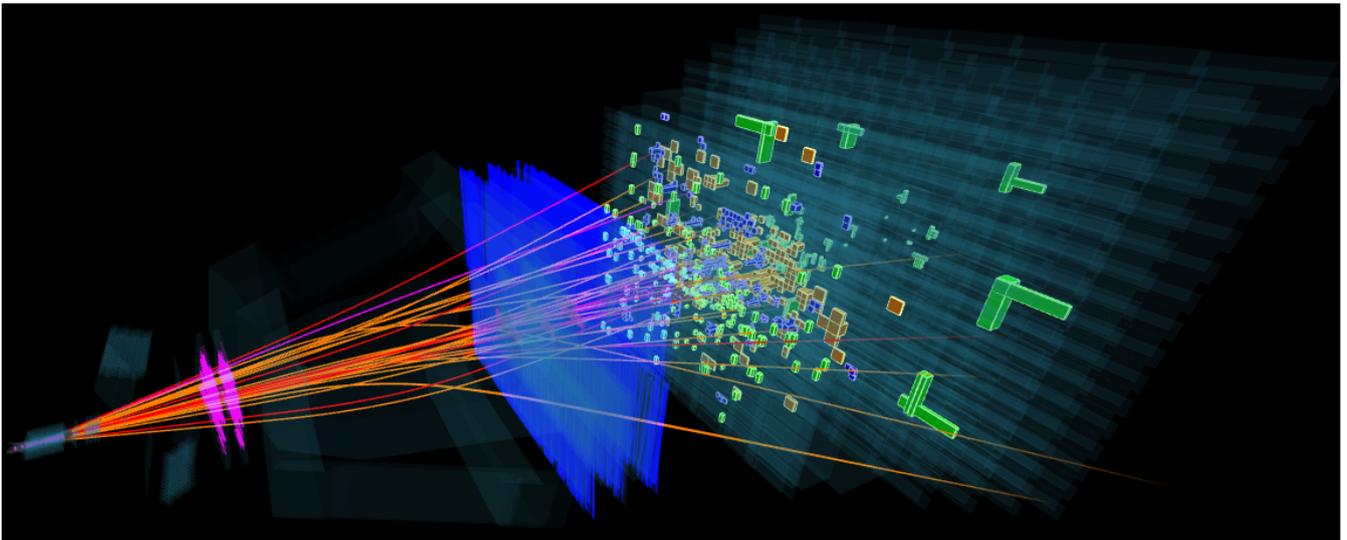


# Development of the Selection Procedure for a Measurement of the Ratio of Branching Fractions $R_{pK} = \mathcal{B}(\Lambda_b^0 \rightarrow pK\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \rightarrow pKe^+e^-)$ as a Test of Lepton Universality

Sean Kirwan  
Department of Physics and Astronomy  
University of Southampton  
Southampton  
SO17 1BJ  
UK

15th May 2016



## Abstract

A selection procedure for the non-rare decays  $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \ell^+\ell^-)$  (where  $\ell = \mu, e$ ) is presented, involving PID selection, removal of peaking backgrounds and multivariate analysis. Unforeseen delays in the acquisition of up to date stripping lines have slowed progress, but a verification of the selection through the measurement of  $\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+\mu^-))/\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow e^+e^-))$  is expected very soon. The same selection will then be used to isolate the rare decays  $\Lambda_b^0 \rightarrow pK^- \ell^+\ell^-$  (where  $\ell = \mu, e$ ) in order to make the first ever observation of the decay  $\Lambda_b^0 \rightarrow pK^- e^+e^-$ , and to test lepton universality by making the first ever measurement of  $R_{pK} = \mathcal{B}(\Lambda_b^0 \rightarrow pK^- \mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \rightarrow pK^- e^+e^-)$ . The expected sources of systematic error are also discussed.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Overview of the Standard Model . . . . .	2
1.2	Charged Weak Interactions . . . . .	3
1.3	The $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ Decay . . . . .	3
1.4	Effective Hamiltonian and New Physics . . . . .	4
1.4.1	Extended Higgs Interactions . . . . .	5
1.4.2	Supersymmetric Effects . . . . .	5
<b>2</b>	<b>Motivation for Measuring <math>R_{pK}</math></b>	<b>5</b>
2.1	Lepton Universality . . . . .	5
2.2	Testing Lepton Universality Using $R_{pK}$ . . . . .	5
2.3	Previous Lepton Universality Measurements . . . . .	6
<b>3</b>	<b>The LHCb Detector</b>	<b>6</b>
3.1	Vertex Locator (VELO) . . . . .	8
3.2	Dipole Magnet . . . . .	8
3.3	Ring Imaging Cherenkov (RICH) Detectors . . . . .	9
3.4	Tracking Stations . . . . .	10
3.5	Calorimeter System . . . . .	10
3.6	Muon Chambers . . . . .	10
3.7	Trigger System . . . . .	11
3.7.1	L0 Trigger . . . . .	11
3.7.2	High level trigger . . . . .	11
<b>4</b>	<b>Analysis Strategy</b>	<b>11</b>
4.1	Data . . . . .	12
4.2	The $J/\psi$ Control Channel . . . . .	12
4.3	Bremsstrahlung Radiation . . . . .	13
4.3.1	Energy Recovery Procedure . . . . .	13
4.3.2	HOP . . . . .	15
4.4	Dilepton Invariant Mass ( $q^2$ ) . . . . .	15
<b>5</b>	<b>Event Selection</b>	<b>17</b>
5.1	Particle Identification (PID) . . . . .	17
5.2	Rejection of Peaking Backgrounds . . . . .	18
5.2.1	$B_s^0 \rightarrow \phi(\rightarrow pK^-)J/\psi(\rightarrow \ell^+\ell^-)$ . . . . .	18
5.2.2	$\bar{B}^0 \rightarrow J/\psi(\rightarrow \ell^+\ell^-)K^*(\rightarrow K\pi)$ . . . . .	20
<b>6</b>	<b>Multivariate Analysis</b>	<b>22</b>
<b>7</b>	<b>Fitting</b>	<b>25</b>
7.1	Use of Probability Density Functions . . . . .	25
7.2	Determination of Signal Yield . . . . .	26
<b>8</b>	<b>Systematic Errors and Efficiencies</b>	<b>26</b>
<b>9</b>	<b>Conclusion</b>	<b>27</b>

# 1 Introduction

This report presents the event selection procedure that has been developed to facilitate a measurement of the parameter  $R_{pK}$ , defined as the ratio of the branching fractions of the rare decays  $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$  and  $\Lambda_b^0 \rightarrow pK^-e^+e^-$ . The theoretical principles behind these decays will be discussed in this introduction, followed by the motivation for a measurement of  $R_{pK}$  in section 2. The LHCb detector will be described in section 3, with the strategy for the analysis and its current status, as well as the main expected sources of systematic uncertainty, following in later sections.

## 1.1 Overview of the Standard Model

All of the known processes in particle physics are governed by a relativistic quantum field theory known as the Standard Model (SM). The SM specifies two types of fundamental particle: fermions, with spin 1/2, and bosons, with integer spin. The fermion category includes leptons, which are further divided into three generations, and quarks, which exist in three mass generations and six flavours. The leptons and quarks of the SM are shown in tables 1 and 2 respectively.

Table 1: The three generations of leptons in the Standard Model.

Generation	Lepton	Charge (e)
1	electron $e$	-1
	electron neutrino $\nu_e$	0
2	muon $\mu$	-1
	muon neutrino $\nu_\mu$	0
3	tau $\tau$	-1
	tau neutrino $\nu_\tau$	0

Table 2: The three generations and six flavours of quarks in the Standard Model.

Generation	Quark Flavour	Charge (e)
1	up	$+\frac{2}{3}$
	down	$-\frac{1}{3}$
2	charm	$+\frac{2}{3}$
	strange	$-\frac{1}{3}$
3	top	$+\frac{2}{3}$
	bottom	$-\frac{1}{3}$

The SM also describes three fundamental forces mediated by gauge bosons. These are the electromagnetic, weak and strong forces, and are mediated by the photon ( $\gamma$ ), the vector bosons  $W^\pm$  and  $Z$ , and the gluon ( $g$ ) respectively. All leptons and quarks couple to the vector bosons and therefore participate in weak interactions. Quarks and the charged leptons interact with the photon and therefore feel the electromagnetic force.

Only quarks have colour charge, which couples them to gluons, and they are affected by the strong force to the extent that quarks cannot exist as free particles. This property, known as quark confinement, requires that they exist in colour neutral states known as hadrons, made up of mesons (quark-antiquark pair) and baryons (three quarks or three antiquarks).

The final fundamental particle described by the standard model is the spinless, scalar Higgs boson, which couples to all massive particles thereby giving them that property. Its discovery at the LHC in 2012 provided the final piece of experimental evidence for the SM.

The Standard Model is the current best description of fundamental particles and their interactions, but it is known to be incomplete. For example, the SM provides no insight into the fourth fundamental force, gravity, and shows no signs of being compatible with general relativity. It also has nothing to say about the large matter/antimatter imbalance in the universe or dark matter/energy. A number of theories have been proposed that go beyond the Standard Model, two of which will be introduced briefly in sections 1.4.1 and 1.4.2, but so far none have been verified experimentally.

## 1.2 Charged Weak Interactions

In the Standard Model, all interactions are required to conserve quark flavour and mass generation except for charged weak interactions mediated by  $W^+$  and  $W^-$  bosons. In these interactions, so-called weak eigenstates are created in the form of superpositions of down-type quarks ( $d, s, b$ ), also known as mixing. The choice of down-type mixing rather than up-type ( $u, c, t$ ) is purely a convention and the two are mathematically equivalent [1].

The mixed state between  $d$  and  $s$  quarks in a four-quark system is expressed as a rotation performed on the unmixed state, shown in equation 1.

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \quad (1)$$

where  $\theta_c$  is known as the Cabbibo angle, and has a value of approximately 0.23 radians [1].

After the top and bottom quarks were discovered, this procedure was extended via the Cabbibo-Kobayashi-Maskawa (CKM) matrix shown in equation 2, where the various trigonometric terms have been recast as  $V_{ij}$ , where  $|V_{ij}|^2$  is proportional to the amplitude of an  $i - j$  quark transition.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2)$$

The weak eigenstates couple to the unmixed up-type quarks to form three doublet states ( $u, d'$ ), ( $c, s'$ ) and ( $t, b'$ ), known as weak interaction generations.

It is these weak interaction generations that are conserved by charged weak interactions, while quark flavour and mass generation are not. The off diagonal elements of the CKM matrix have been measured to be much smaller than the diagonal elements [2], meaning that such transitions between different quark mass generations are highly suppressed in the SM. The low values of  $V_{ub}$  and  $V_{cb}$  mean that  $b$  hadrons tend to have much longer lifetimes than other heavy flavour hadrons and are therefore reasonably easy to study in large numbers at dedicated detectors like LHCb.

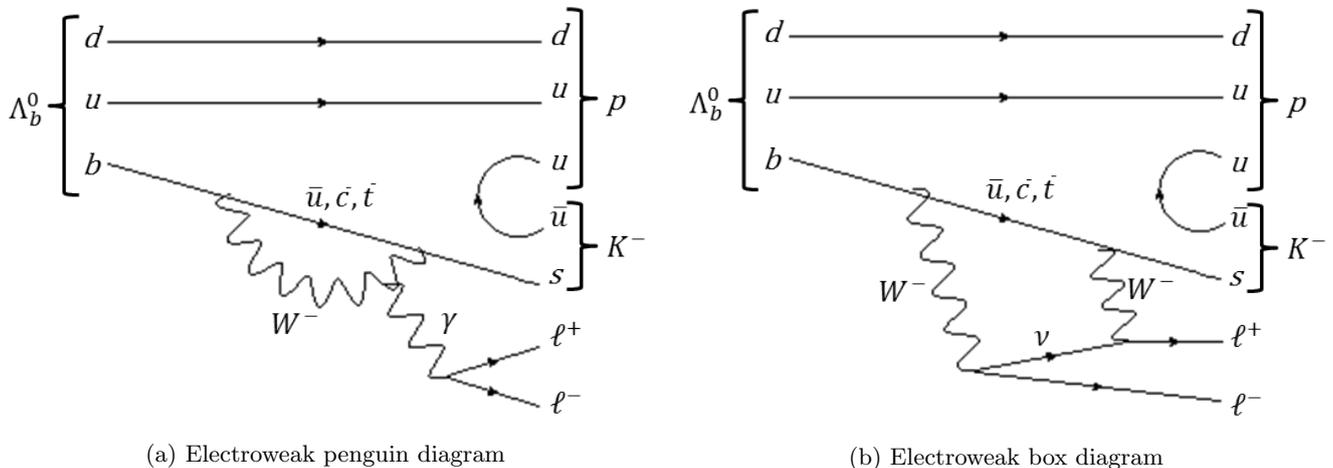
## 1.3 The $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ Decay

The main goal of this analysis is to observe and measure the branching fractions of the rare decays  $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ , where  $\ell = \mu, e$ . This decay proceeds via the flavour changing neutral current (FCNC) process  $b \rightarrow s\ell^+\ell^-$ , which, like all FCNC processes, changes the quark flavour but conserves charge. This has the effect of also changing the mass generation, making the decay an example of the charged weak interactions described in section 1.2.

