\, pK$ yield.

Section 3.9 details the calculation of the statistical uncertainty, which, on the combined 2011 and 2012 data is 1.68%. This gives a total corrected $\Lambda_b^0 \to J/\psi\, pK$ yield of

$$N_{\Lambda_b^0 \to J/\psi\, pK}^{corrected} = (7.83 \pm 0.131 \pm 0.648) \times 10^6 \tag{3.39}$$

3.11   Post Unblinding Strategy

- After the mass window of the Λ_b^0 with the $\Lambda_b^0 \to pK\mu^+\mu^-$ selection is unblinded, the branching fraction of $\Lambda_b^0 \to pK\mu^+\mu^-$ will be measured by extracting the $\Lambda_b^0 \to$

$pK\mu^+\mu^-$ efficiency corrected yield using the same methods used for $\Lambda_b^0 \to J/\psi pK$. As mentioned in 3.8, a five dimension efficiency model is used, with the inclusion of q^2.

- Using s-weighted candidates from the fits to $\Lambda_b^0 \to J/\psi pK$ and $\Lambda_b^0 \to pK\mu^+\mu^-$, the relative differential branching fraction will be measured as,

$$\frac{\mathrm{d}\mathcal{B}\left(\Lambda_b^0 \to pK\mu^+\mu^-\right)/\mathrm{d}q^2}{\mathcal{B}\left(\Lambda_b^0 \to J/\psi pK\right)} = \frac{N_{pK\mu^+\mu^-}^{corrected}}{N_{J/\psi pK}^{corrected}} \frac{1}{\Delta q^2}, \qquad (3.40)$$

where Δq^2 is the width of q^2 interval and $N^{corrected}$ for each channel is,

$$N^{corrected} = \sum_i \frac{w_i}{\epsilon_i} \qquad (3.41)$$

where w_i and ϵ_i is the s-weight and efficiency correction weight for the candidate respectively.

- The un-blinding will enable us to evaluate the systematic uncertainty from the fit model, using the method for calculating uncertainties for $N_{J/\psi pK}^{corrected}$, which is detailed in section 3.9. When performing the calculation, all the relevant variations will be done simultaneously on $\Lambda_b^0 \to pK^-\mu^+\mu^-$ and $\Lambda_b^0 \to J/\psi pK^-$ decays and the effect of variation on the relative branching fraction will be estimated.

- The signal significance in each of the q^2 bins will be estimated using Wilk's theorem, where the fits with and without the signal will be compared. In these fits, the mass of the signal peak will be fixed to what the measured value in in the $\Lambda_b^0 \to J/\psi pK$ fit, while the widths will be fixed to the values determined from MC and scaled by the ratio of the widths in data and MC for $\Lambda_b^0 \to J/\psi pK$. The systematic uncertainties will be incorporated by a Gaussian convolved with the likelihood.

- In q^2 bins where we see significant signal, the background subtracted candidates (using s-weighting) will be corrected for efficiency using the efficiency model and used to measure the branching fraction.

- The q^2 distribution is unknown, and although we expect to see $\Lambda_b^0 \to pK\mu^+\mu^-$ there may be q^2 bins where no significant signal is obvserved. If this is the case then the upper limit will be calculated by intergrating the profile likelihood. However, efficiency correction cannot be performed in this case however and instead the phase space efficiency evaluated for that particular q^2 interval will be used. The systematic uncertainties will be incorporated by convolving the profile likelihood with a Gaussian of appropriate width.

Chapter 4

Conclusions

An analysis for the first observation and branching fraction measurement of the rare FCNC decay $\Lambda_b^0 \to pK\mu^+\mu^-$ has been performed using data collected at the LHCb experiment corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$. The analysis awaits approval from the LHCb Collaboration to unblind so that $\frac{BR(\Lambda_b^0 \to pK\mu^+\mu^-)}{\mathcal{B}(\Lambda_b^0 \to J/\psi\, pK)}$ can be measured.

The measurement is normalised to the control channel $\Lambda_b^0 \to J/\psi\, pK$, and candidates from both of these decays are corrected by multidimensional efficiency models to account for the unknown decay structure of $\Lambda_b^0 \to pK\mu^+\mu^-$. In anticipation of unblinding, an efficiency corrected yield of $\Lambda_b^0 \to J/\psi\, pK$ candidates has been measured, corresponding to $(7.83 \pm 0.131 \pm 0.648) \times 10^6$ $\Lambda_b^0 \to J/\psi\, pK$ decays.

It is hoped that the measurement of this decay will motivate the development of a more complete theoretical description, along with predictions for observables related to $\Lambda_b^0 \to pK\mu^+\mu^-$. A parallel analysis of the CP asymmetry of $\Lambda_b^0 \to pK\mu^+\mu^-$ at LHCb is also to be published soon. The analysis for the search for $\Lambda_b^0 \to p\pi\mu^+\mu^-$ is currently under review and a measurement of $\mathcal{B}(\Lambda_b^0 \to p\pi\mu^+\mu^-)$ along with $BR(\Lambda_b^0 \to pK\mu^+\mu^-)$ opens up interesting aspects for the measurement of V_{td}/V_{ts} [91].

Appendices

.1 Matrices for Cholesky toys

Number	Parameter
1	c_0^{comb}
2	c_1^{comb}
3	c_2^{comb}
4	m^{sig}
5	N_{comb}
6	N^{sig}
7	σ_2^{sig}
8	σ^{sig}

Table 1: Map of parameter number to name for $\Lambda_b^0 \to J/\psi pK$.

Parameter	1	2	3	4	5	6	7	8
1	1.	-0.205	0.243	-0.024	0.172	-0.015	-0.055	-0.014
2	-0.205	1.	-0.401	0.037	-0.195	0.131	0.187	0.094
3	0.243	-0.401	1.	0.001	0.404	-0.381	-0.368	-0.262
4	-0.024	0.037	0.001	1.	0.006	0.007	0.02	-0.055
5	0.172	-0.195	0.404	0.006	1.	-0.571	-0.54	-0.375
6	-0.015	0.131	-0.381	0.007	-0.571	1.	0.519	0.367
7	-0.055	0.187	-0.368	0.02	-0.54	0.519	1.	0.33
8	-0.014	0.094	-0.262	-0.055	-0.375	0.367	0.33	1.

Table 2: Covariance matrix for $\Lambda_b^0 \to J/\psi pK$ yield extraction fit to 2011 data

Parameter	1	2	3	4	5	6	7	8
1	1.	0.	0.	0.	0.	0.	0.	0.
2	-0.205	0.979	0.	0.	0.	0.	0.	0.
3	0.243	-0.359	0.901	0.	0.	0.	0.	0.
4	-0.024	0.033	0.021	0.999	0.	0.	0.	0.
5	0.172	-0.163	0.337	0.009	0.911	0.	0.	0.
6	-0.015	0.131	-0.367	0.01	-0.465	0.795	0.	0.
7	-0.055	0.18	-0.322	0.019	-0.431	0.221	0.791	0.
8	-0.014	0.093	-0.25	-0.053	-0.299	0.156	0.088	0.897

Table 3: Lower triangular matrix from the Cholesky decomposition of the $\Lambda_b^0 \rightarrow J/\psi pK$ 2011 fit parameter covariance matrix shown in table 2

Parameter	1	2	3	4	5	6	7	8
1	1.	-0.001	0.314	-0.015	0.248	-0.072	-0.141	-0.019
2	-0.001	1.	-0.002	0.	-0.001	0.001	0.002	0.
3	0.314	-0.002	1.	0.008	0.292	-0.321	-0.35	-0.149
4	-0.015	0.	0.008	1.	0.008	0.011	0.032	-0.07
5	0.248	-0.001	0.292	0.008	1.	-0.501	-0.467	-0.238
6	-0.072	0.001	-0.321	0.011	-0.501	1.	0.495	0.243
7	-0.141	0.002	-0.35	0.032	-0.467	0.495	1.	0.072
8	-0.019	0.	-0.149	-0.07	-0.238	0.243	0.072	1.