The CKM matrix presented in equation 2 provides no term to account for down-type to down-type or up-type to up-type transitions, meaning that FCNCs can only proceed at one-loop level in the SM and are therefore highly suppressed. The dominant one-loop processes for  $b \rightarrow s\ell\ell$  transitions are the electroweak penguin and box diagrams

shown in figure 1. These processes are highly sensitive to possible new physics mediated by heavy virtual particles entering into the loops [3].

Figure 1: The dominant one-loop processes that facilitate  $b \rightarrow s\ell\ell$  transitions.



## 1.4 Effective Hamiltonian and New Physics

This section provides a brief mathematical framework for the  $\Lambda_b^0$  decay introduced in section 1.3, and gives a brief overview of the ways in which new physics can emerge.

The effective Hamiltonian of the  $b \rightarrow s\ell\ell$  transition is expressed in terms of operators  $O_i$  and Wilson coefficients  $C_i$ , as shown in equation 3.

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{i=1}^{10} [C_i(\mu) O_i(\mu) + C'_i(\mu) O'_i(\mu)] \quad (3)$$

Here,  $\mu$  is the energy scale of the interaction, defined for  $b \rightarrow s\ell\ell$  as the  $b$  quark mass,  $\mu = m_b$ . Primed terms represent right-handed currents, which are highly suppressed in the SM. The processes illustrated in figure 1 are dominated by the operators and coefficients for which  $i = 7, 9, 10$  [4]. The operators  $O_{7,9,10}$  are defined in equation 4 according to Ali et al. [5].

$$\begin{aligned} O_7 &= \frac{e}{g_s^2} m_b (\bar{s}_L \sigma^{\mu\nu} b_R) F_{\mu\nu} \\ O_9 &= \frac{e^2}{g_s^2} (\bar{s}_L \gamma_\mu b_L) \Sigma_\ell (\bar{\ell} \gamma^\mu \ell) \\ O_{10} &= \frac{e^2}{g_s^2} (\bar{s}_L \gamma_\mu b_L) \Sigma_\ell (\bar{\ell} \gamma^\mu \gamma_5 \ell) \end{aligned} \quad (4)$$

### 1.4.1 Extended Higgs Interactions

The SM also allows for the exchange of a neutral Higgs boson, adding contributions to the effective Hamiltonian from scalar and pseudo-scalar operators  $O_S$  and  $O_P$  with Wilson coefficients displayed in equation 5 [6].

$$\begin{aligned} C_S &= \frac{e^2}{16\pi^2} (\bar{s}P_R b)(\bar{\ell}\ell) \\ C_P &= \frac{e^2}{16\pi^2} (\bar{s}P_R b)(\bar{\ell}\gamma_5\ell) \end{aligned} \tag{5}$$

In equation 5,  $P_R = (1 + \gamma_5)/2$  is the right-handed projection operator. In the SM, these coefficients scale as  $C_{S,P} \sim m_\ell m_b / m_W^2$  and are therefore very small [6], but new physics models that involve an extended Higgs sector, for example the Two Higgs Doublet Model (2HDM) [7], can add extra components such that  $C_S$  and  $C_P$  are no longer negligible.

### 1.4.2 Supersymmetric Effects

The other main avenue for new physics to enter into the  $b \rightarrow s\ell\ell$  transition is described by supersymmetry (SUSY). This is an umbrella term for a variety of extensions to the SM, all of which predict that every SM fermion and boson has associated with it a much heavier particle of the opposing classification, known as a 'superpartner'.

The simplest SUSY model, and the one discussed here, is known as the Minimal Supersymmetric Standard Model (MSSM), so called because it contains the minimal number of new particles and interactions required to be consistent with phenomenology [8]. The Higgs sector of this model contains the same particles and tree-level Yukawa couplings as the 2HDM described above [9].

In the MSSM, new sources of flavour violation are introduced due to supersymmetric particles participating in the loops [10] (for example a gluino-mediated FCNC [11]). The effect of this is to cause the right-handed currents in equation 3 to become significant, leading to a change in the effective Hamiltonian of the  $b \rightarrow s\ell\ell$  transition. Flavour violation would also be the mechanism by which lepton universality is violated as discussed in the next section.

## 2 Motivation for Measuring $R_{pK}$

This section discusses the relevance of this particular measurement to the possibilities for new physics described above.

### 2.1 Lepton Universality

A stipulation of the Standard Model is that the weak vector bosons couple equally to all lepton flavours, in a concept known as lepton universality. The implication is that weak interactions do not distinguish between different leptons and a large sample of interactions should produce each flavour in equal amounts. Physics beyond the SM may contain new interactions, such as the those described in sections 1.4.1 and 1.4.2, that do not couple equally to all lepton flavours and therefore violate lepton universality.

### 2.2 Testing Lepton Universality Using $R_{pK}$

If all physics respects lepton universality, we expect to find equal amounts of electrons and muons produced by the decay  $\Lambda_b^0 \rightarrow pK^{-\ell^+\ell^-}$  (where  $\ell = \mu, e$ ), such that the branching fractions for  $\ell = \mu$  and  $\ell = e$  are equal and we have [6]

$$R_{pK} = \frac{\mathcal{B}(\Lambda_b^0 \rightarrow pK^{-\mu^+\mu^-})}{\mathcal{B}(\Lambda_b^0 \rightarrow pK^{-e^+e^-})} = 1 \pm \mathcal{O}(10^{-3}) \tag{6}$$

Since the SM requires lepton universality to hold, any significant deviation of this ratio from unity would suggest the presence of new physics, making its measurement a useful probe of extensions to the SM.

### 2.3 Previous Lepton Universality Measurements

The value of  $R_{pK}$  in the search for new physics can easily be seen by looking at previous tests of lepton universality performed using other rare decays. The most notable of these is the measurement of  $R_K$ , defined in equation 7.

$$R_K = \frac{\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-)}{\mathcal{B}(B^+ \rightarrow K^+ e^+ e^-)} \quad (7)$$

Three experiments have measured this ratio to date: Belle in 2009 [12], BaBar in 2012 [13] and LHCb in 2014 [14]. A summary of their results is shown in figure 2, and demonstrates clear tension between recent measurements and the SM prediction. The most recent measurement, from LHCb, of  $R_K = 0.745^{+0.090}_{-0.074}$  (stat)  $\pm 0.036$  (syst) [14], lies 2.6 standard deviations away from the SM prediction. It is therefore of great interest to investigate these findings further by measuring the equivalent ratio in a new  $b \rightarrow s\ell\ell$  process.

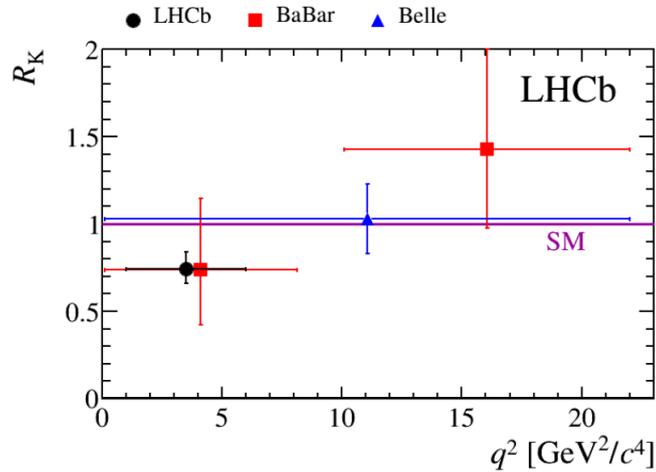


Figure 2: Plot showing the values of  $R_K$  obtained by Belle, BaBar and LHCb. The purple line indicates the SM prediction, showing the tension between measurement and theory.

A measurement of

$$R_{K^*} = \frac{\mathcal{B}(B \rightarrow K^{(*)} \mu^+ \mu^-)}{\mathcal{B}(B \rightarrow K^{(*)} e^+ e^-)}$$

is also currently underway at LHCb, and the ratio

$$R(D^*) = \frac{\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau)}{\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)}$$

which involves a different current ( $b \rightarrow c\ell\ell$ ), was recently found to be 2.1 standard deviations away from the prediction consistent with lepton universality [15].

## 3 The LHCb Detector

LHCb is a single-arm forward spectrometer designed to probe the effects of new physics through precision measurements of decays involving  $b$  and  $c$  quarks. The detector has an acceptance covering the pseudorapidity range

$2 < \eta < 5$ , where pseudorapidity is defined in equation 8.

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \quad (8)$$

In the above equation,  $\theta$  is the angle between the particle's trajectory and the beam axis. A justification for this narrow acceptance range can be seen in figure 3, which shows the forward and backward two-dimensional rapidities of  $b$  and  $\bar{b}$  hadrons produced in  $pp$  collisions at the LHC's design energy of  $\sqrt{s} = 14$  TeV. LHCb consists of six different sub-detector systems: the VERtex LOcator (VELO), Ring Imaging CHERenkov (RICH) detectors, the dipole magnet, calorimeter systems, tracking stations and muon chambers. These will be described in more detail in the following sections. A schematic of the detector is shown in figure 4.

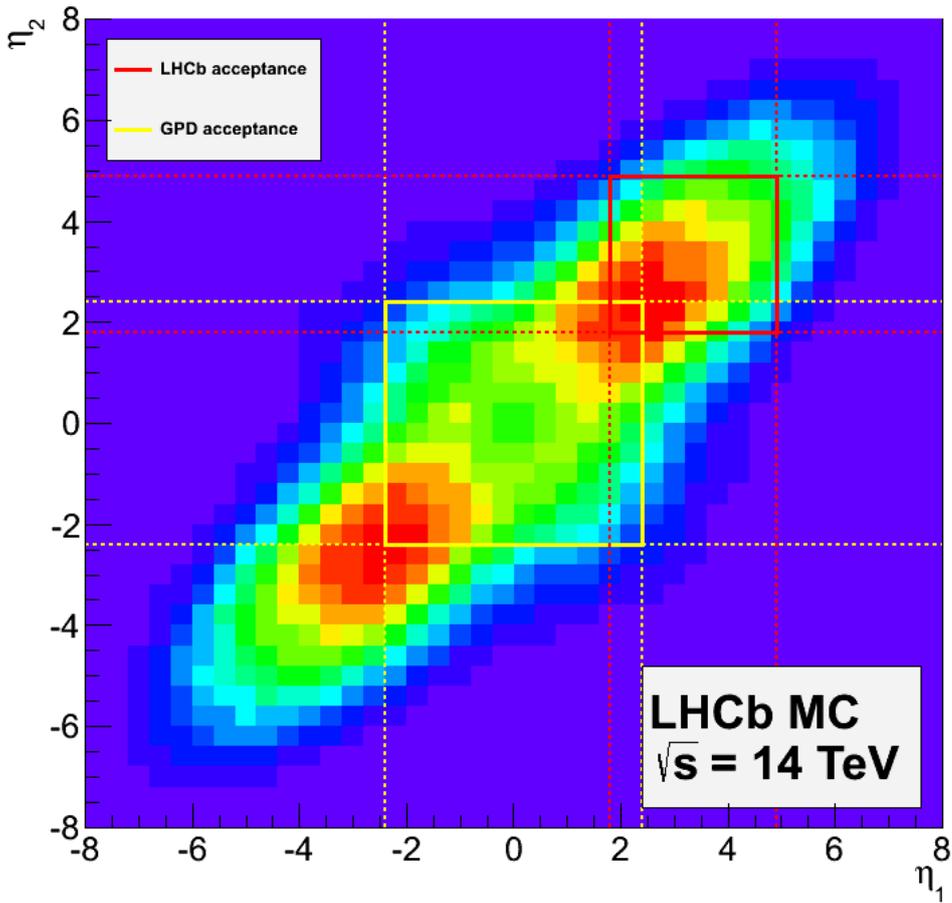


Figure 3: Plot showing the forward and backward two-dimensional rapidities of  $b$  and  $\bar{b}$  hadrons produced in LHCb at the design energy of  $\sqrt{s} = 14$  TeV. The red and yellow squares indicate the acceptance of LHCb vs. that of a general purpose detector (GPD), showing how LHCb's design lends itself to  $b$  hadron physics.