Table 4: Covariance matrix for $\Lambda_b^0 \rightarrow J/\psi pK$ yield extraction fit to 2012 data

Parameter	1	2	3	4	5	6	7	8
1	1.	0.	0.	0.	0.	0.	0.	0.
2	-0.001	1.	0.	0.	0.	0.	0.	0.
3	0.314	-0.002	0.949	0.	0.	0.	0.	0.
4	-0.015	-0.	0.013	1.	0.	0.	0.	0.
5	0.248	-0.001	0.226	0.009	0.942	0.	0.	0.
6	-0.072	0.001	-0.314	0.014	-0.438	0.839	0.	0.
7	-0.141	0.002	-0.322	0.034	-0.382	0.257	0.814	0.
8	-0.019	-0.	-0.151	-0.068	-0.211	0.123	-0.109	0.949

Table 5: Lower triangular matrix from the Cholesky decomposition of the $\Lambda_b^0 \rightarrow J/\psi pK$ 2012 fit parameter covariance matrix shown in table 4

.2   Tables for systematic checks

q^2 (GeV/c^2)	ϵ_{PID}^{α} (%)	$\sigma\left(\epsilon_{PID}^{\alpha}\right)$ (%)	ϵ_{PID}^{β} (%)	$\sigma\left(\epsilon_{PID}^{\beta}\right)$ (%)	$\sigma\left(\epsilon_{PID}^{total}\right)$ (%)
0.1-2	77.934	0.045	78.763	0.019	0.83
2-4	77.488	0.048	78.298	0.021	0.81
4-6	76.557	0.055	77.639	0.024	1.08
6-8	75.334	0.069	76.640	0.029	1.31
11-12.5	71.51	0.017	73.454	0.074	1.94
15-17.5	68.39	0.058	72.37	0.29	3.98
Integrated	76.397	0.024	77.513	0.010	1.12

Table 6: PID efficiencies and uncertainties for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagUp' polarity, 2011

q^2 (GeV/c^2)	ϵ_{PID}^{α} (%)	$\sigma\left(\epsilon_{PID}^{\alpha}\right)$ (%)	ϵ_{PID}^{β} (%)	$\sigma\left(\epsilon_{PID}^{\beta}\right)$ (%)	$\sigma\left(\epsilon_{PID}^{total}\right)$ (%)
0.1-2	76.871	0.027	78.448	0.015	1.58
2-4	79.368	0.029	78.030	0.016	1.34
4-6	75.381	0.033	77.212	0.018	1.83
6-8	74.230	0.041	76.320	0.023	2.09
11-12.5	69.670	0.100	72.856	0.057	3.19
15-17.5	66.880	0.400	71.17	0.25	4.31
Integrated	75.258	0.015	77.1587	0.0081	1.90

Table 7: PID efficiencies and uncertainties for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagDown' polarity, 2011

q^2 (GeV/c^2)	ϵ_{PID}^{α} (%)	$\sigma\left(\epsilon_{PID}^{\alpha}\right)$ (%)	ϵ_{PID}^{β} (%)	$\sigma\left(\epsilon_{PID}^{\beta}\right)$ (%)	$\sigma\left(\epsilon_{PID}^{total}\right)$ (%)
0.1-2	79.650	0.021	80.752	0.011	1.10
2-4	79.325	0.022	80.509	0.012	1.18
4-6	78.449	0.025	79.862	0.014	1.41
6-8	77.019	0.031	78.787	0.017	1.77
11-12.5	74.121	0.078	76.360	0.043	2.24
15-17.5	70.970	0.290	73.89	0.016	2.93
Integrated	78.264	0.011	79.6773	0.0060	1.41

Table 8: PID efficiencies and uncertainties for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagUp' polarity, 2012

q^2 (GeV/c^2)	ϵ_{PID}^{α} (%)	$\sigma\left(\epsilon_{PID}^{\alpha}\right)$ (%)	ϵ_{PID}^{β} (%)	$\sigma\left(\epsilon_{PID}^{\beta}\right)$ (%)	$\sigma\left(\epsilon_{PID}^{total}\right)$ (%)
0.1-2	80.937	0.021	81.942	0.011	1.01
2-4	80.722	0.022	81.769	0.012	1.05
4-6	79.995	0.025	81.236	0.014	1.24
6-8	78.882	0.031	80.421	0.017	1.54
11-12.5	75.104	0.079	77.621	0.044	2.52
15-17.5	72.58	0.330	75.40	0.19	2.84
Integrated	79.749	0.011	81.0331	0.0060	1.28