LHCb took data during Run I at a maximum luminosity of  $4 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$  [16]. It collected  $1 \text{ fb}^{-1}$  in 2011 and  $2 \text{ fb}^{-1}$  in 2012. The detector is designed to run at a much lower luminosity than that of the wider LHC because,

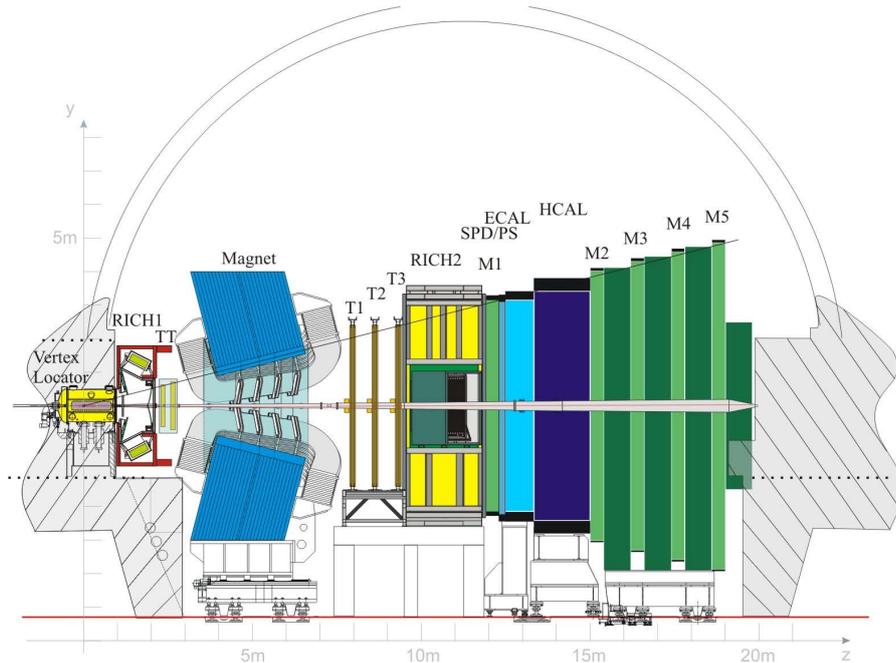


Figure 4: Schematic of the LHCb detector

as well as reducing radiation damage to the components, this reduces the interaction rate to just one per  $pp$  bunch crossing [17], allowing for optimal reconstruction to be used in precision measurements. The reduction in luminosity inside LHCb is achieved by slightly offsetting the beams at the interaction point.

### 3.1 Vertex Locator (VELO)

The VELO is formed from a set of silicon strip detectors that operate in a cylindrical formation along the beam axis around the  $pp$  interaction point, at a distance of 7mm [18]. This is smaller than the minimum beam aperture required during the injection phase, so to prevent radiation damage the VELO is constructed in two semi-circular halves that can be retracted until the beams are stable.

The purpose of the VELO is to make precision measurements of charged particle tracks around the interaction point, and to use these tracks to identify the coordinates of the primary and secondary vertices. The primary vertex refers to the production vertex of a  $b$  hadron (i.e. the  $pp$  collision point), while secondary vertex refers to the decay vertex of a  $b$  hadron.

Due to its unusually long lifetime, a  $b$  hadron's flight distance before decaying is on the order of millimetres, meaning that the primary and secondary vertices are clearly separated in the detector. The impact parameter is defined as the shortest distance between a particle's extrapolated track and the primary vertex, and it is an extremely useful variable for distinguishing between  $b$  hadrons, produced at the primary vertex, and their decay products, produced at the secondary vertex.

### 3.2 Dipole Magnet

The dipole magnet at LHCb is used for measuring the momenta of charged particles. The momentum is also an important variable for  $b$  physics as the decay products of  $b$  quarks tend to be much lighter than the quark itself and

therefore have higher transverse momentum relative to the  $b$  flight direction. The magnet produces an integrated magnetic field of 4 Tm across the magnet region [17], which deflects charged particles as they pass through. Observing the magnitude of the deflection enables a particle's momentum to be determined (see also section 3.4).

Since the magnet deflects positively and negatively charged particles in opposite directions in the  $x - z$  plane, a difference in performance between the left and right sides of the detector could lead to bias. To avoid this, the polarity of the magnet is reversed from time to time. This happened roughly twice per month during Run I [16].

Inside the magnet region itself, there is a very low density of material [17]. This feature ensures that Bremsstrahlung photons produced by electrons in or before the magnet region do not experience refractive effects which would hamper reconstruction. Bremsstrahlung radiation will be discussed further in section 4.3.

### 3.3 Ring Imaging Cherenkov (RICH) Detectors

There are two RICH detector systems at LHCb, namely RICH1 and RICH2, located before and after the magnet respectively. Their function is to identify charged hadrons from the Cherenkov radiation that the particles produce while passing through a radiator medium. RICH1 is designed to identify low-momentum particles (2-40 GeV/c), while RICH2 focuses on higher-momentum particles (15-100 GeV/c) [19].

Cherenkov radiation is produced when a particle travels through a dielectric medium faster than the phase velocity of photons in that medium. The radiation takes the form of a cone of light centred on the particle's trajectory, with an opening angle  $\theta_c$ , known as the Cherenkov angle. This angle is related to the particle's velocity  $v$  by

$$\cos \theta_c = \frac{1}{\eta\beta}$$

where  $\eta$  is the refractive index of the medium and  $\beta = v/c$ . Combining the value of  $\beta$  with the momentum of the particle gives a determination of the mass and therefore the identity of the particle.

RICH1 uses aerogel and  $C_4F_{10}$  as radiator materials, while RICH2 uses  $CF_4$  [19]. Figure 5 shows the relationship between  $\theta_c$  and momentum for particles in the different radiator materials, and illustrates why the RICH detectors are not reliable discriminators between electrons and muons in the momentum acceptance range. The small masses of  $e$  and  $\mu$  mean that  $\beta \sim 1$  and the Cherenkov angle is at its maximum for both particles – it is said to be saturated.

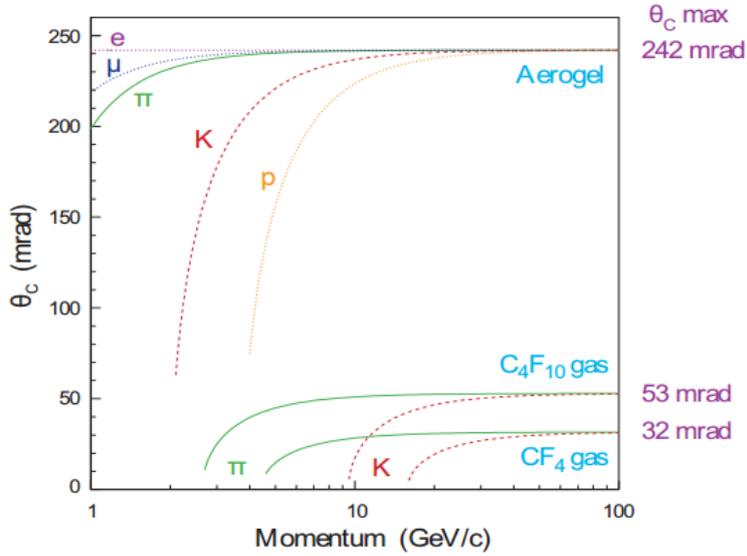


Figure 5: The relationship between the Cherenkov angle  $\theta_c$  and the momenta of particles in the RICH detectors.

### 3.4 Tracking Stations

In addition to the VELO, LHCb has four other tracking stations: the Tracker Turicensis (TT), located before the magnet, and the tracking stations T1-T3 located after the magnet. They work in tandem with the VELO to measure the tracks of charged particles. The TT uses silicon strip detectors similar to those used in the VELO, while T1-T3 use silicon strips close to the beam pipe and straw-tubes in the outer region [17]. Track measurements from the VELO and TT are combined with those from T1-T3 in order to determine the curvature of particles in the magnet region, thereby allowing a particle's momentum and charge to be calculated.

### 3.5 Calorimeter System

The calorimeter system at LHCb consists of four separate sub-systems, namely the Scintillating Pad Detector (SPD), Pre-Shower Detector (PS), Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). Their function is to identify and measure the energy of electrons, photons and hadrons. An absorber layer of lead is placed between the SPD and PS and is also interwoven between the scintillating tiles (known as cells) of the ECAL. The HCAL uses an iron absorber [17]. The interaction of a particle with the absorber layers produces a shower of charged particles that cause flashes of light in the scintillating tiles. The intensity of these flashes is used to determine the energy of the incident particle [20]. The longitudinal and transverse segmentation of the calorimeter system provides good discriminating power between hadrons and electrons/photons as hadronic showers tend to be longer and broader than electromagnetic ones [20].

### 3.6 Muon Chambers

The most downstream sub-detectors in LHCb are the muon chambers M1-M5. Their function is to identify and measure the momenta of muons, which interact minimally with the preceding parts of the detector. M1 is located before the calorimeter system and is designed to provide transverse momentum measurements, while M2-M5 lie after the HCAL and are separated by 80cm-thick iron absorbers in order to select only penetrating muons [17].

Those muons that reach all the way to M5, which requires a momentum of approximately 6 GeV/c [17], are assigned the `IsMuonLoose` classification. Effective muon identification is one of the fundamental requirements of LHCb, and the muon detector system has been designed such that a binary classification `IsMuon` can be assigned to a particle with an efficiency above 97% [21].

### 3.7 Trigger System

The trigger system at LHCb is divided into two sub-systems: Level 0 (L0) and High level trigger (Hlt). The L0 triggers are hardware based and run synchronously with the detector at the 40 MHz bunch crossing rate [16]. The function of L0 is to reduce the data collection rate to below 1 MHz such that the entire detector readout can be passed to the Hlt for more detailed selection [16].

#### 3.7.1 L0 Trigger

The L0 trigger is divided into 3 parts: the calorimeter trigger, the muon trigger and the pile-up trigger. The last of these is used only in determining the luminosity and not for selecting events. The L0-Calorimeter trigger uses information from the calorimeter systems described in section 3.5 and selects hadron, electron and photon candidates with the highest transverse energy,  $E_T$ . If an event contains at least one candidate that exceeds a given threshold the event is stored [22].

The L0-Muon trigger looks for the two muon tracks with the largest and second largest transverse momentum,  $p_T$ . The tracks must also define a straight line through each of the five muon stations and point towards the interaction point in the vertical plane. The muon trigger places a threshold on either the largest  $p_T$  or the product of the largest and second largest [22]. The thresholds for both L0-Muon and L0-Calorimeter are shown in table 3.

Table 3: The L0 trigger thresholds used during Run I [16]

	$p_T$ or $E_T$	
	2011	2012
single muon	1.48 GeV/c	1.76 GeV/c
dimuon $p_{T_1} \times p_{T_2}$	(1.30 GeV/c) <sup>2</sup>	(1.60 GeV/c) <sup>2</sup>
hadron	3.50 GeV	3.70 GeV
electron	2.50 GeV	3.00 GeV
photon	2.50 GeV	3.00 GeV

#### 3.7.2 High level trigger

The Hlt is a software programme that runs on a farm of multi-processor PCs known as the Electronic Filter Farm (EFF) [22]. The reduction in input rate achieved by L0 means that more time is available for each evaluation. The Hlt is split into two stages; Hlt1 performs a partial reconstruction of each track and places thresholds on various kinematic variables to reduce the event rate to 43 kHz [22], which is low enough for a more complete kinematic analysis to be performed at the Hlt2 stage. Hlt2 then performs selections based on invariant mass or the direction of the  $b$  hadron momentum, required to be oriented towards the primary vertex [17]. This reduces the rate to 3 kHz [22], at which point the data are stored for use in analysis.

## 4 Analysis Strategy

This section describes in detail the strategy for performing the measurement of  $R_{pK}$ .

## 4.1 Data

The final analysis to measure  $R_{pK}$  will use the full  $3 \text{ fb}^{-1}$  of data collected by LHCb in 2011 and 2012. The intermediate analysis performed on the  $J/\psi$  control channel (see section 4.2) uses smaller samples of electron and muon data to reduce processing time.

Prior to analysis, the data are screened by the trigger system as described in section 3.7. A process known as stripping is then carried out in which a series of loose selection criteria, known as a stripping line, is imposed in order to reduce the sample to a more manageable size while keeping a high signal yield. The stripping lines used so far in this analysis are actually designed to select  $B \rightarrow K^+ \ell^+ \ell^-$  decays, but as the selections are very loose and the  $B$  and  $\Lambda_b^0$  masses are similar, it is possible to reuse them here for the initial development of the selection. However, decays involving  $B$  mesons are a considerable source of background in this analysis, so new bespoke stripping lines will be adopted as soon as they become available. The selection criteria for the muon and electron lines are presented in table 4.

Table 4: The stripping selections used in this analysis.

Particle	Requirement
B	Flight Distance $\chi^2 > 100$ DIRA $> 0.9995$ Vertex $\chi^2 < 16$
Dilepton	Flight Distance $\chi^2 > 16$ Impact Parameter $\chi^2 > 9$
Lepton	Impact Parameter $\chi^2 > 16$ $p_T > 800 \text{ MeV}/c$
Kaon	Impact Parameter $\chi^2 > 16$ $p_T > 800 \text{ MeV}/c$
Muon (Muon line only)	<b>IsMuonLoose</b>
Electron (Electron line only)	$\text{DLL}_{e\pi} > -2.0$

In table 4, the DIRA variable is defined as the cosine of the angle between the  $\Lambda_b^0$  flight path and the resultant momentum vector of the final state particles. This angle should be small, hence the large value of DIRA specified here. The  $\chi^2$  quantity is a statistical test of the fit quality of a variable. DLL will be defined in section 5.1.

Smaller samples of simulated Monte Carlo data will also be used for modelling background effects and for calculating the efficiency of the selection process, as well as for training the BDT (see section 6).

## 4.2 The $J/\psi$ Control Channel

In order to calculate  $R_{pK}$  using the  $\Lambda_b^0 \rightarrow pK^- \ell^+ \ell^-$  rare decay, this analysis first makes use of the control channel  $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \ell^+ \ell^-)$  (where  $\ell = \mu, e$ ) in order to determine the relative, rather than absolute, branching fractions of the rare modes as illustrated in equation 9.

$$\frac{\mathcal{B}(\Lambda_b^0 \rightarrow pK \ell^+ \ell^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow pK J/\psi (\rightarrow \ell^+ \ell^-))} = \frac{N_{\Lambda_b^0 \rightarrow pK \ell^+ \ell^-}}{N_{\Lambda_b^0 \rightarrow pK J/\psi (\rightarrow \ell^+ \ell^-)}} \times \frac{\epsilon_{\Lambda_b^0 \rightarrow pK J/\psi (\rightarrow \ell^+ \ell^-)}}{\epsilon_{\Lambda_b^0 \rightarrow pK \ell^+ \ell^-}} \times \mathcal{B}(J/\psi \rightarrow \ell^+ \ell^-) \quad (9)$$

Here,  $N$  represents the number of decays appearing in the data and  $\epsilon$  is the selection efficiency. Dividing two of these equations (one for each of  $\ell = \mu, e$ ) gives the equation for finding  $R_{pK}$ .

$$R_{pK} = \frac{\frac{N_{\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-}}{N_{\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)}} \times \frac{\epsilon_{\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)}}{\epsilon_{\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-}}}{\frac{N_{\Lambda_b^0 \rightarrow pK^- e^+ e^-}}{N_{\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow e^+ e^-)}} \times \frac{\epsilon_{\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow e^+ e^-)}}{\epsilon_{\Lambda_b^0 \rightarrow pK^- e^+ e^-}}} \times \frac{\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)}{\mathcal{B}(J/\psi \rightarrow e^+ e^-)} \quad (10)$$

The values of  $\mathcal{B}(J/\psi \rightarrow \ell^+ \ell^-)$  will be obtained from the Particle Data Group (PDG) [2].

The advantages of using the control channel in this way are fourfold. Firstly, the control channel and rare mode have the same final state, meaning the same selection can be used for both with the exception of a criterion to distinguish between rare and control events. This is useful because the selection can be developed more easily with the control channel as it proceeds at tree-level in the SM and therefore contains many more events.

Secondly, the control channel provides an easy means of testing the selection as the branching fraction of  $\Lambda_b^0 \rightarrow pK^- J/\psi$  has been measured recently [23], and by combining it with the values of  $\mathcal{B}(J/\psi \rightarrow \ell^+ \ell^-)$  we expect

$$\frac{\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-))}{\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow e^+ e^-))} = 1$$

Thirdly, the common final state between the control channel and the rare modes means that when forming the double ratio in equation 10, the efficiencies and systematic uncertainties will tend to cancel out, leading to a much more precise measurement of  $R_{pK}$ .

Finally, fit models can be developed more easily using the control channel due to the larger number of events, and the shared final state means they can be readily scaled to fit the rare modes.

### 4.3 Bremsstrahlung Radiation

It is necessary at this point to discuss the complications arising in electron physics due to the effects of Bremsstrahlung radiation, as well as the correction procedure employed to reduce the associated errors.

All accelerating charged particles radiate Bremsstrahlung photons. This causes a problem inside particle detectors because it means that the momentum and energy of electrons, and therefore the reconstructed masses of parent particles, will be significantly underestimated due to kinetic energy losses.

The power lost in a vacuum by an accelerating particle of mass  $m$ , charge  $q$ , and initial energy and velocity  $E$  and  $\vec{v}$  respectively is given by equation 11, in which  $\vec{\beta} = \vec{v}/c$ .

$$P = \frac{q^2 E^4}{6\pi\epsilon_0 m^4 c^9} \left( \dot{\beta}^2 + \frac{(\vec{\beta} \cdot \dot{\vec{\beta}})^2}{1 - \beta^2} \right) \quad (11)$$

Muons are around 200 times heavier than electrons, so it can be clearly seen in equation 11 why Bremsstrahlung losses are neglected in the muon channel.

#### 4.3.1 Energy Recovery Procedure

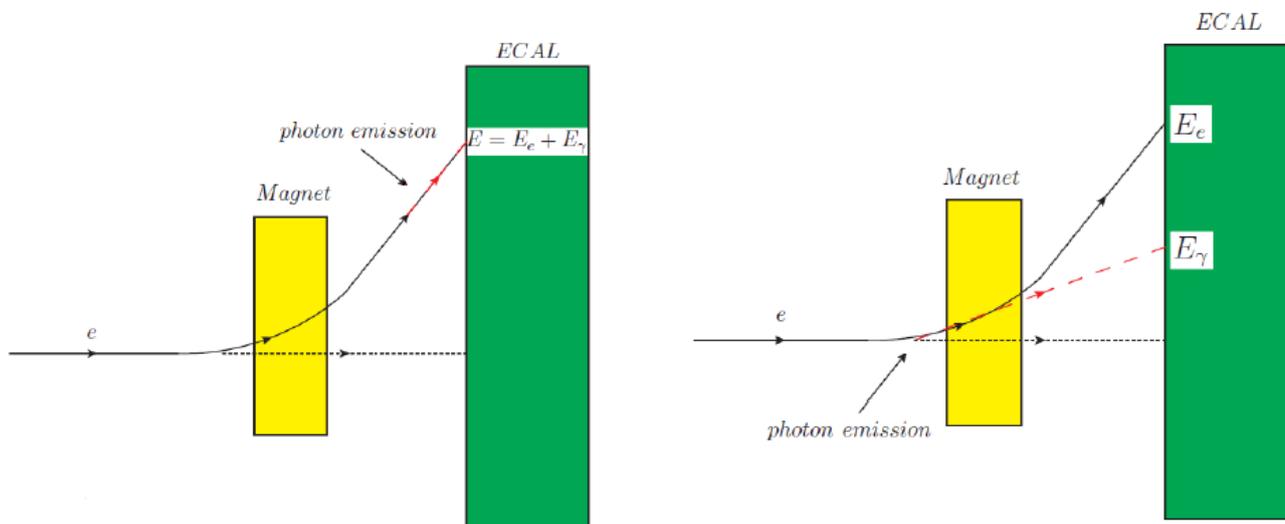
As an electron traverses the detector, there is a probability that it will emit Bremsstrahlung radiation. The vast majority of Bremsstrahlung photons are emitted collinear to the particle track, giving two distinct scenarios which

are discussed below and shown in figure 6.

If the electron emits a Bremsstrahlung photon after the magnet region, both particles will continue along the same trajectory and deposit their energy in the same ECAL cell, leading to no loss of measured energy. The curvature of the electron's track in the magnet region is unaffected by the emission, hence the measurement of its initial momentum is also correct. In this scenario, no special recovery procedure is necessary.

The electron may also emit a photon either before or in the magnet region, both of which give rise to the same effect. The photon continues along its initial trajectory while the electron track is curved by the magnetic field, and the particles deposit their energy in different ECAL cells meaning some of the electron's initial energy is lost on detection. The measurement of the initial momentum of the electron is also incorrect since the magnitude of its deflection in the magnet region is affected by the loss of kinetic energy.

Figure 6: The two scenarios for Bremsstrahlung emission in LHCb



(a) Photon emitted after the magnet region

(b) Photon emitted before the magnet region. Apart from the position of the photon's production vertex, this setup would be the same for a photon originating inside the magnet region.

The correction procedure applied in the second scenario described above is as follows. Electron tracks are extrapolated linearly from the production vertex to the ECAL (black dashed line in figure 6b). If a photon is detected in the ECAL between this extrapolation and the actual electron hit (red dashed line in figure 6b) and lies in the same horizontal plane, its energy is added to that of the electron.

The Bremsstrahlung correction procedure described here is an essential tool for improving accuracy in the electron channel, but it is still not perfect. Photons can be lost or associated with the wrong electrons, and other particles such as pions can be mis-identified as both photons and electrons. These factors lead to under- and over-corrections to the electron energy and subsequent reconstructed variables, causing a substantial drop in resolution compared to the muon channel.

### 4.3.2 HOP

Related to the above process is a procedure known as HOP, which can also be a useful tool for recovering lost electron energy. Its principles are described here according to Borsato et al. [24], but the technique has not been previously investigated with  $\Lambda_b^0$  decays and the process of incorporating it into this analysis is still ongoing.

During a  $\Lambda_b^0 \rightarrow pK^-e^+e^-$  decay (which may or may not proceed via a  $J/\psi$  resonance), the proton and kaon are actually produced via a  $\Lambda^*(1520)$  resonance. This particle is highly off-mass-shell and only exists for an extremely short time ( $\Gamma = 15.6$  MeV [2]) before decaying into  $pK^-$ .

When a  $\Lambda_b^0$  decays, the products  $\Lambda^*$  and  $e^+e^-$  (which may or may not come via a  $J/\psi$  decay) are observed to have equal and opposite momenta transverse to the initial  $\Lambda_b^0$  flight direction, i.e.  $\vec{P}_t(\Lambda^*) = -\vec{P}_t(ee)$ . Note here that  $P_t$  refers to momentum transverse to the  $\Lambda_b^0$  flight direction and not to momentum transverse to the beam axis as is the case elsewhere. If the electrons go on to lose energy through Bremsstrahlung radiation, their measured momentum transverse to the  $\Lambda_b^0$  flight direction will no longer equal that of  $\Lambda^*$ , and the ratio

$$\alpha_{\text{HOP}} = \frac{|\vec{P}_t(\Lambda^*)|}{|\vec{P}_t(ee)|}$$

will be greater than one. Over-corrections made by the procedure described in section 4.3.1 cause the ratio to be less than one.