Table 9: PID efficiencies and uncertainties for $\Lambda_b^0 \to pK\mu^+\mu^-$ 'MagDown' polarity, 2012

q^2 (GeV/c^2)	Efficiency	Uncertainty(%)
0.1-2	0.07688	2.58
2-4	0.09032	2.26
4-6	0.09626	2.08
6-8	0.09726	2.03
8-11	0.09593	2.10
11-12.5	0.08881	2.30
12.5-15	0.08173	2.42
15-17.5	0.06517	2.98

Table 10: Λ_b^0 lifetime systematic uncertainty as a percentage of efficiency for reconstructed candidates passing stripping and pre-selection, relative to candidates within the geometric acceptance of LHCb

q^2 (GeV/c^2)	Efficiency	Uncertainty(%)
0.1-2	0.5351	0.04
2-4	0.6039	0.05
4-6	0.6723	0.08
6-8	0.7297	0.09
8-11	0.7809	0.09
11-12.5	0.8225	0.1
12.5-15	0.8573	0.1
15-17.5	0.8653	0.1

Table 11: Λ_b^0 lifetime systematic uncertainty as a percentage of efficiency for TOS candidates relative to reconstructed candidates that pass stripping and pre-selection. LHCb

.3 Generator level decay distribution in efficiency parametrisation

The efficiency for $\Lambda_b^0 \to pK\mu^+\mu^-$ is parameterised in a multidimensional space and fully determines the kinematics of the decay, using an expansion in legendre polynomials. This method assumes that all variables of the model were generated in with flat distributions in MC. In the phase-space MC for four-body $\Lambda_b^0 \to pK\mu^+\mu^-$ this is the case for the anglular distributions, but not for the two mass variables $m(pK)$ and q^2.

In order to make the parameterisation method valid, an analytical approach is used to generate weights to correct the two dimensional mass space of q^2 and $m(pK)$ to be uniform. Starting with the genereral expression to describe a multi-body phase-space [7], the $m(pK)$, q^2 distribution is described by,

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}m(pK^-)\mathrm{d}q^2} = |p_1||p_3||q_J|$$

where $|p_1|$ is the magnitude of proton's 3-momentum in the pK rest frame, $|p_3|$ is the magnitude muon's 3-momentum in dimuon rest frame and $|q_J|$ is the 3-momentum magnitude of the dimuon system in the rest frame of the Λ_b^0. These are related to $m(pK^-)$ and q^2 as,

$$|p_1| = \frac{\left[\left(m(pK^-)^2 - (m_p + m_K)^2\right)\left(m(pK^-)^2 - (m_p - m_K)^2\right)\right]^{1/2}}{2m(pK^-)},$$

$$|p_3| = \frac{\left[\left(q^2 - (2m_\mu)^2\right)(q^2)\right]^{1/2}}{2\sqrt{q^2}},$$

$$|q_J| = \frac{\left[\left(m(\Lambda_b^0)^2 - \left(m(pK^-) + \sqrt{q^2}\right)^2\right)\left(m(\Lambda_b^0)^2 - \left(m(pK^-) - \sqrt{q^2}\right)^2\right)\right]^{1/2}}{2m(\Lambda_b^0)}.$$

Weighting events by $1/\mathrm{d}\Gamma/(\mathrm{d}m(pK^-)\mathrm{d}q^2)$ will produce a uniform distribution in $m(pK^-)$-q^2. The variables $m(pK)$ and q^2 also need to be transformed to between -1 and +1, while preserving the uniformity. This is done by multiplying the phase-space distribution weight by the length of the allowed q^2 interval.

Bibliography

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics A **22** (1961) 579.