This ratio provides a straightforward correction to the overall momentum of the dielectron:  $\vec{P}^{\text{corr}}(ee) = \alpha_{\text{HOP}} \times \vec{P}(ee)$ . The corrected value of the dielectron momentum can then be used to reconstruct a corrected value of the  $\Lambda_b^0$  mass. Implicit in this correction is the assumption that an electron emits Bremsstrahlung photons collinear to its flight path and therefore maintains a constant direction.

The effectiveness of the HOP tool as a Bremsstrahlung correction procedure is not comparable to that of the procedure described in section 4.3.1. Its resolution is degraded by the quality of the reconstruction of the  $\Lambda_b^0$  direction of flight, which may not always be reliable. The resolution is also affected by that of the angle between the dielectron momentum vector and the  $\Lambda_b^0$  direction of flight, which is required for a trigonometric calculation of  $P_t(ee)$ .

The HOP tool does, however, provide discriminating power between signal and both partially reconstructed and combinatorial background. Partially reconstructed background arises when one or more of the final state particles used to reconstruct the  $\Lambda_b^0$  mass is either mis-reconstructed from a different source, or not reconstructed at all, causing the reconstructed mass to be shifted to lower values. Combinatorial background occurs when final state particles originating from different  $\Lambda_b^0$ 's are mistakenly reconstructed together, producing a fake  $\Lambda_b^0$  signature. In both these cases, the value of  $\alpha_{\text{HOP}}$  will differ significantly from unity, and a veto of such values allows background to be removed.

## 4.4 Dilepton Invariant Mass ( $q^2$ )

The rare mode analysis will be carried out in the dilepton invariant mass squared region  $1 < q^2 < 6$  GeV/c<sup>2</sup>. This region is theoretically favoured as it avoids the  $J/\psi$  mass and is therefore free from  $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \ell^+ \ell^-)$  background. Analysis conducted on the control channel of course has a different  $q^2$  requirement directed at the  $J/\psi$  region. Above  $q^2 = 6$  GeV/c<sup>2</sup>, the presence of  $c\bar{c}$  loops, known as charmonium ( $J/\psi$  and  $\psi(2S)$ ) resonances, becomes a considerable hindrance, while below  $q^2 = 1$  GeV/c<sup>2</sup> the same happens with  $u\bar{u}$ ,  $d\bar{d}$  and  $s\bar{s}$  loops [18].

Additionally, at low  $q^2$  there can be considerable pollution from  $\gamma \rightarrow e^+e^-$  conversions which cannot necessarily be distinguished from signal decays. The theory also diverges as  $q^2 \rightarrow 0$ , exhibiting what is known as a photon pole. Figure 7 shows the favoured region, photon pole and the charmonium resonances. The favoured  $q^2$  region is

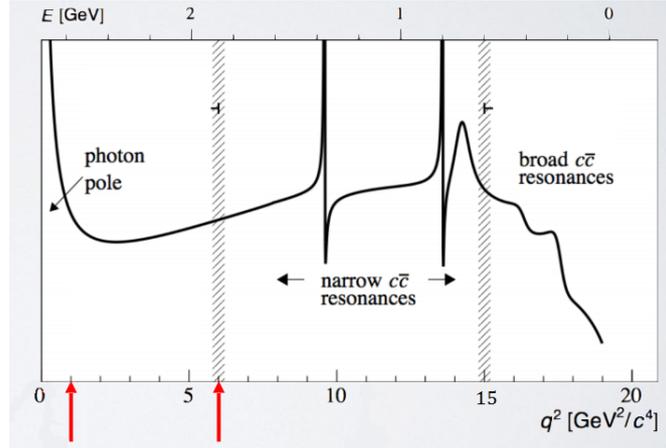
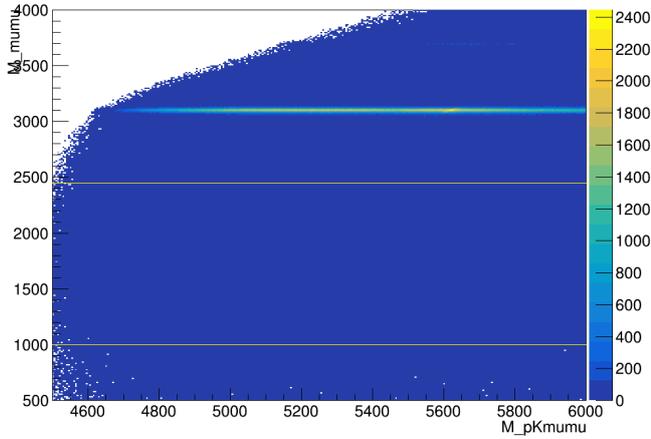


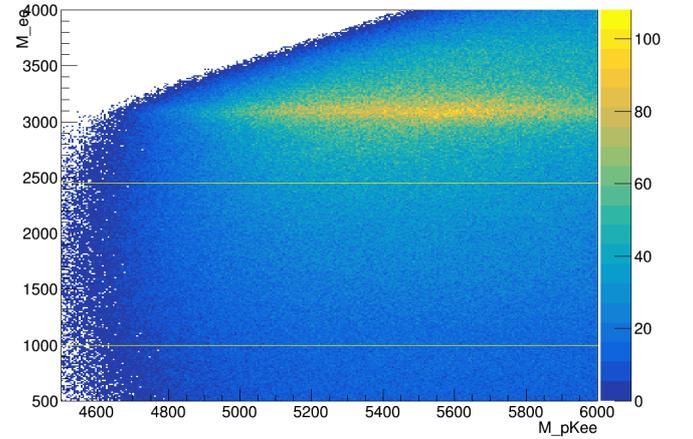
Figure 7: Plot showing the behaviour of the theory with  $q^2$  [25]. The favoured  $q^2$  region is shown by the red arrows, with the photon pole divergence and  $c\bar{c}$  resonances clearly visible outside this area.

marked on data in figure 8, which shows the reconstructed  $\Lambda_b^0$  mass plotted against  $q$  for the muon and electron channels.

Figure 8: Reconstructed  $\Lambda_b^0$  mass plotted against  $q$  in the muon and electron channels. The favoured  $q^2$  region is indicated by the yellow lines.



(a) Plot for the muon channel, showing a well defined band corresponding to the  $J/\psi$  resonance, with the warmest point lying at the  $\Lambda_b^0$  mass. The favoured region for the rare mode is placed well below this. The  $\psi(2S)$  resonance is also faintly visible towards the top right of the plot.



(b) Plot for the electron channel, in which the  $J/\psi$  band is clear but the resolution is noticeably impaired by the Bremsstrahlung correction procedure described in section 4.3.1.

## 5 Event Selection

This section describes the procedure for selecting signal events while rejecting events not useful to the analysis. The selection is split into three main stages: particle identification (PID), rejection of peaking backgrounds, and the use of a multivariate classifier. As discussed in section 4.2, the selection is developed using  $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \ell^+\ell^-)$  events after which it will be applied as is to the rare mode. The exception to this is of course the  $q^2$  requirement as mentioned in section 4.4.

### 5.1 Particle Identification (PID)

At LHCb, two types of particle identification variable are in use: DLL and ProbNN. DLL stands for Difference in Log Likelihood, and is a measure of how much a particular track in the detector is 'like' a known particle. Information from the RICH detectors, calorimeter system and muon chambers is combined to produce numbers representing the likelihood that a detected particle is each of  $p^\pm, K^\pm, \mu^\pm, e^\pm$ . This likelihood is then normalised by comparing it with the likelihood that the particle is a pion to produce the DLL variable. For example  $\text{DLL}_{K\pi} = \ln L(K) - \ln L(\pi)$  is a measure of how kaon-like a particle is.

Combining the DLL variables with information from the tracking stations into a neural network algorithm produces ProbNN, which is the probability of a detected particle being each of  $p^\pm, K^\pm, \mu^\pm, e^\pm, \pi^\pm$ . For example ProbNNk is the probability that a particle is a kaon. There is also ProbNNghost, which is the probability of what the detector interprets as a particle being a phantom trace. This analysis favours ProbNN over DLL as the addition of tracking information has been shown to provide advantages to PID-based selection [16].

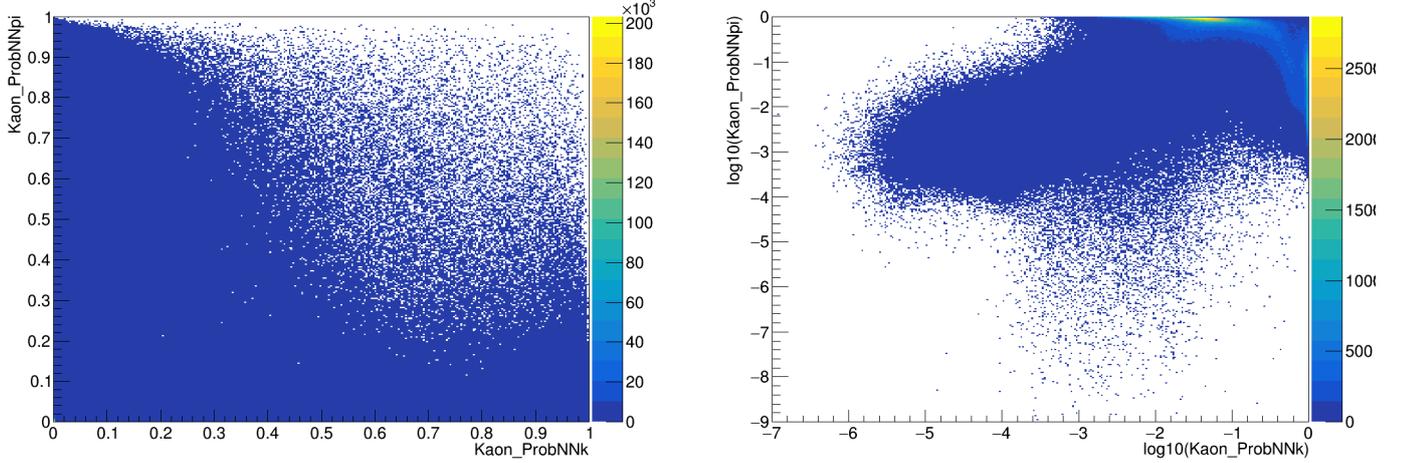
The PID selections imposed on the muon and electron channels are summarised in table 5. No PID selection has been performed on muon tracks due to the already high efficiency of the muon trigger.

Table 5: The PID selection imposed on final state particles in the muon and electron channels.

Particle	$\ell = \mu$	$\ell = e$
$p$	ProbNNp > 0.1 ProbNNpi < 0.7 ProbNNk < 0.8 ProbNNghost < 0.5	log(ProbNNp) > -1.0 log(ProbNNpi) < -0.4 log(ProbNNk) < -0.2 ProbNNghost < 0.5
$K^-$	ProbNNk > 0.2 ProbNNpi < 0.6 ProbNNghost < 0.5	ProbNNk > 0.2 log(ProbNNpi) < -0.4 ProbNNghost < 0.5
$\ell^\pm$	N/A	ProbNNe > 0.4 log(ProbNNpi) < -0.3

It was noticed in the electron channel, for which there are no previous examples of selection in the literature, that using logarithms of PID variables gave a much clearer picture of the ideal selection criteria. This is illustrated in figure 9.

Figure 9: Plots illustrating the advantages of taking logarithms of ProbNN variables when determining appropriate selection criteria in the electron channel.



(a) Plot of  $K\_ProbNNk$  vs.  $K\_ProbNNpi$ , which does not show any kind of discernible detail.

(b) Plot of  $\log(K\_ProbNNk)$  vs.  $\log(K\_ProbNNpi)$ . In this plot, pion traces can be clearly seen in the warm region at the top, while kaon traces are also visible in the thin warm region on the right hand side.

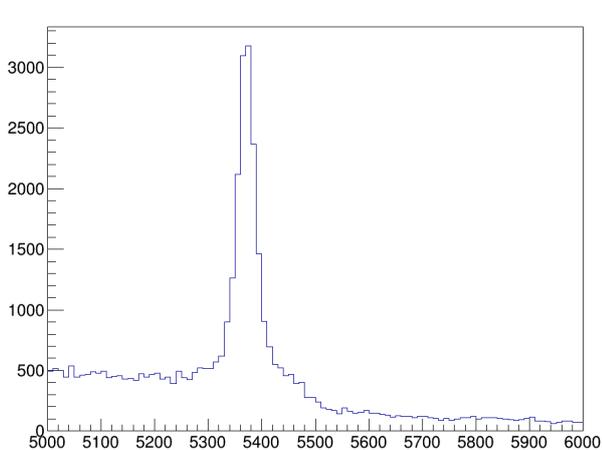
## 5.2 Rejection of Peaking Backgrounds

An observation of a  $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$  decay is made through the detection of final state particles with tracks pointing to the same secondary vertex in the VELO. However, even after PID selection, it is possible for one or more of these particles to be mis-identified, meaning that a final state from a different decay is falsely attributed to a  $\Lambda_b^0$  decay. This leads to peaks in the signal distribution that do not correspond to the  $\Lambda_b^0$  mass, and these peaks can have tails in  $\Lambda_b^0$  mass-space that extend under and therefore contaminate the signal peak. There are two main sources of peaking background in this analysis, whose isolation and removal is described in the following subsections.

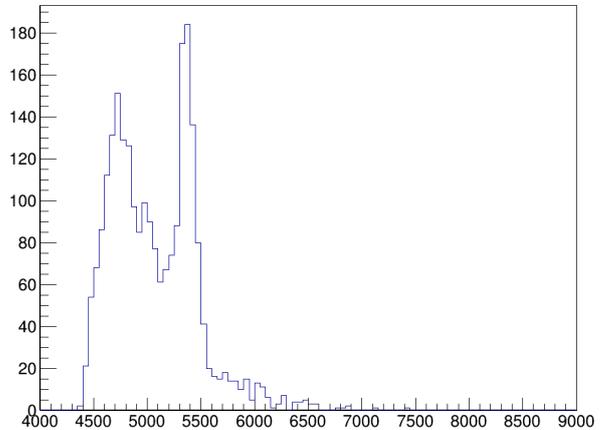
### 5.2.1 $B_s^0 \rightarrow \phi(\rightarrow pK^-)J/\psi(\rightarrow \ell^+\ell^-)$

When a particle is identified by the detector, its signature in the dataset is assigned a ‘mass hypothesis’ corresponding to the mass of its identification. For example, a particle that is identified by the detector as a kaon will be assigned a kaon mass label in the dataset. The particle mis-identification described in the previous paragraph is caused by the incorrect assignment of mass hypotheses. In order to explore the effects of mis-identification, different mass hypotheses can be manually assigned to final state particles before reconstructing and plotting the initial state mass. A peak at the mass of a known non- $\Lambda_b^0$  particle indicates that mis-identification has occurred and a peaking background source is present. Figure 10 shows the reconstructed mass distribution after the proton in the final state has been assigned a kaon mass, such that the hypothesised final state is  $KK\ell\ell$ .

Figure 10: The reconstructed initial state mass after the final state has been hypothesised as  $KK\ell\ell$  in the muon and electron channels, showing peaks at the mass of the  $B_s$  meson,  $m_{B_s} = 5367 \text{ MeV}/c^2$ .



(a) The reconstructed mass of the  $KK\mu\mu$  final state.



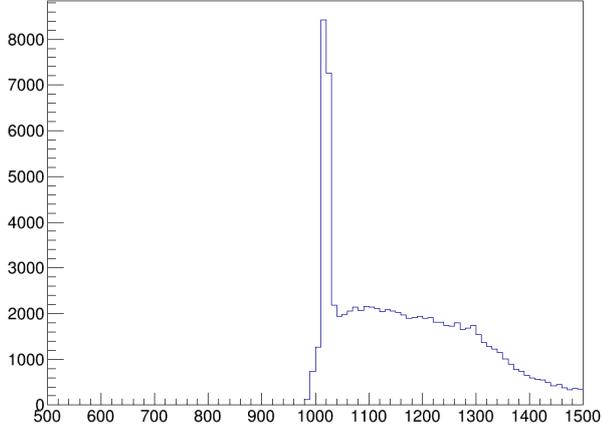
(b) The reconstructed mass of the  $KKee$  final state, showing the  $B_s$  peak as well as a second peak to the left which likely comes from partially reconstructed background.

The histogram in figure 10a displays a very pronounced peak at the  $B_s$  mass, while the plot in figure 10b shows the same peak with the addition of a second smaller peak, which is likely a sign of partially reconstructed background.

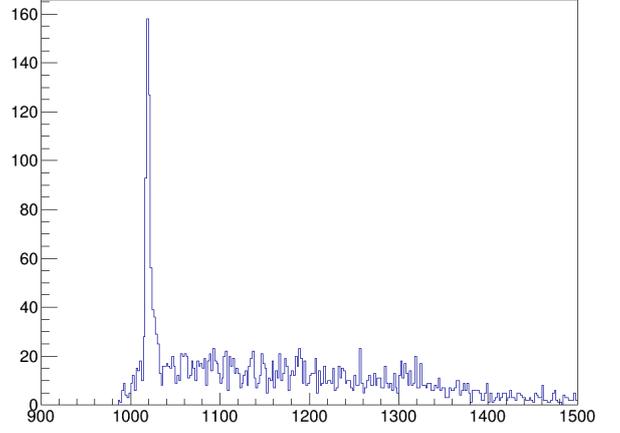
The presence of peaks at the  $B_s$  mass indicates that the dataset is contaminated by decays of  $B_s$  mesons into the final state  $KK\ell\ell$ . A similar plot using the hypothesised two-body final state  $KK$  in place of  $pK$ , shown in figure 11, produces a peak at the mass of the  $\phi$  meson, indicating that the background decay present here is  $B_s \rightarrow \phi(\rightarrow K^+K^-)J/\psi(\rightarrow \ell^+\ell^-)$ .

This background source was removed from the muon channel by vetoing all events for which  $m_{KK\ell\ell}$  lies within a region around the  $B_s$  peak and all events for which  $m_{KK}$  lies within a region around the  $\phi$  peak. In the electron channel, Bremsstrahlung broadening means that the  $B$  decay events cannot easily be distinguished from the signal peak. A veto was therefore only imposed on  $\phi \rightarrow KK$  events in this channel as these are negligibly affected by Bremsstrahlung losses.

Figure 11: The reconstructed initial state mass for a decay into the  $KK$  final state, showing well-defined peaks at the  $\phi$  mass,  $m_\phi = 1019 \text{ MeV}/c^2$ .



(a) The reconstructed  $KK$  mass in the muon channel.

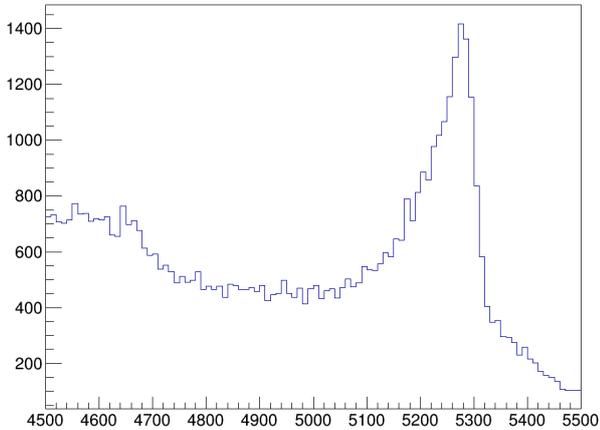


(b) The reconstructed  $KK$  mass in the electron channel.

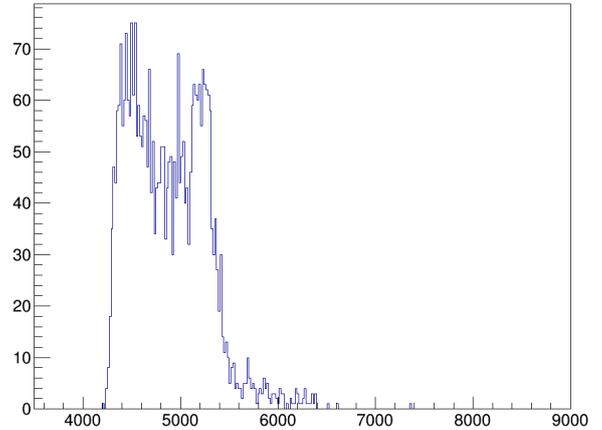
### 5.2.2 $\bar{B}^0 \rightarrow J/\psi(\rightarrow \ell^+\ell^-)K^*(\rightarrow K\pi)$

This background is produced when either the proton is misidentified as a pion or both the proton and the kaon are misidentified as a kaon and pion respectively. Both the singly and doubly misidentified cases were investigated and both produced similar peaks indicating the presence of a  $\bar{B}^0$  decay. The distributions of the singly misidentified  $K\pi\ell\ell$  reconstructed mass are shown in figure 12, while figure 13 shows the distributions of the singly misidentified  $K\pi$  reconstructed mass.

Figure 12: The reconstructed initial state mass after the final state has been hypothesised as  $K\pi\ell\ell$  in the muon and electron channels, showing peaks at the mass of the  $\bar{B}^0$  meson,  $m_{\bar{B}^0} = 5280 \text{ MeV}/c^2$ .



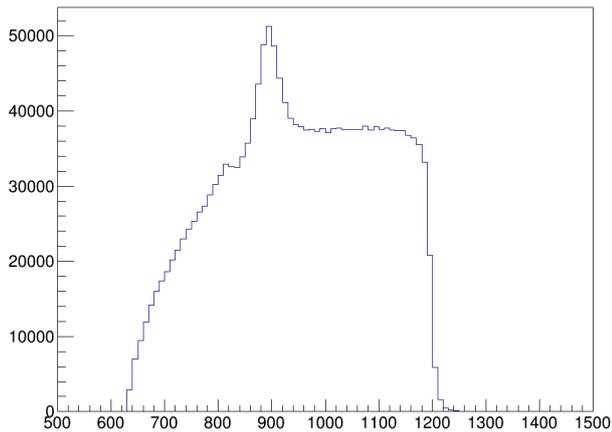
(a) Reconstructed mass of the  $K\pi\mu\mu$  final state.



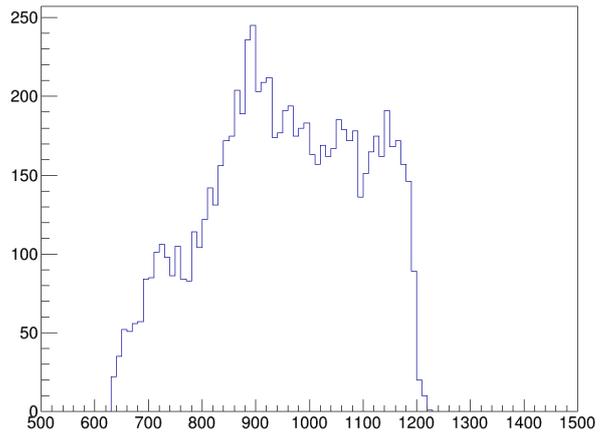
(b) Reconstructed mass of the  $K\pi ee$  final state. The peak at the  $\bar{B}^0$  mass is still visible here despite the comparably large peak to the left which is likely partially reconstructed background.

The large peak to the left of figure 12b is likely to be partially reconstructed background, but could also include traces of  $B_s \rightarrow \phi(\rightarrow K^+K^-)J/\psi(\rightarrow \ell^+\ell^-)$  decays that have been further misidentified as  $\bar{B}^0$  and have migrated left due to the  $K\pi\ell\ell$  final state used in this reconstruction having a lower mass than the  $KK\ell\ell$  final state used above.

Figure 13: The reconstructed initial state mass for a decay into the  $K\pi$  final state, showing peaks at the  $K^*$  mass,  $m_{K^*} = 891.7 \text{ MeV}/c^2$ .



(a) Reconstructed  $K\pi$  mass in the muon channel.



(b) Reconstructed  $K\pi$  mass in the electron channel.

As in section 5.2.1, vetoes are imposed on both the primary and secondary decays in the muon channel, but only on the secondary decays in the electron channel.

The  $\pi\pi$  final state, corresponding to background from  $\rho \rightarrow \pi\pi$  decays, was also investigated but no discernible peak could be seen at the  $\rho$  mass.