[2] CDF Collaboration, F. Abe *et al.*, *Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab*, Phys. Rev. Lett. **74** (1995) 2626.

[3] ATLAS Collaboration and CMS Collaboration, 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. **114** (2015) 191803.

[4] *Table of standard model particles*, `http://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg`. Accessed: 2016.

[5] D. Griffiths, *Introduction to elementary particles, 1987*, John Willey & Sons Inc.

[6] The LHCb Collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[7] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

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

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[10] R. H. Bernstein and P. S. Cooper, *Charged lepton flavor violation: An experimenters guide*, Physics Reports **532** (2013), no. 2 27 , Charged Lepton Flavor Violation: An Experimenter's Guide.

[11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[12] H. Burkhardt *et al.*, *First evidence for direct cp violation*, Physics Letters B **206** (1988), no. 1 169.

[13] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[14] The LHCb Collaboration, S. Ali, *Measurement of the CKM angle γ at LHCb*, in *50 Years of CP Violation London, UK, July 10-11, 2014*, 2014. `arXiv:1410.8779`.

[15] The LHCb Collaboration, *Determination of the quark coupling strength— vub— using baryonic decays*, Nature Physics **11** (2015), no. 9 743.

[16] CKMfitter Group, J. Charles *et al.*, *Current status of the standard model ckm fit and constraints on $\Delta f = 2$ new physics*, Phys. Rev. D **91** (2015) 073007.

[17] D. V. Perepelitsa, *Sakharov conditions for baryogenesis*, PHYS G6050 Particle Phenomenology class project, Columbia University, Department of Physics (2008).

[18] A. D. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,61(1991)].

[19] R. Fleischer, *CP violation beyond the standard model*, in *Heavy flavor physics. Proceedings, 7th International Symposium, Santa Barbara, USA, July 7-11, 1997*, pp. 155–164, 1997. `arXiv:hep-ph/9709291`.

[20] L. Maiani, *The GIM Mechanism: origin, predictions and recent uses*, in *Proceedings, 48th Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 2-9, 2013*, pp. 3–16, 2013. `arXiv:1303.6154`.

[21] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[22] G. Barenboim, F. J. Botella, and O. Vives, *Tree-level flavor-changing neutral currents in the b system: From* CP *asymmetries to rare decays*, Phys. Rev. D **64** (2001) 015007.

[23] A. Eichhorn *et al.*, *The Higgs Mass and the Scale of New Physics*, JHEP **04** (2015) 022, `arXiv:1501.02812`.

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

[25] E. Kraus and S. Groot Nibbelink, *Renormalization of the electroweak standard model*, in *4th National Summer School for German-speaking Graduate Students of Theoretical Physics Saalburg, Germany, August 31-September 11, 1998*, 1998. `arXiv:hep-th/9809069`.

[26] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the LHC*, Ann. Rev. Nucl. Part. Sci. **65** (2015) 113, `arXiv:1501.03309`.

[27] K. G. Chetyrkin, M. Misiak, and M. Munz, *Weak radiative B meson decay beyond leading logarithms*, Phys. Lett. **B400** (1997) 206, `arXiv:hep-ph/9612313`, [Erratum: Phys. Lett.B425,414(1998)].

[28] T. DeGrand and C. DeTar, *Lattice methods for quantum chromodynamics*, World Scientific, 2006.

[29] V. M. Braun, *Light-cone sum rules*, arXiv preprint hep-ph/9801222 (1998).

[30] C. De Melis, *The CERN accelerator complex. Complexe des acclrateurs du CERN*, , General Photo.

[31] L. Evans and P. Bryant, *Lhc machine*, Journal of Instrumentation **3** (2008), no. 08 S08001.

[32] G. Aad *et al.*, *The atlas experiment at the cern large hadron collider*, Journal of Instrumentation **3** (2008), no. S08003.

[33] The CMS Collaboration, *The cms experiment at the cern lhc*, JInst **3** (2008), no. 08 S08004.

[34] K. Aamodt *et al.*, *The alice experiment at the cern lhc*, Journal of Instrumentation **3** (2008), no. 08 S08002.