## 6 Multivariate Analysis

In order to remove the bulk of the combinatorial background, a multivariate analysis technique was used. Specifically, this consisted of a method known as a Boosted Decision Tree (BDT), built within the Toolkit for Multivariate Analysis (TMVA) software package [26]. The principles of the method are described in the following paragraphs according to Roe et al. [27].

The BDT process starts with two samples of data, one to represent signal and one to represent combinatorial background. This analysis uses simulated Monte Carlo data to mimic the  $\Lambda_b^0$  signal, while the combinatorial background is represented by samples of real data in the upper mass sidebands ( $m_{pK\mu\mu} > 5800 \text{ MeV}/c^2$ ,  $m_{pKee} > 6000 \text{ MeV}/c^2$ ), where the datasets are free from other sources of background. A higher upper sideband is used in the electron channel to accommodate the broader signal peak. Each sample is split into two parts; one part is used to train the BDT while the other is used to test the BDT's performance. This analysis uses a 4:1 training to testing sample size ratio.

In order to train, the BDT makes use of a list of variables that provide good separation between signal and background. The variables used in this analysis are listed below, and plots of these variables in the muon channel are shown in figure 14.

- $\log(K^- \chi_{\text{IP}}^2)$
- $\log(p \chi_{\text{IP}}^2)$
- $\log(\Lambda_b^0 \chi_{\text{IP}}^2)$
- $p \oplus K^- p_{\text{T}}$
- $\Lambda_b^0$  flight distance
- $\Lambda_b^0 \chi_{\text{vertex}}^2/\text{NDOF}$

NDOF stands for Number of Degrees Of Freedom, and is defined as the difference between the number of detector hits used to fit a particle track and the number of parameters used to make the fit.

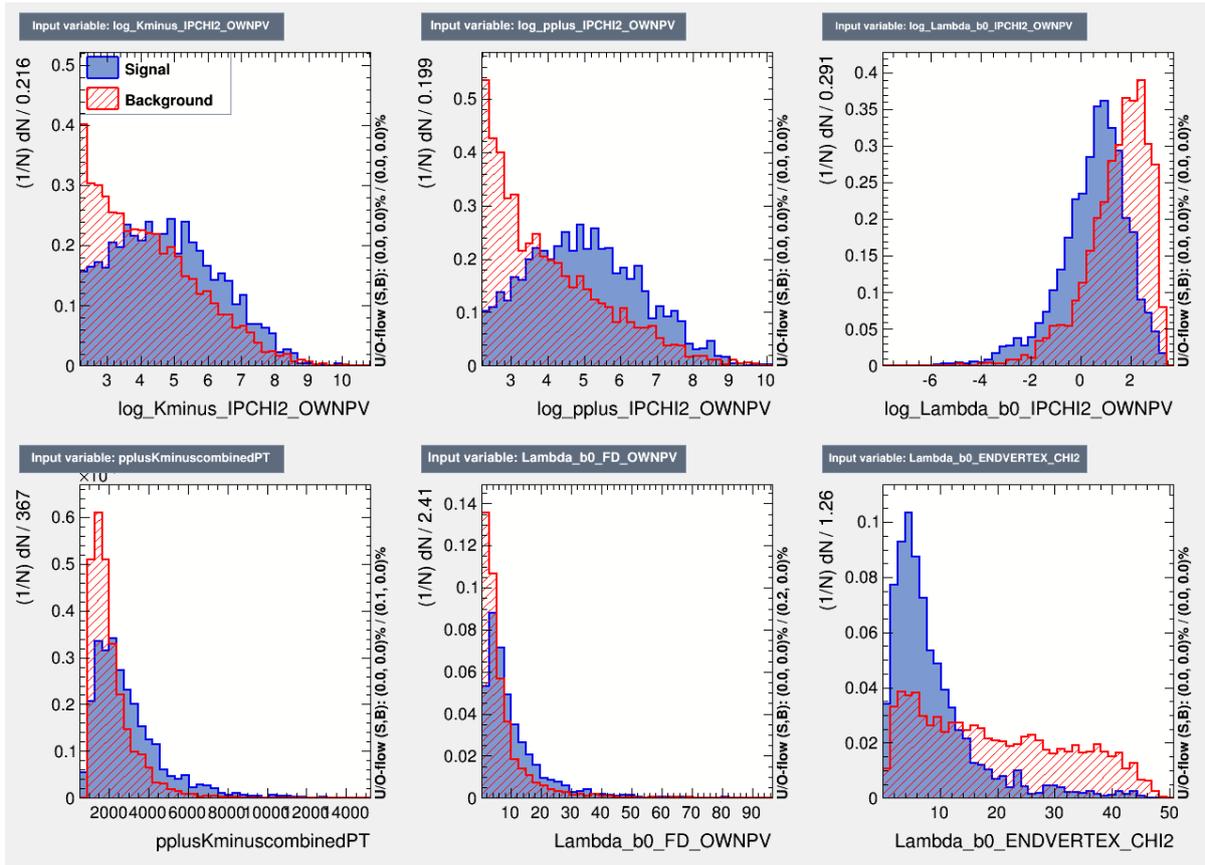


Figure 14: The variables used to train the BDT in the muon channel, showing the discriminating power they provide between signal and background.

A decision tree is created by taking the first variable in the list and sorting the events into two halves depending on the value of this variable. The splitting value is optimised by the algorithm as the value that produces the best separation of signal and background between the two halves. This is then repeated for each of the variables, such that the initial ‘node’ comprising the unsorted data is split into two ‘branches’ ideally each containing a majority of either signal or background events.

The classifier then performs the same process on the branch that is likely to most improve the separation, and so on. Once a branch is considered to contain mostly signal or mostly background events, it becomes known as a signal or background ‘leaf’ and is exempt from further splitting. The process stops either when a specified number of leaves have been obtained or when every leaf is pure signal or pure background, or has too few events to be split further. A schematic of the decision tree process is shown in figure 15.

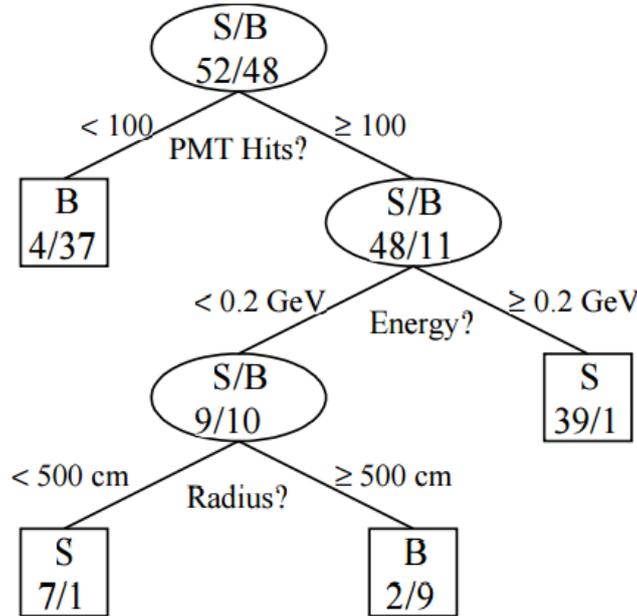


Figure 15: A schematic of the decision tree-making process with some example variable names and splitting values [27]. Leaves are marked with squares, and the numbers of signal and background events are indicated by S and B respectively.

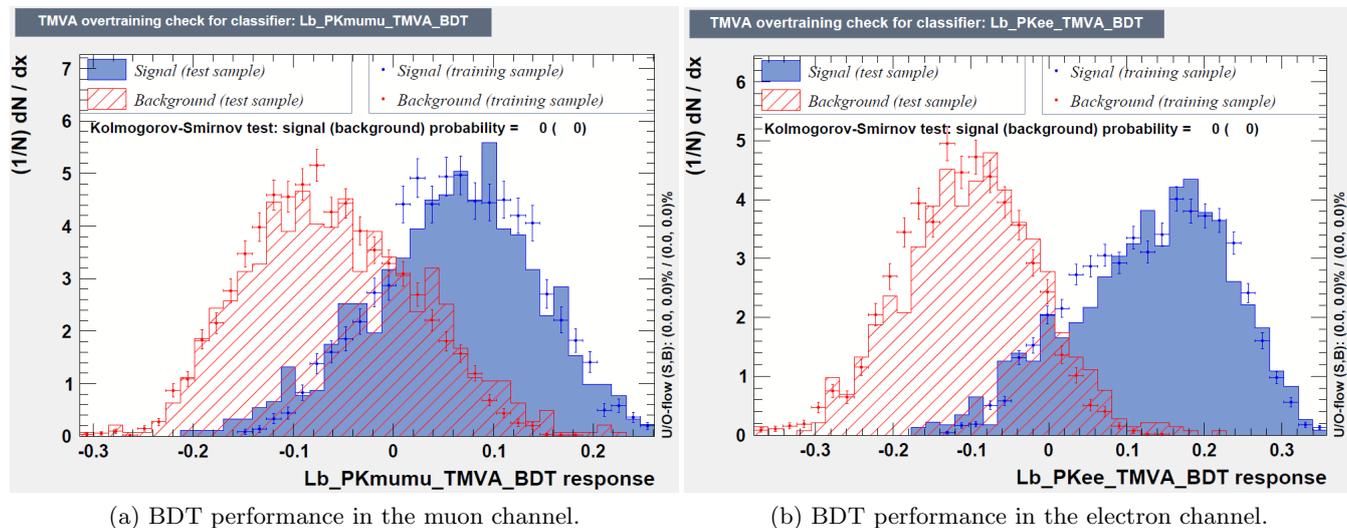
The process of boosting adds weights to the events in order to dramatically improve the accuracy of the tree. A decision tree is first created as above using unweighted events. If a signal or background event is misclassified and ends up on a leaf of the opposing type, the weight of the event is increased. Another tree is then made using the new boosted weights and the process is repeated until many trees have been created to form a 'forest'. Each event is tracked across all the trees and given a score of 1 if it lands on a signal leaf and -1 if it lands on a background leaf. If the normalised sum of these scores is high, the event is likely to be signal, while low scores indicate background events. By selecting events with scores above a certain value, a sample of mostly signal events can be obtained.

In order to provide a reliable separation between signal and background, the BDT requires that the sample of data used for selection be statistically independent from the sample used for training and testing. To accomplish this, only 50% of the initial data sample was used for training and testing, while the selection was performed on the other 50%. The 50% used for training and testing was not used in further analysis.

In the electron channel, LHCb has a lower detection efficiency than in the muon channel, so to preserve the number of events the entire sample was used for selection. To maintain statistical independence, two BDTs were created. One was trained on the first 50% of the data and events were selected from the second 50%, while the other was trained on the second 50% and events were selected from the first half.

The performance of the BDT in the muon and electron channels (just one of the electron channel BDTs is shown for illustration purposes) is illustrated in figure 16, which shows the separation between high-scoring signal events and low-scoring background events.

Figure 16: Plots showing the performance of the BDT in the muon and electron channels.



The plots in figure 16 show the separation achieved by the BDT between signal and background in both the training and test samples. The slight discrepancy between the two samples in each channel indicates a small amount of overtraining, which occurs when the BDT is not given enough data points to accommodate the number of variables it uses for training. A small amount of overtraining is to be expected given the reduced sizes of the datasets used to develop the selection, but it is not sufficient to invalidate the choice of training variables listed above.

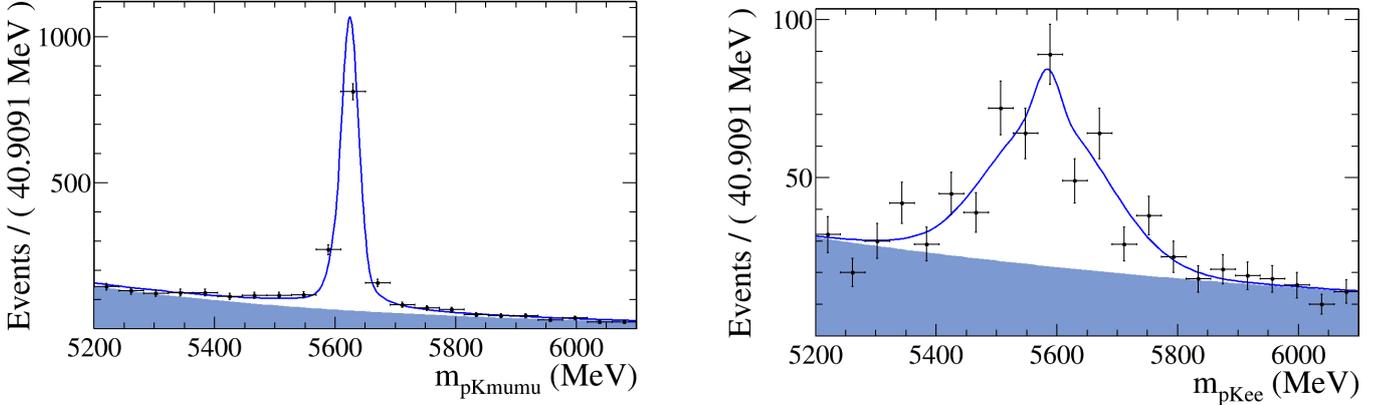
As suggested by figure 16, events were selected in both channels by requiring that the BDT response variable be greater than zero, however this is still to be refined.