[35] A. A. Alves Jr *et al.*, *The lhcb detector at the lhc*, Journal of instrumentation **3** (2008), no. 08 S08005.

[36] G. Anelli *et al.*, *The totem experiment at the cern large hadron collider*, Journal of Instrumentation **3** (2008), no. 08 S08007.

[37] K. Noda, *The lhcf experiment*, New Journal of Physics **11** (2009) 065012.

[38] MoEDAL Collaboration, *Technical design report of the moedal experiment*, Technical design report CERN-LHCC-2009-006, CERN, Geneva. http:// cds. cern. ch/ record **1181486** (2009).

[39] M. Lamont, *Status of the lhc*, Journal of Physics: Conference Series **455** (2013), no. 1 012001.

[40] L. Rossi, *Superconducting magnets for the lhc main lattice*, IEEE transactions on applied superconductivity **14** (2004), no. 2 153.

[41] M. Lamont, *Status of the large hadron collider*, Journal of Physics: Conference Series **347** (2012), no. 1 012001.

[42] The LHCb Collaboration, *Prompt charm production in pp collisions at*, Nuclear Physics B **871** (2013), no. 1 1 .

[43] The LHCb Collaboration, *Measurement of at in the forward region*, Physics Letters B **694** (2010), no. 3 209 .

[44] *b-quark production*, `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`. Accessed: 2016.

[45] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, , Comments: 5 pages, contribution to the ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders, CERN, Geneva, Switzerland, 18-22 Mar 2013.

[46] LHCb Collaboration, *Lhcb velo tdr: Vertex locator*, Technical design report (2001).

[47] The LHCb Collaboration, *Opposite-side flavour tagging of b mesons at the lhcb experiment*, The European Physical Journal C **72** (2012), no. 6 1.

[48] *Velo event display*, `http://lhcb-reconstruction.web.cern.ch/lhcb-reconstruction/Panoramix/PRplots/CERN-Press/default.html`. Accessed: 2016.

[49] R. Aaij *et al.*, *Performance of the lhcb vertex locator*, Journal of Instrumentation **9** (2014), no. 09 P09007.

[50] M. Brice, *LHCb installs its precision silicon detector, the VELO. L'quipe de LHCb installe son dtecteur de prcision au silicium, le VELO*, Oct, 2007.

[51] The LHCb Collaboration, *Rich technical design report*, tech. rep., CERN-LHCC-2000-037, 2000.

[52] D. Perego, L. R. Group *et al.*, *The rich detectors of the lhcb experiment*, Nuclear Physics B-Proceedings Supplements **197** (2009), no. 1 358.

[53] C. P. Buszello, {*LHCB*} {*RICH*} *pattern recognition and particle identification performance*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **595** (2008), no. 1 245 , {RICH} 2007Proceedings of the Sixth International Workshop on Ring Imaging Cherenkov Detectors.

[54] The LHCb Collaboration, *Performance of the lhcb rich detector at the lhc*, The European Physical Journal C **73** (2013), no. 5 1.

[55] LHCb Collaboration, *Lhcb magnet technical design report*, CERN/LHCC **7** (2000) 2000.

[56] The LHCb Collaboration, *Lhcb detector performance*, International Journal of Modern Physics A **30** (2015), no. 07 1530022, `arXiv:http://www.worldscientific.com/doi/pdf/10.1142/S0217751X15300227`.

[57] P. Perret, X. Vilasís-Cardona *et al.*, *Performance of the lhcb calorimeters during the period 2010-2012*, in *Journal of Physics: Conference Series*, vol. 587, p. 012012, IOP Publishing, 2015.

[58] A. A. A. Jr *et al.*, *Performance of the lhcb muon system*, Journal of Instrumentation **8** (2013), no. 02 P02022.

[59] F. Sauli, *Gem: A new concept for electron amplification in gas detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **386** (1997), no. 2 531.

[60] J. Albrecht *et al.*, *The lhcb trigger system: Present and future*, in *Journal of Physics: Conference Series*, vol. 623, p. 012003, IOP Publishing, 2015.