## 7 Fitting

### 7.1 Use of Probability Density Functions

In order to assess the performance of the selection, and eventually to calculate the number of  $\Lambda_b^0$  events, probability density functions (PDFs) designed to approximate the signal and background distributions are fitted to the selected events. The PDFs used here comprise two Crystal Ball functions (CBFs) for the signal and an exponential decay to approximate the remaining combinatorial background. A CBF is a Gaussian distribution with a power-law tail on one side, and the two CBFs used in each PDF are constructed with tails on opposite sides. Fits for the muon and electron control channels are shown in figure 17.

Figure 17: Fits to the signal (blue line) and background (blue shading) distributions in the muon and electron control channels after all selection has been performed.



(a) The signal and background fits in the muon channel, showing a very clean peak at the  $\Lambda_b^0$  mass as expected.

(b) The fits in the electron channel show clearly the loss in  $\Lambda_b^0$  mass resolution caused by Bremsstrahlung effects and corrections.

The fits in figure 17 show that the selection described in the previous section is effective, especially in the muon channel. The electron channel is much less clean, due mainly to the effects of Bremsstrahlung radiation and the imperfect corrections described in section 4.3.1, but also due to the generally lower number of electron events detected by LHCb. It can also be seen by eye that the signal to background ratio is much lower in figure 17b, illustrating that refinement of the selection in this channel is still ongoing – for example the HOP tool is still to be added.

## 7.2 Determination of Signal Yield

In order to determine how many signal and background events are present in the final sample, the PDFs and exponential functions in figure 17 are integrated over a signal region defined around each peak. A much wider signal region is clearly required in the electron channel. Integrating each PDF gives the total number of events in each channel, so to compute the signal yield the integral of the respective exponential function (i.e. the area of the shaded regions in figure 17) must be subtracted.

This process is still ongoing and remains highly rudimentary until new stripping lines are available and the selection has been completely refined, especially in the electron channel. However, the discrepancy between the numbers of muon and electron events is roughly consistent with the detection efficiencies of LHCb.

## 8 Systematic Errors and Efficiencies

Using the control channel as described in section 4.2 causes many of the errors and selection efficiencies to cancel out. However, some kinematic differences between the control channels and rare modes may remain, leading to a residual systematic error in the calculation of  $R_{pK}$ . This section summarises the main sources of error in this analysis and discusses how the selection efficiencies will be calculated. The drop in efficiency when detecting electrons compared to muons must also be taken into account.

The overall selection efficiency of the signal events depends on the detector's geometric acceptance, the trigger efficiencies and the efficiency of the selection developed for this analysis. The contributions of each of these will be calculated using simulated data, but the efficiency of LHCb's geometric acceptance for a wide range of decays is already well studied.

The trigger efficiencies will be determined specifically for  $\Lambda_b^0$  decays, but it is known that for  $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$  the efficiencies of L0, Hlt1 and Hlt2 are 89%, 92% and 87% respectively [16]. A calculation of the trigger efficiencies using simulated data is expected to contribute the most significant component of the overall systematic uncertainty.

The intrinsic particle identification efficiency of LHCb is around 90% for both electrons and kaons, with a mis-ID probability of 5%. For muons the efficiency is 97%, with a mis-id probability of 1–3% [28]. The efficiencies of the PID selection used in this analysis will be calculated using simulated data, as will the efficiencies of the other selection criteria. The use of simulated data for this calculation will add to the overall systematic uncertainty.

An efficiency will also be calculated for the dilepton invariant mass cut used in the control channel. This cut has its own intrinsic error in the electron channel as Bremsstrahlung under- and over-corrections can cause events to migrate into or out of the  $q^2$  window. This process is not expected to be accurately modelled by simulated data and therefore contributes a further uncertainty.

In the muon channel, the choice of signal model has only a very small effect on the signal yield and therefore contributes a negligible error [14]. In the electron channel, however, the discrepancy between the control channel and the rare mode introduces a larger systematic error to the fit. The magnitude of this error can be determined by investigating the difference between simulated distributions of control channel and rare mode events.

## 9 Conclusion

Using small test samples of data collected by LHCb in 2011 and 2012, a procedure has been developed to select events in the control channel,  $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \ell^+ \ell^-)$ , where  $\ell = \mu, e$ . Fits have been performed on muon and electron channel data, showing a discrepancy between the numbers of muon and electron events that is broadly consistent with the detection efficiencies of LHCb.

Unforeseen delays in the acquisition of new stripping lines have slowed the overall progress of the analysis, and selection is still ongoing. New data is expected in the coming weeks, after which the efficiencies will be calculated and it is expected that the selection procedure developed here will be verified with a measurement of

$$\frac{\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-))}{\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow e^+ e^-))} = 1 \pm \mathcal{O}(10^{-3})$$

At this point, the  $q^2$  requirement imposed on the data will be moved away from the  $J/\psi$  resonance region to focus on the theoretically favoured rare mode region  $1 < q^2 < 6 \text{ GeV}/c^2$ , leading to the first ever observation of the decay  $\Lambda_b^0 \rightarrow pK^- e^+ e^-$ . The size of the data samples will be extended to make use of the full  $3.0 \text{ fb}^{-1}$  of data collected in 2011 and 2012, and it is expected that the same selection procedure will remain valid in the rare mode region.

Selection of rare mode events will allow for the eventual first ever measurement of  $R_{pK} = \mathcal{B}(\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-) / \mathcal{B}(\Lambda_b^0 \rightarrow pK^- e^+ e^-)$ . This process is expected to start within the coming weeks.

## References

- [1] I. S. Hughes, *Elementary Particles*, Cambridge University Press, Cambridge, 1991
- [2] K. A. Olive et al. (Particle Data Group), *Chin. Phys. C*, 38, 090001 (2014) and 2015 update
- [3] BaBar Collaboration, B. Aubert et al., *Evidence for the Flavor Changing Neutral Current Decays  $B \rightarrow K\ell^+\ell^-$  and  $B \rightarrow K^*\ell^+\ell^-$* , 2002, [arXiv:hep-ex/0207082](#)
- [4] C. Bobeth et al., *Angular Distributions of  $B \rightarrow K\ell\ell$  Decays*, *JHEP* 0712:040, 2007, [arXiv:0709.4174v2](#) [[hep-ph](#)]
- [5] A. Ali et al., *Improved Model-Independent Analysis of Semileptonic and Radiative Rare B Decays*, 2002, [arXiv:hep-ph/0112300v2](#)
- [6] G. Hiller and F. Krüger, *More Model-Independent Analysis of  $b \rightarrow s$  Processes*, *Phys. Rev. D*. 69:074020, 2004, [arXiv:hep-ph/0310219v2](#)
- [7] C. Bobeth et al., *Analysis of neutral Higgs-boson contributions to the decays  $\bar{B}_s \rightarrow l^+l^-$  and  $\bar{B} \rightarrow Kl^+l^-$* , *Phys. Rev. D* 64, 074014, 2001, doi:10.1103/PhysRevD.64.074014
- [8] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events*, Cambridge University Press, Cambridge, 2006
- [9] H. E. Logan and U. Nierste,  *$B_{s,d} \rightarrow \ell^+\ell^-$  in a Two-Higgs-Doublet Model*, *Nucl. Phys.* B586 (2000) 39-55, [arXiv:hep-ph/0004139v2](#)
- [10] C.-S. Huang and X.-H. Wu,  *$B_s \rightarrow \mu^+\mu^-$  and  $B \rightarrow X_s\mu^+\mu^-$  in MSSM*, *Nucl. Phys.* B657 (2003) 304-332, [arXiv:hep-ph/0212220v2](#)
- [11] Y. G. Kim et al., *Possible New Physics Signals in  $b \rightarrow s\gamma$  and  $b \rightarrow sl^+l^-$* , *Nucl. Phys.* B544 (1999) 64-88, [arXiv:hep-ph/9810336v2](#)
- [12] Belle Collaboration, J.-T. Wei et al., *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for  $B \rightarrow K^{(*)}\ell^+\ell^-$* , *Phys. Rev. Lett.* 103:171801, 2009, [arXiv:0904.0770v2](#) [[hep-ex](#)]
- [13] BaBar Collaboration, J. P. Lees et al., *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays  $B \rightarrow K^{(*)}\ell^+\ell^-$* , *Phys. Rev. D* 86, 032012, 2012, [arXiv:1204.3933v2](#) [[hep-ex](#)]
- [14] LHCb Collaboration, R. Aaij et al., *Test of lepton universality using  $B^+ \rightarrow K^+\ell^+\ell^-$  decays*, *Phys. Rev. Lett.* 113, 151601 (2014), [arXiv:1406.6482](#) [[hep-ex](#)]
- [15] LHCb Collaboration, R. Aaij et al., *Measurement of the ratio of branching fractions  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , *Phys. Rev. Lett.* 115, 111803 (2015), [arXiv:1506.08614](#) [[hep-ex](#)]
- [16] LHCb Collaboration, R. Aaij et al., *LHCb Detector Performance*, *Int. J. Mod. Phys. A* 30, 1530022 (2015), [arXiv:1412.6352](#) [[hep-ex](#)]
- [17] LHCb Collaboration, A. A. Alves et al., *The LHCb Detector at the LHC*, *JINST* 3 (2008) S08005, doi:10.1088/1748-0221/3/08/S08005
- [18] N. Skidmore, *Measurement of  $R_K = \mathcal{B}(B^+ \rightarrow K^+\mu^+\mu^-)/\mathcal{B}(B^+ \rightarrow K^+e^+e^-)$  as a Test of Lepton Universality at LHCb*, Master's thesis, University of Southampton, May 2013
- [19] 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](#) [[physics.ins-det](#)]

- [20] C. W. Fabjan and F. Gianotti, *Calorimetry for particle physics*, Rev. Mod. Phys. 75, 1243 (2003), doi:10.1103/RevModPhys.75.1243
- [21] F. Archilli et al., *Performance of the Muon Identification at LHCb*, JINST 8 (2013) P10020, arXiv:1306.0249 [physics.ins-det]
- [22] R. Aaij et al., *The LHCb Trigger and its Performance in 2011, 2013*, arXiv:1211.3055v2 [hep-ex]
- [23] LHCb Collaboration, R. Aaij et al., *Study of the production of  $\Lambda_b^0$  and  $\bar{B}^0$  hadrons in pp collisions and first measurement of the  $\Lambda_b^0 \rightarrow J/\psi p K^-$  branching fraction*, Chinese Phys. C, 40, 011001 (2016), arXiv:1509.00292
- [24] M. Borsato et al., *HOP: an additional tool for decays involving electrons.*, 2015, LHCb Internal Publication LHCb-INT-2015-037
- [25] L. Pescatore, *Testing lepton flavour universality in rare decays at the LHCb experiment*, presented at the Joint Annual HEPP and APP Conference, Brighton, UK, 21st March 2016
- [26] A. Hoecker et al., *TMVA: Toolkit for Multivariate Data Analysis*, PoS A CAT 040 (2007), arXiv:physics/0703039
- [27] B. P. Roe et al., *Boosted Decision Trees as an Alternative to Artificial Neural Networks for Particle Identification*, Nucl. Instrum. Meth. A543 (2005) 577-584, arXiv:physics/0408124v2 [physics.data-an]
- [28] LHCb Speakers' Bureau, *LHCb Performance Numbers*, 2015, <http://lhcb.web.cern.ch/lhcb/speakersbureau/html/PerformanceNumbers.html>, accessed 9th May 2016