[61] R. Aaij *et al.*, *The lhcb trigger and its performance in 2011*, Journal of Instrumentation **8** (2013), no. 04 P04022.

[62] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[63] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[64] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[65] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[66] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, `arXiv:hep-ph/0506026`.

[67] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[68] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[69] B. Couturier, E. Kiagias, and S. B. Lohn, *Systematic profiling to monitor and specify the software refactoring process of the lhcb experiment*, Journal of Physics: Conference Series **513** (2014), no. 5 052020.

[70] *Description of electromagnetic physics lists in geant4*, `http://geant4.web.cern.ch/geant4/`. Accessed: 2016.

[71] G. Hiller, M. Knecht, F. Legger, and T. Schietinger, *Photon polarization from helicity suppression in radiative decays of polarized Lambda(b) to spin-3/2 baryons*, Phys. Lett. **B649** (2007) 152, `arXiv:hep-ph/0702191`.

[72] T. Mannel and S. Recksiegel, *Flavor changing neutral current decays of heavy baryons: The Case $\Lambda_b^0 \to \Lambda \gamma$*, J. Phys. **G24** (1998) 979, `arXiv:hep-ph/9701399`.

[73] LHCb collaboration, R. Aaij *et al.*, *Study of the productions of Λ_b^0 and $\bar{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \to J/\psi p K^-$ branching fraction*, Chin. Phys. C **40** (2015) 011001, `arXiv:1509.00292`.

[74] L. Mott and W. Roberts, *Rare dileptonic decays of Λ_b in a quark model*, Int. J. Mod. Phys. **A27** (2012) 1250016, `arXiv:1108.6129`.

[75] D. Melikhov and B. Stech, *Weak form factors for heavy meson decays: An update*, Phys. Rev. D **62** (2000) 014006.

[76] LHCb collaboration, *Davinci-the lhcb analysis program*, Web page http://lhcbrelease-area. web. cern. ch/LHCb-release-area/DOC/davinci **144**.

[77] M. Clemencic *et al.*, *The lhcb simulation application, gauss: design, evolution and experience*, in *Journal of Physics: Conference Series*, vol. 331, p. 032023, IOP Publishing, 2011.

[78] L. Anderlini *et al.*, *The pidcalib package*, tech. rep., 2016.

[79] I. Belyaev *et al.*, *Python-based physics analysis environment for lhcb*, tech. rep., CERN, 2004.

[80] M. Feindt and U. Kerzel, *The neurobayes neural network package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **559** (2006), no. 1 190.

[81] LHCb collaboration, 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 **08** (2014) 143, `arXiv:1405.6842`.

[82] M. Pivk and F. Diberder, *splot: A statistical tool to unfold data distributions, 275 nucl. instrum. meth. a555 (2005) 356*, arXiv preprint physics/0402083 **276**.

[83] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. `arXiv:physics/0308063`.

[84] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[85] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, ArXiv Physics e-prints (2003) `arXiv:physics/0306116`.

[86] The LHCb Collaboration, R. Aaij *et al.*, *Measurements of the $\Lambda_b^0 \to J/\psi\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B724** (2013) 27, `arXiv:1302.5578`.

[87] E. Parzen, *On estimation of a probability density function and mode*, The annals of mathematical statistics **33** (1962), no. 3 1065.

[88] A. Poluektov, *Kernel density estimation of the multidimensional efficiency profile*, Tech. Rep. LHCb-INT-2014-008. CERN-LHCb-INT-2014-008, CERN, Geneva, Feb, 2014.

[89] M. Whitehead *et al.*, *Observation of $B^0 \to \bar{D}^0 K^+ K^-$*, .

[90] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the Λ_b^0 baryon lifetime*, Phys. Rev. Lett. **111** (2013) 102003, `arXiv:1307.2476`.

[91] C. S. Kim, T. Morozumi, and A. I. Sanda, *Determination of $-V(td) / V(ts)-$ within the standard model*, in *B physics and CP violation. Proceedings, 2nd International Conference, Honolulu, USA, March 24-27, 1997*, pp. 387–391, 1997. `arXiv:hep-ph/9706380`.